LEXICOGRAPHICA

International Annual for Lexicography
Revue Internationale de Lexicographie
Internationales Jahrbuch für Lexikographie

Edited by
Ulrich Heid, Stefan Schierholz,
Wolfgang Schweickard, Herbert Ernst Wiegand
and Werner Wolski

in conjunction with the
Dictionary Society of North America (DSNA)
and the
European Association for Lexicography (EURALEX)

24/2008

Sonderdruck

Max Niemeyer Verlag
Tübingen

Zita Hollós

Kollokationen und weitere typische Mehrwortverbindungen in der ungarischen Lexikographie

0	Einleitung		Nutzung großer Sprachkorpora
1	Kollokationen in der ungarischen Lexikographie	5	Mikrostruktur
		6	Resümee
2	Potenzielle Benutzer von KOLLEX	7	Literatur
3	Bestimmung des Wörterbuchtyps von KOLLEX	7.1	Korpora
		7.2	Wörterbücher
4	Wörterbuchbasis und effiziente	7.3	Sekundärliteratur in Auswahl

Abstract
Collocations, related to Hungarian or Hungarian and other languages, are rarely discussed as a theoretical problem. Some studies contrasting Hungarian with German focus on the didactic aspects of collocations, especially on the learning process and language production, because collocations are basically problematical language entities in both cases. They are difficult to produce in a foreign language, but relative easily understood in the mother tongue, although students are mostly not able to notice their own mistakes by speaking or writing. Some new Hungarian-German translation dictionaries in both directions do not only specify and define collocations, but furthermore work with a concept of interlingual collocations. Now the idea of a collocations dictionary in its widest sense appeared in Hungary in relation to the German language. The following article introduces a learner-friendly, data- and corpusbased German-Hungarian dictionary with two word lists, first of all for the language production: shortly called KOLLEX. The conception and the main characteristics of this bilingual dictionary will be demonstrated on the basis of dictionary typology. The presentation of the dictionary type from different significant aspects and diverse concrete examples of dictionary articles provide a direct insight into a completely new collocations dictionary for advanced learners of German as a foreign language, viz. into KOLLEX.

0 Einleitung

Da die Kollokationsforschung im deutschen Sprachgebiet auf eine nicht allzu lange Vergangenheit zurückblicken kann, ist es nicht verwunderlich, dass auch für das Sprachenpaar Deutsch-Ungarisch bisher kein Spezialwörterbuch für Kollokationen existiert. Sowohl die ungarische Lexikographie – als kulturelle Praxis im Sinne von Wiegand – wie die ungarische Wörterbuchforschung befanden sich im Vergleich zum Stand Westeuropas in den letzten Jahrzehnten stets im Hintertreffen, so auch gegenüber Deutschland, deshalb kann das Fehlen eines einsprachigen ungarischen Kollokationswörterbuchs (oder breiter gefasst: eines syntagmatischen Wörterbuchs) oder aber überhaupt eines zweisprachigen mit Ungarisch nicht überraschen, wenn auch in den letzten Jahren mehrere neu konzipierte und moderne Sprachwörterbücher auf dem ungarischen Markt erschienen. Nach dem offiziellen Stand

DOI 10.1515/lexi.2008.010

werden lexikographische Nachschlagewerke, die die lexikalische Syntagmatik als ihren Gegenstand gewählt haben, in Ungarn zur Zeit nicht erarbeitet, und solche Unternehmungen sind sowieso vornehmlich einsprachig ausgerichtet.[1] Damit stellt die Konzeption und Verwirklichung von KOLLEX (ung. SZÓkapTÁR), eines korpusbasierten, produktionsbezogenen deutsch-ungarischen syntagmatischen Lernerwörterbuchs in der zweisprachigen Lexikographie eine echte Herausforderung dar. Im vorliegenden Beitrag soll über diesen lexikographischen Prozess, der seit 2004 mit fünf wissenschaftlichen Hilfskräften und der finanziellen Unterstützung des ungarischen Nationalfonds für Wissenschaft (OTKA) umgesetzt wird, mit folgenden Themenschwerpunkten berichtet werden:

- Potenzielle Benutzer
- Bestimmung des Wörterbuchtyps von KOLLEX: nach Benutzerbezug, Wörterbuchgegenstand, Wörterbuchform und nach trägermedialen Eigenschaften
- Wörterbuchbasis und effiziente Nutzung großer Sprachkorpora
- Mikrostruktur

Bevor auf den lexikographischen Prozess als solchen näher eingegangen wird, soll das erwähnte Wörterbuchprojekt in einem ersten orientierenden Schritt in den Horizont des ungarischen lexikographischen Kontextes eingeordnet werden.

1 Kollokationen in der ungarischen Lexikographie

In der ungarischen Wörterbuchlandschaft gibt es zur Zeit drei generisch engstens verwandte zweisprachige Wörterbücher, die den Begriff „Kollokation" in ihrer jeweiligen Einleitung nicht nur erwähnen, sondern auch anwenden, nämlich das von R. Hessky: *Német-magyar kéziszótár* (=HESS-NM ^{3}2005, Deutsch-ungarisches Handwörterbuch), *Magyar-német kéziszótár* (=HESS-MN ^{2}2005, Ungarisch-deutsches Handwörterbuch) und das auf deren Grundlage erstellte R. Hessky-B. Iker: *Német-magyar, magyar-német tanulószótár* (=HESSIKE 2005, Deutsch-ungarisches, ungarisch-deutsches Lernerwörterbuch). Unter Kollokation verstehen die Autoren des ersten Wörterbuchs

> (...) typische, für die deutsche Sprache charakteristische Wortverbindungen, deren Übersetzung mit der ungarischen Entsprechung, die unter der entsprechenden Bedeutung angegeben wird, nicht möglich ist oder es ergäbe sich eine ungewöhnliche Lösung.[2]

Diese Auffassung kommt auch bei den oben genannten Nachfolgewörterbüchern zur Anwendung. Kollokationen werden infolgedessen von den Beispielsätzen mit dem Symbol Doppellvirgel „||" getrennt und sogar typographisch, fett kursiv, hervorgehoben. Der in diesem zweisprachigen Wörterbuch praktizierte Kollokationsbegriff ist kontrastiv, und wird in Hollós 2004 übernommen und *interlinguale* Kollokation genannt, unabhängig davon, ob

---

[1] Wie z.B. auf dem englischen Sprachgebiet das erste neuere einsprachige Kollokationswörterbuch OCDSE (2002), dessen Vorläufer das zu seiner Zeit beispielhafte BBI (1989) war.

[2] Übersetzt nach HESS-NM ^{3}2005, X.

die Kollokation noch einen *intralingualen* Status besitzt oder nicht. In diesem Sinne wird auch im KOLLEX neben einem interlingualen Kollokationsbegriff auch ein intralingualer angewandt. Dem intralingualen Status einer Kollokation, also Untersuchung, Nachweis und Kodifizierung des beschränkten Kollokationspotenzials mit Hilfe von Korpora, das wichtigste Spezifikum von KOLLEX, wurde meines Wissens in früheren ein- und zweisprachigen Wörterbuchprojekten bislang noch nicht Rechnung getragen, wie auch die innere Selektion aufgrund von Korpusstatistiken als lexikographische Zwischendatei ein recht neuer methodischer Weg ist – vor allem in der zweisprachigen Lexikographie.

Erfreulicher Weise ist die Beschäftigung mit Kollokationen im Bereich der Fremdsprachendidaktik in Ungarn auf dem Vormarsch (vgl. Réder 2006), was immerhin ein stärkeres Interesse für dieses Phänomen beweist, wie auch der erste, vom Verein ungarischer angewandter Linguisten organisierte runde Tisch zu diesem Thema im Herbst 2006.[3]

2 Potenzielle Benutzer von KOLLEX

Die Zielgruppe, also die Adressaten des Wörterbuchs, sind in erster Linie fortgeschrittene Deutschlerner, die sich den Grundwortschatz „Zertifikat Deutsch" (ZD) von 2600 Lexemen bereits angeeignet haben, genauer gesagt: die die mehreren tausend dazugehörigen Einzelbedeutungen/Sememe aktiv beherrschen. Darüber hinaus verfügen sie über die wichtigsten Verknüpfungsmöglichkeiten der Sememe, sowohl auf der grammatischen Ebene in Form von konkreten Valenzrealisierungen – beruhend auf den Valenzeigenschaften des jeweiligen Lexems – als auch auf der lexikalischen Ebene in Form von usuellen Syntagmen, wie z.B. freie Wortkombinationen und Kollokationen.

3 Bestimmung des Wörterbuchtyps von KOLLEX

Das Endprodukt wird folgende Wörterbuchtypen nach den 4 Klassen der Wörterbuchtypologie repräsentieren:

a. nach dem Benutzerbezug: Produktionswörterbuch
b. nach dem Wörterbuchgegenstand: synchronisches polyselektives syntagmatisches Spezialwörterbuch
c. nach der Wörterbuchform: bialphabetisches Wörterbuch mit Register
d. nach trägermedialen Eigenschaften: sowohl Printwörterbuch als auch digitales Wörterbuch

Im Nachfolgenden wird der konkrete Wörterbuchtyp von KOLLEX näher beschrieben:

---

[3] Die ungarische Fassung der Beiträge und die nachfolgende Diskussion sind auf der folgenden Internetseite zu finden: http://www.manye.pte.hu/Kollokaciok.rtf

a. Produktionswörterbuch

Diese Zuordnung erfolgt anhand von typischen Benutzungssituationen, die für die potenziellen Benutzer des Wörterbuchs charakteristisch sind. Der wichtigste anvisierte Benutzungsanlass bei KOLLEX ist die gestörte Produktion von fremdsprachlichen Texten, und zwar in Produktionssituationen mit einer lexikalischen Lücke oder mit Zweifeln hinsichtlich der Kombinierbarkeit der Wörter.[4] Im ersten Fall der lexikalischen Lücke lautet eine konkrete Suchfrage beispielsweise folgendermaßen: „Wie sagt man auf Deutsch *buja* (*saftig*) in Verbindung mit *grün*?" und im zweiten Fall, wenn der Benutzer Korrektheitszweifel hat: „Kann man *eine Fahne aufziehen* wie *eine Uhr*?", da im Ungarischen das Verb *felhúz* (*aufziehen*) ein entsprechendes zusätzliches Semem hat, weshalb beide Objekte in Verbindung mit ein und demselben Verb möglich sind.

In beiden Fällen, sei es Ausdrucksfindungsschwierigkeit oder Ausdruckswahlunsicherheit, ist ein Bestandteil der Kombination bekannt, nämlich *grün* bzw. *Fahne*. Sie stellen den eigentlichen Kern der Aussage dar, worüber etwas gesagt wird oder der modifiziert werden kann. Ausgangspunkt sind also die bereits erworbenen Wortschatzelemente, und diese Tatsache begründet auch die Sprachrichtung deutsch-ungarisch, da auf dem Sprachniveau des anvisierten Benutzerkreises die Sprachproduktionsprozesse idealerweise vornehmlich in der Fremdsprache ablaufen sollten. Das zu erarbeitende Wörterbuch wird dementsprechend in Konsultationssituationen wegen lexikonbedingter Ausdrucksfindungsschwierigkeit oder Ausdruckswahlunsicherheit gute Dienste leisten, da man unter dem Lemmazeichen **grün** *saftig*, unter **Fahne** *hissen* finden wird.

Ein weiterer wichtiger Grund für die Wahl dieser Sprachrichtung besteht darin, dass das Wörterbuch den deutschen usuellen Sprachgebrauch zu dokumentieren hat, was nur mit der Einbeziehung realer, normal schriftsprachlicher Texte, also alleine durch die Verwendung deutscher elektronisch gespeicherter und statistisch auswertbarer großer Sprachkorpora zu leisten ist.

Es seien hier noch einige weitere typische Beispiele für Ausdrucksfindungsschwierigkeiten erwähnt:[5] *utolérhetetlenül* ***olcsó*** (.?. *billig – konkurrenzlos* ***billig***), *bársonyosan* ***puha*** (.?. *weich – samtig* ***weich***) oder *fogas* ***kérdés*** (.?. *Frage – eine knifflige* ***Frage***), *beugratós* ***kérdés*** (.?. *Frage – eine verfängliche* ***Frage***), aber bei der folgenden Kombination ist Vorsicht geboten: *találós* ***kérdés*** (≠ .?. *Frage – Rätsel* !).

Konkrete Beispiele aus Seminaren für Ausdrucksfindungsschwierigkeiten sind:

?*zusammengesetztes* ***Thema*** – richtig: *komplexes Thema* oder ***Interesse*** ?*erwecken/aufwecken* – korrekt: *Interesse wecken.*

Gerade zur Beantwortung der obigen Suchfragen ist also ein syntagmatisches Lernerwörterbuch dadurch hervorragend geeignet, dass der Benutzer-in-actu unter dem bekannten Wort mit großer Wahrscheinlichkeit das fehlende treffende Wort findet.

---

4 Zur Typologie der Wörterbuchbenutzungssituationen vgl. Wiegand 1998a.

5 Fett hervorgehoben werden die bekannten fremdsprachigen Bestandteile (Basen) der jeweiligen ungarischen oder deutschen Kombination.

b. Synchronisches polyselektives syntagmatisches Spezialwörterbuch

Den Wörterbuchgegenstandsbereich stellt die deutsche Gegenwartssprache dar, untersucht und abgebildet werden die syntagmatischen Relationen der Lemmazeichen, insbesondere auf der lexikalischen Ebene. Der Wörterbuchgegenstand ist demzufolge die Syntagmatik der ausgewählten Substantive, Adjektive, Verben und weniger Adverbien des „Zertifikat Deutsch"-Wortschatzes.[6] Das Wörterbuch ist polyselektiv, weil 4 Klassen von Lemmazeichen berücksichtigt werden, nämlich diejenigen, die als Basis einer Kollokation fungieren können. An dieser Stelle ist die Erwähnung der von Hausmann (1984) geprägten Kollokationsdefinition unerlässlich:

> Wörter mit begrenzter Kombinierbarkeit verbinden sich entsprechend differenzierter semantischer Regeln und einer gewissen zusätzlichen Üblichkeit mit Wörtern, zu denen sie in Affinität stehen. Affinität sei definiert als die ‚Neigung zweier Wörter, kombiniert aufzutreten'.[7] (Hausmann 1984, 398)

Laut Hausmann haben die Bestandteile einer Kollokation unterschiedlichen Status: den semantisch autonomen Teil nennt er *Basis*, die zur Wortklasse Substantiv, Adjektiv, Verb oder eventuell Adverb gehört, und den dazu affinen Teil *Kollokator*. Die beiden zusammen bilden eine Kollokation nach folgenden 5 Strukturtypen:[8]

Substantiv + Substantiv	z.B. *ein Fleckchen* ***Gras***
Substantiv + Adjektiv	z.B. *ohrenbetäubender* ***Lärm***
Substantiv + Verb	z.B. ***Bücher*** *verschlingen*
Verb + Adverb	z.B. *verschmitzt* ***lächeln***
Adjektiv + Adverb	z.B. *saftig* ***grün***

Außerdem hält das KOLLEX für größere, jedoch nicht idiomatische Einheiten einen eigenen Angabetyp bereit, nämlich den Angabetyp der Kombinationen, im Wörterbuchartikel mit dem Angabesymbol „<KOMB>" gekennzeichnet. Kombinationen stellen entweder mögliche Kombinationen der obigen Strukturtypen dar (z.B. Substantiv + Adverb + Adjektiv: *häufig gestellte* ***Frage***), oder setzen sich aus einem durch eine adjektivische/adverbiale Wortgruppe realisierten Kollokator und einer Basis zusammen (z.B. adverbiale Wortgruppe + Adjektiv: *zum Verwechseln* ***ähnlich***). Zur Bestimmung und Typologisierung weiterer Typen und Untertypen soll das bereits redigierte Material eingehend untersucht werden.

Anhand des erläuterten Produktionsbezuges ist einleuchtend, dass die Kollokationen im KOLLEX unter der meistens bereits bekannten Basis aufgeführt werden, denn

> jegliches Formulieren geht vom Substantiv zum Adjektiv oder Verb, nicht umgekehrt [...] Nur bei der Textrezeption reichen somit die üblichen einsprachigen Wörterbücher aus. Bei der Textproduktion müssen sie jämmerlich versagen. (Hausmann 1984, 401–402)

Die Kürze des Beitrages erlaubt nicht, diese Behauptung empirisch zu untermauern, aber jeder weiß aus eigener Erfahrung von ihrer Richtigkeit. Sogar in den großen deutschen

---

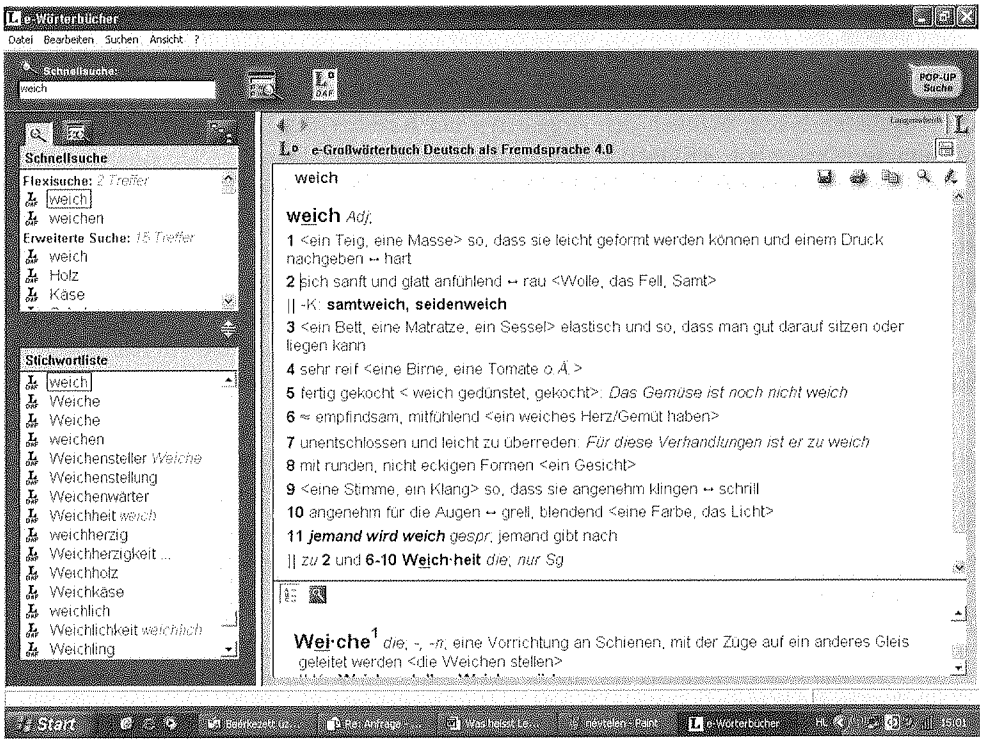
[6] Zur Zeit enthält das Wörterbuch 2284 Lemmata. Die kleinere Hälfte des Lemmabestandes machen die Verb- und Adjektivlemmata aus, die zu ihnen gehörigen Wörterbuchartikel sind bereits vollständig erarbeitet und wurden einmal redigiert und lektoriert.

[7] Auf diesen letzten Teil der Definition wird noch bei der Behandlung der Wörterbuchbasis im Kap. 4 eingegangen.

[8] Fett kursiv hervorgehoben sind jeweils die Basen.

Lernerwörterbüchern von Langenscheidt (LGWDaF 2003), de Gruyter (DGWDaF 2000), Pons (PGWDaF 2004) oder Duden (DDaFSTW 2002) sind die meisten Kollokationen unter dem Kollokator verzeichnet, also unter *dem* Bestandteil, der meistens in der Produktionssituation unbekannt ist, wodurch die Kollokationen nur in der Rezeptionssituation auffindbar sind. Darüber hinaus listet der Kollokatorartikel nur einige mögliche Basen auf, nicht jedoch alle möglichen und/oder typischen. Neu sind allerdings die Erkenntnisse, die sich im Rahmen des Wörterbuchprojekts bei der Volltextsuche im LGWDaF ergeben haben, dass nämlich:

1. unter den adverbial gebrauchten Adjektiv-Kollokatoren bzw. Adverb-Kollokatoren als Lemmazeichen meist keine Verb- oder Adjektiv-Basen zu finden sind: z.B. unter dem Lemmazeichen *glimpflich* ist die Basis *ausgehen* gar nicht verzeichnet oder unter *seidig* finden wir kein *weich*,
2. zu den meisten Verb- oder Adjektiv-Basen, die gleichzeitig Lemmata im KollexX sind, auch die Volltextsuche (sehr) wenige adverbial gebrauchte Adjektiv-Kollokatoren bzw. Adverb-Kollokatoren verzeichnet – hier besteht also ein großer Nachholbedarf,
3. auch unter der modifizierbaren verbalen oder adjektivischen Basis der Kollokation, wie z.B. *ausgehen* oder *weich*, man im dazugehörigen Wörterbuchartikel meist sehr wenige oder überhaupt keine adverbial gebrauchten Adjektive oder Adverbien findet. Wie die Basis *weich* modifiziert werden kann, wird aus dem folgenden Wörterbuchartikel – wie aus vielen ähnlichen – nicht ersichtlich:



1. Abbildung Wörterbuchartikel zu *weich* aus der CD-ROM-Version von LGWDaF 2003

Das sind weitere Argumente für die Notwendigkeit der Korpusarbeit und spricht für die systematische und effiziente Nutzung großer Korpora mit ihren entsprechenden Analysetools, worauf im später folgenden Kapitel „Wörterbuchbasis" noch näher eingegangen werden soll.

c. Bialphabetisches Wörterbuch mit Register

Das KOLLEX enthält nicht nur *ein* Wörterverzeichnis – das sogenannte Basis-Wörterverzeichnis –, in dem die Basen, also die Autosemantika des ZD-Wortschatzes als Lemmazeichen verzeichnet sind, sondern das Wörterbuch verfügt auch über ein zweites alphabetisches Wörterverzeichnis – Kollokator-Wörterverzeichnis –, das sukzessive entsteht. Dessen Lemmata bilden zukünftig diejenigen Kollokatoren, deren Kombinationsfähigkeit, genauer *Kollokationspotenzial*, sehr eingeschränkt ist. Das heißt, dass anhand der statistischen Kollokationsanalyse eine kleine, aufzählbare Menge von maximal 5–10 Basen für sie angegeben werden kann. Außerdem findet der potenzielle Benutzer sogar alle deutschen Kollokatoren in den Basis-Wörterbuchartikeln, wenn er das Register der im Wörterbuch verzeichneten Kollokatoren benutzt.

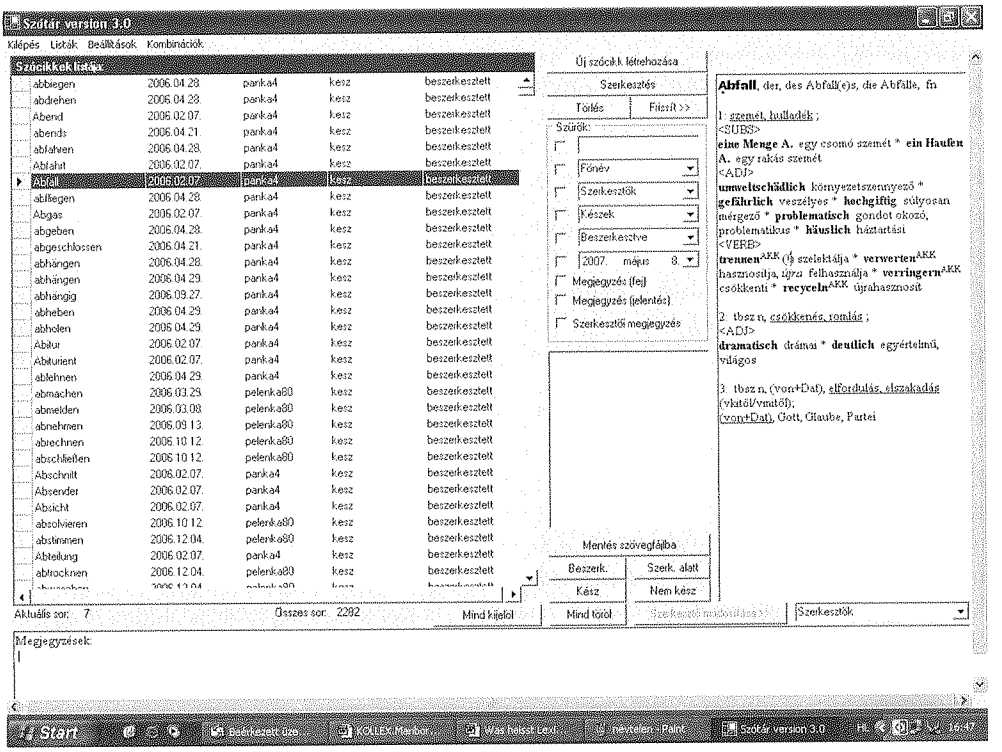
Die Typologisierung nach der Wörterbuchform gibt gleichzeitig Auskunft über die wichtigsten Elemente der Makrostruktur von KOLLEX. Dabei liegt eine glattalphabetische Makrostruktur – ohne Nischen- oder Nestbildung – mit zwei makrostrukturellen Zugriffsstrukturen für die zwei Wörterverzeichnisse und mit einem Register der Kollokatoren mit beschränktem Kollokationspotenzial vor.

d. Printwörterbuch und/oder digitales Wörterbuch

Eine spezielle Datenbank und ein darauf abgestimmtes Wörterbuchprogramm wurden extra für die Bedürfnisse der Erstellung von KOLLEX entwickelt und sind das Ergebnis erfolgreicher interdisziplinärer Zusammenarbeit. Das verwendete Datenbankmanagementsystem ermöglich nicht nur die manuelle, sondern auch die z.T. automatische Speicherung der selektierten Daten. Die Datenbank verfügt über internetseitenähnliche Arbeitsoberflächen, die gleichzeitig auch das vorläufige Ergebnis der Datenauswertung, d.h. die Wörterbuchartikel selbst zeigen.

Das Programm verfügt über 3 Benutzeroberflächen: eine für die studentischen Hilfskräfte, eine für den Lektor und eine für die Redakteurin. Darüber hinaus erlaubt es die Sortierung der Wörterbuchartikel nach verschiedenen Filter-Kriterien (in Abbildung 2: Szűrők): nach Buchstabe, nach Wortklasse, nach Status (erstellt und fertig bzw. redigiert), nach Anmerkungen der Hilfskräfte oder der Redakteurin und macht u.a. die Speicherung der Daten in eine Textdatei möglich.

Die Organisation der Daten in einer Datenbank erlaubt, dass das ganze Datensortiment sowohl als Printwörterbuch als auch als digitales Wörterbuch publiziert werden kann. Aus demselben Datenbestand können also Printwörterbuch, Offline-Wörterbuch (in Form von CD-ROM oder DVD) und Online-Wörterbuch erstellt werden.



2. Abbildung Startseite des Wörterbuchprogramms zur erstellten Datenbasis

4 Wörterbuchbasis und effiziente Nutzung großer Sprachkorpora

Die Wörterbuchbasis umfasst alle benutzten Quellen im aktuellen lexikographischen Prozess. Sie besteht aus den primären, sekundären und tertiären Quellen. Für das KOLLEX sind die ersten zwei von Relevanz, dabei haben die Primärquellen eine Sonderstellung.

Die fast ausschließliche Nutzung großer deutscher Sprachkorpora bei der Datenerhebung für ein Wörterbuch ist eine Pionierarbeit in der zweisprachigen Lexikographie und befindet sich auch in der einsprachigen Lexikographie erst im Anfangsstadium.

Die Zusammenarbeit mit der Leipziger Universität und die Nutzung des Korpus „Deutscher Wortschatz", das mit 20 Mio. Sätzen neben den IDS-Korpora die größte deutsche elektronische Textsammlung darstellt, ermöglicht, dass für die Datenauswertung eine enorme Menge von Korpusdaten zur Verfügung steht. Die Primärquelle für KOLLEX bildet daher das Leipziger Korpus. Zu allen Lemmazeichen des Wörterbuchs gibt es statistisch ermittelte Wortlisten, anhand derer die innere Selektion für die Mikrostruktur, d.h. die Auswahl der Kotextangaben erfolgen kann. Da in den gängigen Wörterbüchern zu bestimmten Kollokationstypen – wie bereits erläutert – kaum Daten vorhanden sind, zudem nicht unter der für die Sprachproduktion notwendigen Basis, und dazu noch die Mehrheit der Angaben durch Exzerpieren der Primärquellen – und dabei selbstverständlich auf frühere Sekundärquellen gestützt – ermittelt wurde, kann von Objektivität oder Repräsentativität der Wör-

terbuchangaben kaum die Rede sein. Durch spezifische Algorithmen kommt man bei solchen großen Sprachkorpora bereits zu wertvollen sprachlichen Daten in Mengen, die früher unvorstellbar gewesen wären und die durch die intellektuelle Weiterverarbeitung erstmalig die Möglichkeit einer adäquaten Dokumentation des tatsächlichen Sprachgebrauchs eröffnen. Hausmanns Affinitätsbegriff wurde nämlich damit mathematisch fassbar und durch Signifikanzgrenzen ermittelbar gemacht.[9] Dabei ist Vorsicht geboten, d.h. die aus den Korpora gewonnenen Daten werden mit Hilfe der Sekundärquellen überprüft und wenn nötig ergänzt. Zu den sekundären Quellen gehören diejenigen lexikographischen Nachschlagewerke, die bei der Datenerhebungsphase obligatorisch oder fakultativ einzusehen sind. Die Ermittlung der Sememe und damit der Äquivalente zu den einzelnen Lemmazeichen, also die semantische Mediostruktur, entstand mit Hilfe von ein- und zweisprachigen Wörterbüchern, bei den Äquivalenzangaben wurden die neukonzipierten Hessky-Wörterbücher herangezogen. Die Referenzwörterbücher, also die jederzeit obligatorisch zu konsultierenden einsprachigen Wörterbücher sind: LGWDAF (2003) und PGWDAF (2004), die zweisprachigen: HESS-NM ([3]2005) und HESSIKE (2005). Eine besondere Rolle kommt noch den deutschen Valenzwörterbüchern, unter anderem dem korpusbasierten VALBU (2004) und dem ungarischen Synonymwörterbuch zu.

Mit alldem wird für ein ausgewogenes Verhältnis zwischen den primären und sekundären Quellen gesorgt, womit die Repräsentativität und Objektivität der Wörterbuchangaben in möglichst hohem Maße gewährleistet wird.

5 Mikrostruktur

Artikelmikrostrukturen (kurz: Mikrostrukturen) gelten für hoch verdichtete und standardisierte Wörterbuchartikel. Die zwei unmittelbaren Textkonstituenten des Artikels sind der Formkommentar, einschließlich der Angabe zur Wortart, und der semantische Kommentar, bis zum Ende des Artikels. Zu monosemen Lemmazeichen in der wiegandschen Terminologie gehören nach dem ersten Typologiekriterium, nämlich Anzahl, Art und Reihenfolge der Kommentare, einfache Mikrostrukturen, wie es beim monosemen Lemmazeichen *analysieren* der Fall ist. Bei den Wörterbuchartikeln zu polysemen Lemmazeichen spielt das zweite Typologiekriterium, die Verteilung der Bedeutungsangaben und der von den Bedeutungen mitbestimmten Angaben im semantischen Kommentar, eine Rolle, deshalb weisen sie entweder integrierte, oder, seltener, annexierte Mikrostrukturen auf, worauf später eingegangen wird.

---

[9] Die genaue Methode der Datenselektion für das Wörterbuch habe ich bereits auf dem 4. lexikographischen Kolloquium in Maribor und auf Tagungen in Ungarn unter unterschiedlichen Gesichtspunkten dargestellt. Zugunsten der Vorstellung der Wörterbuchartikel und der Mikrostruktur wird darauf hier nicht weiter eingegangen.

analysieren analysierte, hat analysiert, AKK, ige, vál

elemez, analizál VMIT
AKK Roman, Text, Satz, Musikstück, Traum, Lage, Situation, Probleme, Beziehung
<ADV> **sorgfältig** gondosan * **eingehend** behatóan, tüzetesen * **gründlich** alaposan * **wissenschaftlich** tudományosan * **exakt** pontosan, egzakt módon * átv **näher** közelebbről * **statistisch** stilisztikailag * **systematisch** szisztematikusan * **logisch** logikusan, logikailag * vál **minuziös** (!) aprólékosan * **präzis(e)** pontosan, precízen

An dem obigen Musterartikel zum Lemmazeichen *analysieren* kann man die verschiedenen Arten der syntagmatischen Angaben auf beiden Ebenen – grammatisch und lexikalisch – deutlich erkennen. Syntagmatik wird im vorliegenden Wörterbuch sowohl auf der lexikalischen als auch auf der grammatischen Ebene – gemäß dem „Britischen Kontextualismus" – gedeutet und dargestellt, allerdings wird die *grammatische* anhand der Wörterbücher und nicht mit Korpusanalysen ermittelt. Das bedeutet konkret für die Valenzangaben, dass sie zu den Verblemmata aus dem korpusbasierten VALBU (2004) stammen, wie im Falle von *analysieren* die obligatorische Akkusativergänzung. Auch die Angaben zur Valenzrealisierung (z.B. nach dem Kürzel „Akk" aufgelistete konkrete Akkusativobjekte) haben ihre Quellen in den neueren großen einsprachigen Wörterbüchern (vor allem LGWDAF und PGWDAF), notfalls allerdings mit Korpusdaten ergänzt.

Die syntagmatischen Relationen auf der *lexikalischen* Ebene werden im KOLLEX anhand der fünf vorgestellten, für die Beschreibung der Kollokationen erarbeiteten Strukturtypen erfasst. Freie Wortkobinationen und Kollokationen wechseln sich innerhalb der einzelnen Kategorien ab, da keine zuverlässigen Abgrenzungskriterien erarbeitet werden konnten und außerdem nicht nur im Bereich der Kollokationen bei den Lernern Defizite herrschen. Durch diese Praxis kann die diskurstypische Kombinatorik der lexikalischen Einheiten am adäquatesten dargestellt werden. Andererseits variiert der Kollokationsbegriff in der einschlägigen Literatur stark, worauf in diesem eng gesteckten Rahmen nicht näher eingegangen werden kann.[10]

Ein weiteres Charakteristikum der Beschreibung der lexikalischen Syntagmatik ist, dass bei den Wörterbuchartikeln zu Adjketivlemmata außer der Kollokatorkategorie „Adverb" (ADV) auch Basen, also Substantive, gegebenenfalls Verben und/oder Adjektive aufgelistet werden. Sie erfüllen eine ähnliche Funktion wie die substantivischen Valenzrealisierungen bei den Verben, die auch als Basen zum verbalen Kollokator aufgefasst werden können. Ziel dieses Angabetyps ist es, außer diskurstypische semantische Netze aufzubauen, die

[10] Näheres dazu in Hollós (2004). Dort wurde auch eine eigene, für die Zwecke des Wörterbuchs entwickelte Kollokationstypologie vorgestellt, die sich mit gewissen Einschränkungen auch in der Praxis bewährt hat. Sie basiert gleichzeitig auf strukturellen, semantisch orientierten und korpuslinguistischen Auffassungen und bezieht außer der intralingualen Perspektive auch die interlinguale mit ein.

Syntagmen in einen möglichen Kommunikationszusammenhang zu stellen. Der folgende Musterartikel zum polysemen Adjektivlemma *weich* demonstriert diesen Angabetyp, außerdem wird daran die geplante Textarchitektur und der Einsatz der vorläufigen Typographie im zukünftigen Wörterbuch ersichtlich:

weich mn
❶ *nem kemény* lágy
❷ *nem éretlen v nyers* puha *gyümölcs*; lágy *tojás*
❸ *sima v selymes* puha *bőr*
❹ lágy *víz*
❺ *kellemes hatású* lágy *pl. vonások, hangzás*
❻ *engedékeny* lágy(szívű) *ember*
❼ *gyenge hatású* lágy *drogok*

❶ *nem kemény* lágy
+Subs Sessel, Bett, Matratze, Holz, Metall, Bleistift, Käse, Brot, Teig, Masse
<ADV> biz **schön** (!) jó * **wunderbar** csodásan * **angenehm** kellemesen * **vergleichsweise** viszonylag
❷ *nem éretlen v nyers* puha; lágy
+Subs Äpfel, Gemüse, Tomate, Ei, Fleisch
+Verb kochen
+Adj gekocht, gedünstet
<ADV> **innen** belül * biz **eher** *összehasonlításnál* inkább * **auffallend** feltűnően
❸ *sima v selymes* puha
+Subs Haut, Leder, Pelz, Fell, Wolle, Stoff
+Verb spülen
<ADV> biz **schön** szép, jó * **angenehm** kellemesen * **samtig** bársonyos(an) * **kuschelig** simogatóan
❹ lágy
+Subs Wasser
<ADV> **relativ** viszonylag * **ungewöhnlich** szokatlanul
❺ *kellemes hatású* lágy *pl. vonások, hangzás*
+Subs Gesichtszüge, Züge, Mund, Stimme, Klang, Licht, Farbe, Brauntöne, Landung
+Verb klingen
+Adj klingend
<ADV> biz **schön** szép * **wunderbar** csodá(lato)san * **angenehm** kellemesen
❻ *engedékeny* lágy(szívű)
+Subs Mensch, Gemüt, Herz
+Verb stimmen, machen
<ADV> biz **eher** *viszonyításnál* inkább * **angeblich** állítólag
❼ *gyenge hatású* lágy *drogok*
+Subs Drinks, Zigarette, Drogen
<ADV> **sogenannt** úgynevezett * **relativ** viszonylag

Bei diesem polysemen Lemmazeichen teilt sich der semantische Kommentar in so viele semantische Subkommentare, wie das Lemmazeichen Bedeutungen/Sememe aufweist. Bei der Zuordnung der Mikrostruktur zu einem bestimmten Typ kommt hier das zweite, bereits genannte Typologiekriterium zur Anwendung. Demnach gehören die meisten Wörterbuchartikel zu den integrierten Mikrostrukturen. Für einige wenige Wörterbuchartikel kommen nur die annexierten Mikrostrukturen in Frage, weil nur diese Datenverteilung die unnötige Wiederholung mehrerer gleicher Kotextpartner/Kollokatoren ab einer gewissen Zahl, etwa bei 6, erlaubt, d.h. sie werden am Ende des Wörterbuchartikels, im sogenannten Annex, aufgelistet.

Zusammenfassend kann man schließlich folgende Spezifika der Mikrostruktur nennen:

- Detaillierte Bedeutungsstrukturen mit Kurzübersicht der Sememe
- Sememspezifische Zuordnung der Kollokatoren gemäß den Strukturtypen
- Verzeichnung der substantivischen, verbalen und adjektivischen Kotextpartner in den Adjektivartikeln
- Verzeichnung der grammatischen Kollokationen (Valenzen) mit deren Verwirklichungen (Valenzrealisierungen) in den Verb-, Substantiv- und Adjektivartikeln.

6 Resümee

Zwar richtet sich das Wörterbuch an fortgeschrittene Lerner oder Germanistikstudenten, aber auch Schüler mit geringeren Sprachkenntnissen dürften damit Erfolge verzeichnen können. Außerdem werden auch Lehrer zur Korrektur von Schüleraufsätzen und angehende Übersetzer – vor allem bei der Hinübersetzung – das KOLLEX gezielt einsetzen können. Die besten Dienste kann es allerdings in erster Linie in der Sprachproduktion leisten, u.U. mit anderen modernen ein- und zweisprachigen Wörterbüchern gemeinsam. Der lange Traum der fortgeschrittenen ungarischsprachigen Deutschlerner wird also in naher Zukunft endlich Wirklichkeit.

7 Literatur

7.1 Korpora

IDS-Korpus. Institut für Deutsche Sprache, Mannheim. Internet-Adresse: http://www.ids-mannheim.de/cosmas2

Leipziger Korpus. Deutscher Wortschatz. Internet-Adresse: http://wortschatz.informatik.uni-leipzig.de

7.2 Wörterbücher

BBI 1986 = The BBI Combinatory Dictionary of English. A Guide to Word Combinations, compiled by Morton Benson, Evelyn Benson and Robert Ilson. Amsterdam/Philadelphia 1986.

DDAFSTW 2002 = Duden. Deutsch als Fremdsprache. Standardwörterbuch. Hrsg. vom Wissenschaftlichen Rat der Dudenredaktion: Matthias Wermke, Kathrin Kunkel-Razum, Werner Scholze-Stubenrecht. Mannheim [usw.] 2002.

DGWDAF 2000 = de Gruyter Wörterbuch. Deutsch als Fremdsprache. Hrsg. von Günther Kempcke unter Mitarbeit von Barbara Seelig, Birgit Wolf, Elke Tellenbach und Edelgard Dückert, Margot Richter, Vera de Duiter, Renate Schmidt, Karl Wunsch. Berlin. New York 2000.

HESS-NM ^{3}2005 = Hessky Regina (főszerk.): Német-magyar kéziszótár. Harmadik, javított kiadás. Szeged 2005.

HESS-MN ^{2}2005 = Hessky Regina (főszerk.): Magyar-német kéziszótár. Második, javított kiadás. Szeged 2005.

HESSIKE 2005 = Hessky Regina/Iker Bertalan: Német-magyar, magyar-német tanulószótár. Szeged 2005.

LGWDAF 2003 = Langenscheidts Großwörterbuch Deutsch als Fremdsprache. Das neue einsprachige Wörterbuch für Deutschlernende. Neubearbeitung. Hrsg. v. Dieter Götz, Günther Haensch, Hans Wellman. In Zusammenarbeit mit der Langenscheidt-Redaktion. Leitende Redakteure: Vincent J. Docherty, Günther Jehle. Berlin [usw.] 2003.

OCDSE 2002 = Oxford Collocations Dictionary for students of English. Eds. Jonathan Crowther, Sheila Dignen, Diana Lea. Oxford 2002.

PGWDAF 2004 = Pons Großwörterbuch Deutsch als Fremdsprache. Bearb. v. Susanne Balhar, Frank Mönkemöller, Elke Hippe, Silke Löhle, Martina Lacatena, Lydia Quaas, Ulrich Saßmann, Julia Weißmüller. Barcelona [usw.] 2004.

VALBU 2004 = VALBU – Valenzwörterbuch deutscher Verben. Hrsg. Von Helmut Schuhmacher, Jacqueline Kubczak, Renate Schmidt, Vera de Ruiter. Tübingen 2004.

7.3 Sekundärliteratur in Auswahl

Bahns 1996 = Jens Bahns: Kollokationen als lexikographisches Problem. Eine Analyse allgemeiner und spezieller Lernerwörterbücher des Englischen. Tübingen 1996. (Lexicographica. Series Maior 74)

Hausmann 1984 = Franz Josef Hausmann: Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen. In: Praxis des neusprachlichen Unterrichts 31. 1984, 395–406.

Hollós 2004 = Zita Hollós: Lernerlexikographie: syntagmatisch. Konzeption für ein deutsch-ungarisches Lernerwörterbuch. Tübingen 2004. (Lexicographica. Series Maior 116).

Réder 2006 = Anna Réder: Kollokationen in der Wortschatzarbeit. Wien 2006.

Steyer 2002 = Kathrin Steyer: Wenn der Schwanz mit dem Hund wedelt. Zum linguistischen Erklärungspotenzial der korpusbasierten Kookkurrenzanalyse. In: Haß-Zumkehr, Ulrike/Kallmeyer, Werner/Zifonun, Gisela (Hrsg.): Ansichten zur deutschen Sprache. Festschrift für Gerhard Stickel zum 65. Geburtstag. Tübingen 2002. (Studien zur deutschen Sprache).

Wiegand 1998a = Herbert Ernst Wiegand: Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilbd. Mit 159 Abbildungen im Text. Berlin. New York 1998.

Wiegand 1998b = Herbert Ernst Wiegand (Hrsg.): Perspektiven der pädagogischen Lexikographie des Deutschen. Untersuchungen anhand von „Langenscheidts Großwörterbuch Deutsch als Fremdsprache". Tübingen 1998. (Lexicographica. Series Maior 86).