



universität
wien

Helena Kempf

Mehrsprachige elektronische Phraseographie

Konzeptueller Entwurf eines Wörterbuchs
und Praxistest mit *SDL MultiTerm*

Masterarbeit zur Erlangung
des akademischen Grades

Master of Arts (MA)

Wien, im Oktober 2010

Studienkennzahl: A 060 357 381

Studienrichtung: Masterstudium Übersetzen UG 2002

Betreuer: Univ.-Prof. Mag. Dr. Gerhard Budin

Abstract

Multiword units (phrasemes), such as *to give the game away* or *to have second thoughts*, are frequently underrepresented in bilingual dictionaries, or they lack the functional information needed for active use in the target language. The present thesis examines possible ways of organising comprehensive functional, syntactic, semantic and translation-related information in a multilingual electronic dictionary. Initially, basic notions of monolingual and contrastive phraseology research are presented and discussed. Based on these foundations, a conceptual schema for a multilingual phraseological electronic dictionary is suggested, in which one phraseme is linked to several L2-equivalents, providing translation-related functional information for each phraseme combination. Separating monolingual from translation-related data is a core element of the concept, which allows the re-use of monolingual data for different language combinations. Data categories are discussed in particular, as well as methods of data collection and editing. An emphasis is put on collaborative online-editing, in which usability issues play a crucial role. Subsequently, the designed schema is put into practice with *SDL MultiTerm 2007*. For this purpose, three different entry models are created and tested by means of a small Hungarian-German example corpus. In conclusion, these entry models are compared. General advantages and disadvantages of the use of *MultiTerm* for phraseographic purposes are discussed.

Zusammenfassung

Mehrworteinheiten wie *die Gunst der Stunde nutzen* oder *ein Sprung ins kalte Wasser* (sog. Phraseologismen) sind in bilingualen Wörterbüchern häufig unterrepräsentiert oder mit unzureichenden Angaben zur aktiven Verwendung in der Zielsprache versehen. Die vorliegende Arbeit untersucht daher Möglichkeiten, Phraseologismen in einem elektronischen multilingualen Wörterbuch zu verwalten und praxisrelevante Information in strukturierter Form bereit zu stellen. Eingangs werden Grundbegriffe aus der einsprachigen und kontrastiven Phraseologieforschung eingeführt und diskutiert. Auf dieser Basis wird ein konzeptueller Entwurf einer mehrsprachigen phraseographischen Datenbank vorgeschlagen. Wesentliche Elemente sind die Zuordnung eines Eintrags zu mehreren ZS-Äquivalenten sowie die Trennung einsprachiger und übersetzungsbezogener Daten. Einzelne Datenkategorien sowie Formen der Datengewinnung und -bearbeitung werden besprochen. Ein Schwerpunkt liegt dabei auf kollaborativer Online-Bearbeitung, wobei auch Usability-Aspekte ins Blickfeld rücken. Anschließend wird die entworfene Datenbankstruktur mit SDL MultiTerm 2007 praktisch getestet. Hierfür werden drei verschiedene Eintragsmodelle entworfen und anhand eines ungarisch-deutschen Beispielkorpus verglichen. Abschließend werden allgemeine Vor- und Nachteile der phraseographischen Nutzung von MultiTerm diskutiert.

Danksagung

Folgende Personen haben durch ihre Unterstützung wesentlich zum Gelingen dieser Arbeit beitragen; ihnen möchte ich daher herzlich danken:

- meiner Mutter **Elisabeth Kempf**, die mir dieses Studium durch ihre finanzielle Unterstützung ermöglicht hat;
- meinem Betreuer **Prof. Dr. Gerhard Budin** für fachlichen Rat und Förderung;
- **Mag. Irmgard Soukup-Unterweger** für ihre aufschlussreiche und geduldige Beratung im Zusammenhang mit *SDL MultiTerm*;
- **Luise Kempf, M.A.** für fachlichen Rat in linguistischen Fragen sowie für stilistische Hinweise;
- **Dr. Dóra Solti** für die Durchsicht des ungarisch-deutschen Beispielkorpus, linguistischen Rat sowie das Zuverfügungstellen griechischer Beispiele;
- **Orsolya Mária Birta** für die Durchsicht des ungarisch-deutschen Beispielkorpus und Klärung einschlägiger Fragen;
- **Julija Schellander, Luise Kempf, Elisabeth Kempf und Nadine Bartels** für die formale Revision dieser Arbeit.

Verwendete Abkürzungen:

| | | | |
|---------|-----------------|------|----------------------------------|
| Abb. | Abbildung | jds | jemandes |
| AdjP | Adjektivphrase | L1 | eine von zwei Vergleichssprachen |
| AdvP | Adverbialphrase | L2 | eine von zwei Vergleichssprachen |
| Akk. | Akkusativ | | |
| AS | Ausgangssprache | | |
| Bsp/bsp | Beispiel | NP | Nominalphrase |
| Dat | Dativ | PP | Präpositionalphrase |
| d.h. | das heißt | QUE | Quelle |
| ebd | ebenda | s. | siehe |
| etw | etwas | u.a. | unter anderem, unter anderen |
| Gen | Genitiv | | |
| ggf. | gegebenenfalls | vgl. | vergleiche |
| HK | die Verfasserin | VP | Verbalphrase |
| iSv. | im Sinn von | vs. | versus |
| jd | jemand | z.B. | zum Beispiel |
| jdm | jemandem | ZS | Zielsprache |
| jdn | jemanden | Σ | Satz (Konstituentenstruktur) |

Hinweis:

Die weiblichen und männlichen Bezeichnungen (*die Übersetzerin/ der Übersetzer*) werden im Wechsel verwendet. Beide Formen sind generisch gemeint und stehen stellvertretend für beide Geschlechter.

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einleitung | 1 |
| 2 | Grundbegriffe..... | 5 |
| 2.1 | Merkmale von Phraseologismen..... | 6 |
| 2.1.1 | Festigkeit | 6 |
| 2.1.2 | Polylexikalität..... | 8 |
| 2.1.3 | Idiomatizität..... | 9 |
| 2.2 | Zum Phraseologiebegriff | 10 |
| 2.2.1 | Weite oder enge Definition? | 10 |
| 2.2.2 | Definition auf Basis der Grundmerkmale..... | 12 |
| 2.2.3 | Impulse aus der Korpuslinguistik..... | 13 |
| 3 | Phraseologische Klassen | 15 |
| 3.1 | Verschiedene Merkmale | 16 |
| 3.1.1 | Einteilung nach semiotischen Merkmalen | 16 |
| 3.1.2 | Einteilung nach semantischen Merkmalen..... | 16 |
| 3.1.3 | Einteilung nach syntaktischen Merkmalen | 17 |
| 3.1.3.1 | Syntaktische Klasse..... | 17 |
| 3.1.3.2 | Kontextanbindung | 19 |
| 3.1.3.3 | Interne und externe Leerstellen | 21 |
| 3.1.3.4 | Innere morphosyntaktische Struktur | 23 |
| 3.1.4 | Weitere Merkmale | 24 |
| 3.2 | Systematik..... | 24 |
| 3.2.1 | Basisklassifikation nach Burger (2003) | 24 |
| 3.2.2 | Darstellung im Begriffsfeld..... | 26 |
| 4 | Phraseologie im interlingualen Arbeiten | 29 |
| 4.1 | Phraseologie im translatorischen Gesamtkontext | 29 |
| 4.2 | Kontrastive Phraseologiedefinition..... | 30 |
| 4.2.1 | Mangelnde Analogie..... | 33 |
| 4.2.2 | Schnittstelle zwischen Grammatik und Lexik | 33 |
| 4.3 | Äquivalenzbeziehungen | 34 |
| 4.3.1 | Dimensionen der Äquivalenz..... | 34 |
| 4.3.2 | Übersetzungsverfahren | 38 |
| 4.4 | Phraseologie in Wörterbüchern | 40 |
| 5 | Entwurf einer mehrsprachigen elektronischen Phraseologiedatenbank | 43 |
| 5.1 | Grundeigenschaften..... | 44 |
| 5.2 | Ordnungs- und Suchprinzipien | 46 |
| 5.3 | Datenbankstruktur..... | 47 |
| 5.4 | Datenkategorien | 48 |
| 5.4.1 | Einsprachige Datenkategorien | 48 |
| 5.4.1.1 | Verwaltungsbezogene Daten..... | 48 |
| 5.4.1.2 | Kerninformationen zum Phraseologismus | 49 |
| 5.4.1.3 | Allgemeine benennungsbezogene Angaben | 51 |
| 5.4.1.4 | Funktionale Angaben | 53 |

| | | |
|-------------------------|---|------------|
| 5.4.1.5 | Semantische Angaben | 55 |
| 5.4.1.6 | Syntaktische Angaben..... | 56 |
| 5.4.1.7 | Weitere interessensspezifische Angaben..... | 58 |
| 5.4.1.8 | Zielsprachliche Angaben zum Phraseologismus..... | 59 |
| 5.4.2 | Übersetzungsbezogene Datenkategorien..... | 60 |
| 5.4.3 | Übersicht Datenkategorien..... | 60 |
| 5.4.4 | ER-Modell..... | 60 |
| 5.5 | Contenterstellung und Bearbeitung | 63 |
| 5.5.1 | Datengewinnung | 63 |
| 5.5.2 | Online-Kollaboration..... | 65 |
| 5.5.3 | Eingabe und Repräsentation der Daten..... | 66 |
| 5.5.3.1 | Statistische Datenauswertung | 69 |
| 5.5.4 | Datenabfrage..... | 70 |
| 5.5.5 | Usability..... | 71 |
| 5.6 | Resümee: Kriterienkatalog für ein elektron. phraseologisches Wörterbuch . | 72 |
| 6 | Phraseographie am Beispiel von <i>SDL MultiTerm</i>..... | 75 |
| 6.1 | Verwendetes Korpus | 75 |
| 6.2 | Übliche Verwendung von <i>MultiTerm</i> | 79 |
| 6.2.1 | Allgemeines | 79 |
| 6.2.2 | Eintragsstruktur | 79 |
| 6.2.3 | Anlegen der Datenbank | 82 |
| 6.2.4 | Datenabfrage..... | 86 |
| 6.3 | Phraseographische Nutzung von <i>MultiTerm</i> | 90 |
| 6.3.1 | Einsprachige Daten | 91 |
| 6.3.2 | Übersetzungsbezogene Daten..... | 95 |
| 6.3.3 | Mögliche Datenmodelle..... | 95 |
| 6.3.3.1 | A: Übersetzungsdaten als Indexfeld..... | 95 |
| 6.3.3.2 | B: Übersetzungsdaten auf Eintragsebene | 105 |
| 6.3.3.3 | C: Übersetzungsdaten auf Termebene | 107 |
| 6.3.4 | Vergleich der drei Datenmodelle | 112 |
| 6.4 | Vorteile von <i>MultiTerm</i> | 113 |
| 6.5 | Grenzen von <i>MultiTerm</i> | 114 |
| 7 | Zusammenfassung und Ausblick..... | 117 |
| 8 | Literatur- und Quellenverzeichnis..... | 121 |
| 8.1 | Fachliteratur | 121 |
| 8.2 | Wörterbücher | 123 |
| 8.3 | Internetquellen | 124 |
| 10 | Abbildungsverzeichnis | 125 |
| 11 | Tabellenverzeichnis | 126 |
| 12 | Anhang A: Korpus HU-DE | 127 |
| Lebenslauf | | 131 |

1 Einleitung

Mehrworteinheiten wie die *Gunst der Stunde nutzen* oder *ein Sprung ins kalte Wasser* (sog. **Phraseologismen**) stellen im bilingualen Arbeiten eine besondere Herausforderung dar: Als feste sprachliche „Bausteine“ machen sie einen hohen Prozentsatz sprachlicher Äußerungen aus, werden jedoch häufig erst nach Jahren in der Sprachgemeinschaft erworben. Viele Phraseologismen dienen als Marker für bestimmte Stile, Textsorten oder Sprechaktypen. Phraseologisches Wissen ist somit mehr als bloßer stilistischer Feinschliff; vielmehr beinhaltet es einen wesentlichen Schlüssel zur Kultur.

Umso wichtiger ist daher entsprechende Information in bilingualen Wörterbüchern. Doch erstaunlicherweise sucht man dort oft vergeblich: Wenn ein Phraseologismus überhaupt verzeichnet ist, dann häufig mit unzureichenden Angaben zur aktiven Verwendung in der Zielsprache. Diese Lücke wird in der vorliegenden Arbeit in Angriff genommen: Sie behandelt die Frage, wie Phraseologismen in **mehrsprachige elektronische Wörterbücher** eingebunden werden können, und zwar so, dass Translatorinnen und Sprachlernenden in strukturierter Form praxisrelevante Information an die Hand gegeben wird.

Phraseologie umfasst einerseits bildliche Ausdrücke wie *Feuer und Flamme für etwas sein* oder *jdn. auf die Folter spannen*, auch **Idiome** genannt, und andererseits wörtlich verstandene Ausdrücke, sogenannte **Kollokationen**, wie *ein reibungsloser Ablauf*, *die Flucht ergreifen* oder *die Wurzel ziehen aus*, die besonders im fachsprachlichen Bereich von großer Bedeutung sind. Während bei Idiomen das Verständnis in einer Fremdsprache aufgrund der übertragenen Bedeutung Schwierigkeiten bereiten kann, sind Kollokationen passiv verständlich, wenn die Bedeutung ihrer einzelnen Bestandteile bekannt ist. Die aktive Bildung hingegen ist in beiden Fällen nur mit Vorwissen möglich – so ist für Nichtmuttersprachler des Deutschen keineswegs offensichtlich, dass das zu *Brand* gehörende Verb *stiften* heißen muss und nicht etwa *machen* oder *entfachen*.

Aus diesem Grund besteht Bedarf an einem **aktiven** phraseologischen Wörterbuch, das umfassende Information zur Verwendung eines Phraseologismus in der Zielsprache liefert. Hierzu zählt nicht nur die genaue Bedeutung, sondern auch Angaben über die Einbettung in den syntaktischen Kontext, stilistische Faktoren und vieles mehr. Eine weitere Herausforderung an die mehrsprachige phraseologische Lexikographie (*Phraseographie*) besteht in der Tatsache, dass sich häufig **kein genaues zielsprachliches Äquivalent** zu einem Phraseologismus findet; dennoch sind Hinweise auf mögliche Übersetzungen in einem Wörterbuch wünschenswert. Dass sich in bestehenden gedruckten Wörterbüchern oft nur unzureichende Informationen finden, ist

nicht zuletzt auf die begrenzten Möglichkeiten des Mediums *Buch* zurückzuführen, sowohl was den Umfang als auch die Vernetzung von Informationen betrifft.

Zielsetzung dieser Arbeit ist daher die **Konzeption eines elektronischen mehrsprachigen phraseologischen Wörterbuchs**, das einem **produktiven** Anspruch genügt, d.h., die zur aktiven Verwendung eines zuvor unbekannten Phraseologismus benötigen Informationen liefert. Das Wörterbuch soll in unterschiedliche Übersetzungsrichtungen verwendbar, also **multidirektional** sein. Zudem soll es **fachgebietsunabhängig** sein, wohl aber die sachgemäße Einbeziehung fachsprachlicher Daten ermöglichen. Weiterhin wird die Möglichkeit der Dateneingabe und -bearbeitung durch **Web-Kollaboration** einbezogen, um durch Beteiligung vieler Mitwirkenden Aktualität und einen umfangreichen, repräsentativen Datenbestand zu gewährleisten.

Angestrebtt wird ausschließlich eine **konzeptuelle Datenmodellierung**, einschließlich der Beschreibung von Datenfeldern, ihrer Eigenschaften und wechselseitigen Beziehungen. Die Implementation in ein spezifisches Datenbanksystem ist nicht Gegenstand dieser Arbeit. Das Konzept basiert auf theoretischer Überlegung, die eine hohe Allgemeingültigkeit zum Ziel hat, und wird anhand zahlreicher Beispiele erläutert.

In einem weiteren Schritt wird die zuvor theoretisch erarbeitete Datenbankstruktur in einem Praxistest mit der Terminologieverwaltungssoftware **SDL MultiTerm 2007** umgesetzt. *MultiTerm* erlaubt eine flexible Termbankstrukturierung sowie kollaborative Bearbeitung, ist aber von Haus aus nicht explizit auf Phraseologismen zugeschnitten. Ein weiteres Ziel dieser Arbeit ist es somit, herauszufinden, **ob und inwieweit sich MultiTerm für phraseographische Zwecke eignet**.

Um die beschriebenen phraseographischen Fragestellungen untersuchen zu können, werden zunächst Grundlagen aus der linguistischen Phraseologieforschung benötigt. Kapitel 2 befasst sich daher mit phraseologischen Grundbegriffen und einer näheren begrifflichen Eingrenzung. In Kapitel 3 werden verschiedene phraseologische Klassen und Merkmale besprochen, die eine wichtige Grundlage für die Planung von Datenkategorien im Wörterbuch darstellen.

Während Kapitel 2 und 3 sich nur auf den einsprachigen Bereich beziehen, geht es in Kapitel 4 um die Rolle der Phraseologie im interlingualen Arbeiten. Dabei richtet sich das Augenmerk zunächst auf den translatorischen Gesamtkontext. Daraufhin folgen kontrastive Betrachtungen auf Textebene, die gleichzeitig einen neuen Blickwinkel auf die Begriffsdefinition der Phraseologie eröffnen. Weiterhin werden in diesem Kapitel mögliche Äquivalenzbeziehungen zwischen Phraseologismen zweier Sprachen sowie entsprechende Übersetzungsverfahren untersucht. Dies dient als theoretische Grundlage für die Darstellung von Äquivalenzbeziehungen im Wörterbuch. Abschließend werden einige Mängel gedruckter Phraseographie aufgezeigt, um diese bei der Konzeption eines elektronischen Wörterbuchs zu vermeiden.

In Kapitel 5 wird ein konzeptueller Entwurf für eine mehrsprachige phraseologische Datenbank nach den beschriebenen Kriterien erarbeitet. Dabei werden zunächst die wesentlichen Eigenschaften des Wörterbuchs festgelegt (s. 5.1, 5.2), um daraufhin eine grundlegende Strukturierung der Daten zu skizzieren (5.3). Dem folgt eine detaillierte Beschreibung diverser Datenkategorien und ihrer Eigenschaften (5.4). Weiterhin kommen in diesem Kapitel Aspekte der Datengewinnung und -bearbeitung zur Sprache, wobei ein besonderes Augenmerk auf die Möglichkeit der Web-Kollaboration geworfen wird (5.5). Hieraus ergeben sich Fragestellungen zur Organisation der Datenbearbeitung sowie zur Usability, die einen wesentlichen Eckpfeiler für das Funktionieren eines Projekts dieser Größenordnung darstellen.

Kapitel 6 schließlich dokumentiert den Praxistest der entworfenen Datenbankstruktur mit Hilfe von *MultiTerm*. Zur Durchführung dieser Untersuchung werden drei verschiedene Eintragsmodelle erstellt und anhand eines kleinen ungarisch-deutschen Beispielkorpus getestet. Dabei werden einerseits die drei Datenmodelle auf Praxistauglichkeit hin geprüft und verglichen, andererseits werden allgemeine Stärken und Schwächen der Software bei der Verwaltung von Phraseologismen analysiert.

2 Grundbegriffe

Beim Übersetzen, Dolmetschen oder Erlernen einer Fremdsprache stellt sich häufig die Frage nach einer passenden Übersetzung bzw. der Produktion von Ausdrücken folgender Art:

*einen grünen Daumen haben
sich verbreiten wie ein Lauffeuer
Verdacht schöpfen
blau machen
über die Runden kommen
ein Gespräch führen
hohe Wellen schlagen
Zeit heilt Wunden.
ein reibungsloser Ablauf
die Nacht zum Tag machen
Das Eis ist gebrochen.*

*Zu Risiken und Nebenwirkungen lesen Sie die Packungsbeilage oder fragen Sie Ihren Arzt oder Apotheker.
Das also ist des Pudels Kern!
Die Luft ist rein.
ein Tropfen auf den heißen Stein
der Nabel der Welt
ein spitzer Winkel
meines Erachtens
mit Kind und Kegel
aus heiterem Himmel*

Eine solche fest verbundene Wortgruppe bezeichnet man als Phrasem oder **Phraseologismus**, ihre lexikalischen Bestandteile als **Komponenten** (vgl. Burger 2003:11; 35). In der Computerlinguistik ist häufig auch von **Mehrworteinheiten (multi-word units)** die Rede (vgl. Heid 2007:1036, Bubenhofer 2010:7ff.). Der Fachbereich, der sich mit solchen Ausdrücken befasst, heißt **Phraseologie**, ebenso wie die Gesamtheit dieser Ausdrücke selbst. Das Teilgebiet, das sich mit der Aufzeichnung von Phraseologismen in Wörterbüchern befasst, nennt man, analog zur Lexikographie, **Phraseographie** (vgl. Burger 2003:170).

In diesem Kapitel soll näher eingegrenzt werden, was genau Phraseologie ist, was sie beinhaltet und was nicht. Im Folgenden werden daher zunächst einige grundlegenden Merkmale der Phraseologismen besprochen. Dies soll schließlich zu einer genaueren Begriffsdefinition von Phraseologie führen.

2.1 Merkmale von Phraseologismen

Burger (²2003) weist den Phraseologismen drei Haupteigenschaften zu, nämlich **Festigkeit**, **Polylexikalität** und **Idiomatizität**¹. Diese drei Eigenschaften sollen nun kurz vorgestellt werden.² Im Anschluss soll untersucht werden, in welchem Zusammenhang die drei Haupteigenschaften zueinander stehen.

2.1.1 Festigkeit

Mit **Festigkeit**, manchmal auch Stabilität genannt, ist gemeint, dass die Komponenten in einer festen Verbindung miteinander auftreten. Hierbei spielen mehrere Aspekte eine Rolle:

Zum einen meint Festigkeit, dass der Ausdruck genau in dieser Wortkombination innerhalb einer Sprachgemeinschaft **gebräuchlich** ist, was bedeutet, dass die betreffende Wendung einem Mitglied der jeweiligen Sprachgemeinschaft im Allgemeinen bekannt ist. Dies schließt jedoch nicht die Existenz von individuellen oder regionalen **Varianten** aus, wie die folgenden Beispiele zeigen:

*jdm auf die Nerven gehen/ jdm auf den Nerv fallen
Gut Ding will Weile haben./Gut Ding braucht Weile.*

Jüngere Korpusanalysen haben ergeben, dass Phraseologismen in weitaus mehr Varianten vorkommen als bislang angenommen, wodurch die Grenze zu freien Wortverbindungen unschärfer wird (s. 2.2.3) Auch kann ein Phraseologismus nur in manchen Regionen einer Sprachgemeinschaft verbreitet sein. Des Weiteren spricht man von **psycholinguistischer Festigkeit**. Gemeint ist, dass „der Phraseologismus [...] mental als Einheit ‚gespeichert‘ [ist,] ähnlich wie ein Wort“ und „als ganzer abgerufen und produziert werden“ kann (Burger ²2003:17). Das heißt, ein Phraseologismus ist genau wie ein Lexem in den Köpfen der Sprecherinnen fest verankert und abrufbar. So wären die meisten Sprecher des Deutschen in der Lage, den Satzanfang *Morgenstund hat...* richtig mit *Gold im Mund* zu ergänzen. Ein weiterer Beleg ist die Tatsache, dass Phraseologismen auch nach Abwandlung noch erkennbar bleiben. Der Werbeslogan der Drogeriekette dm *Hier bin ich Mensch, hier kauf ich ein* lässt sich als Modifikation

¹ Fleischer (1997) hingegen unterscheidet **Idiomatizität**, **Stabilität** und **lexikalische Einheit**. Burgers Begriff der Festigkeit umfasst jedoch im Wesentlichen das, was Fleischer unter Stabilität und lexikalischer Einheit einzeln anführt. Die Polylexikalität setzt Fleischer schon als gegeben voraus und beschreibt sie daher nicht näher.

² Eine ausführliche Beschreibung findet sich bei Burger ²2003.

des Ausspruchs *Hier bin ich Mensch, hier darf ich's sein* aus Goethes Faust identifizieren, vorausgesetzt natürlich, dass der Rezipient das Originalzitat kennt.

Auch in **struktureller Hinsicht** weisen Phraseologismen Festigkeit auf. Das zeigt sich z.B. darin, dass sie sich nicht ohne weiteres durch morphosyntaktische und semantisch-lexikalische Operationen abwandeln lassen, die im normalen Sprachgebrauch zulässig wären.³ Beispielsweise kann das Ersetzen einer Komponente durch ein anderes Wort oder Abwandlungen in Tempus und Numerus unzulässig sein:

Der Apfel fällt nicht weit vom Stamm. → **Die Birne fällt nicht weit vom Stamm.*

Der Apfel fällt nicht weit vom Stamm. → **Die Äpfel fielen nicht weit von den Stämmen.*

Man spricht hier von **fester Wortbindung**, d.h., die einzelnen Komponenten können nur in einer bestimmten, festgesetzten Beziehung zueinander stehen. Im Gegensatz dazu steht die **freie Wortbindung** im normalen Sprachgebrauch, bei der ein Ausdruck unter Einhaltung der morphosyntaktischen Regeln einer Sprache beliebig variiert werden kann.

Die strukturelle Festigkeit zeigt sich außerdem in vom üblichen Sprachgebrauch abweichenden syntaktischen und semantischen Strukturen. Beispielsweise gibt es so genannte **unikale Komponenten** also Wörter, die nur im Zusammenhang mit dem betreffenden Phraseologismus, nicht aber im allgemeinen Sprachgebrauch Verwendung finden (*Hehl* in *keinen Hehl aus etwas machen*; *Schliche* in *jdm. auf die Schliche kommen*). Auch **syntaktisch** können **Irregularitäten** auftreten, wie etwa der vorangestellte Genitiv in *in Teufels Küche kommen* oder die untypische Verbvalenz in *Löcher in die Luft starren*.⁴

³ Außer bei Modifikationen als bewusster Spielerei, die auch als solche wahrgenommen wird.

⁴ Für solche Irregularitäten kann es vielerlei Gründe geben. U.a. können sie auf Überresten einer älteren Sprachstufe beruhen, die im Phraseologismus sozusagen „konserviert“ wurden, während die entsprechende Form im allgemeinen Sprachgebrauch ausgestorben ist. Diese „Konservierung“ kann als weiterer Beleg für die Annahme gelten, dass Phraseologismen kognitiv als Einheit erfasst werden, und nicht als ein nach Regeln zusammengesetztes Gebilde aus einzelnen Wörtern. Weiterhin können Irregularitäten auch auf ästhetisch-rhetorischen Faktoren beruhen (Reim, Alliteration).

2.1.2 Polylexikalität

Dies ist das scheinbar eindeutigste Kriterium für die Definition von Phraseologie: Die Polylexikalität besagt, dass ein Phraseologismus aus mehreren lexikalischen Komponenten bzw. aus mehr als einem Wort⁵ bestehen soll.

Eine mögliche Einschränkung erfährt die Polylexikalität durch die von Duhme (1991:60) vorgeschlagenen Klasse der **Einwortphraseologismen**⁶, einem Grenzbereich von Ausdrücken, die zwar nur aus einem Wort bestehen, jedoch im Deutschen auf eine bestimmte, metaphorische Wiese verwendet werden. Beispiele sind:

| | | |
|----------------------------|----------------------|-------------------------|
| <i>bedeutungsschwanger</i> | <i>Karteileiche</i> | <i>Querdenker</i> |
| <i>Dreikäsehoch</i> | <i>Kurschatten</i> | <i>Schubladendenken</i> |
| <i>Draufgänger</i> | <i>Ladenhüter</i> | <i>Spätzunder</i> |
| <i>Eintagsfliege</i> | <i>Langfinger</i> | <i>Staubfänger</i> |
| <i>Elfenbeinturm</i> | <i>Nervensäge</i> | <i>Tapetenwechsel</i> |
| <i>Gratwanderung</i> | <i>Nestwärme</i> | <i>Weichei</i> |
| <i>Grauzone</i> | <i>Papierkrieg</i> | <i>Zaungast</i> |
| <i>Hühnerauge</i> | <i>Quantensprung</i> | <i>Zeitlupe</i> |

Es handelt sich hier also nicht um Ad-hoc-Metaphern, die sich ein Schreiber beim Verfassen eines Textes ausgedacht hat und die nur aus dem Kontext heraus verständlich sind, sondern um allgemein bekannte Begriffe, bei denen nicht nur die Metaphorik als solche, also als häufig herangezogenes Vergleichsbild bekannt ist, sondern auch genau diese sprachliche Realisierung. Häufig handelt es sich um Komposita, die naturgemäß eine Affinität zu polylexikalalen Phraseologismen haben.

Fleischer (1997:249) lehnt das Konzept der Einwortphraseologismen ab, weil er darin „eine Überdehnung des Phraseologismus-Begriffes“ sieht. Für die Zwecke dieser Arbeit sollen sie jedoch nicht grundsätzlich außen vorgelassen werden, da sich bei der Übersetzung eines solchen Ausdrucks ähnliche Probleme ergeben können wie bei polylexikalalen Phraseologismen. Wegen der Tendenz des Deutschen zur Kompositbildung können deutsche Einwortphraseologismen fremdsprachigen Phraseologismen entsprechen, wie etwa bei *Haftbefehl* – portugiesisch *ordem de prisão*, *Quantensprung* – englisch *quantum leap*, aber auch umgekehrt, z.B. ungarisch *szabadlábon* – deutsch *auf freiem Fuß*.

⁵ Als *Wort* wird hier – unabhängig von begrifflichen Zuordnungen – das verstanden, was graphisch als ein Wort erscheint, da zum Auffinden eines Phraseologismus im Wörterbuch das graphisch isolierte Wort maßgeblich ist. Als Ausnahme mögen vom Verb getrennte Präfixe gelten; diese lassen sich jedoch auf eine unigraphe Grundform zurückführen.

⁶ angeführt bei Fleischer 1997: 249.

2.1.3 Idiomatizität

Hier geht es um die Frage, in welchem Maße ein Ausdruck idiomatisch ist. Idiomatisch ist ein Ausdruck dann, wenn seine Bedeutung nicht aus den Bedeutungen der Einzelkomponenten erschlossen werden kann, weil er irgendeine Form von übertragener Bedeutung trägt (*den Gürtel enger schnallen, Die Wände haben Ohren.*).

Ein idiomatischer Ausdruck hat also **zwei Lesarten**, nämlich eine wörtliche und eine übertragene, bzw., um Burgers Terminologie zu übernehmen, phraseologische Bedeutung. Zwischen beiden Ebenen existiert nicht notwendigerweise ein erkennbarer Zusammenhang.

Wichtig ist hierbei die Unterscheidung zwischen *Lesart* und *Bedeutung*: Mit *Lesart* ist die Art und Weise gemeint, wie ein Ausdruck verstanden werden *könnte* – dies entspricht aber nicht zwangsläufig seiner tatsächlichen Bedeutung. Wenn ein Phraseologismus eine wörtliche und eine phraseologische Lesart hat, so entspricht seine Bedeutung in den meisten Fällen nur der phraseologischen Lesart.

Die semantische Idiomatizität ist Burger (2003:31) zufolge „eine graduelle Eigenschaft von Phraseologismen“, d.h., phraselogische und wörtliche Lesart können unterschiedlich stark voneinander abweichen. Daher lässt sich sicher nicht jede Nuance systematisch beschreiben; dennoch kann man mit Burger (2003:38) grob unterscheiden zwischen **Idiomen**, **Teil-Idiomen** und **Kollokationen**. Idiome sind vollständig idiomatische Ausdrücke, die über zwei verschiedene Lesarten verfügen, unabhängig davon, wie stark die beiden Ebenen voneinander abweichen. Teil-Idiome sind Ausdrücke, bei denen eine Komponente idiomatisch ist, eine andere aber im wörtlichen Sinne verwendet wird (*einen Streit vom Zaun brechen, ein blinder Passagier*). Kollokationen schließlich sind Phraseologismen, die nur eine, nämlich ihre wörtliche Lesart haben, wie etwa *sich Gehör verschaffen, Aller Anfang ist schwer oder Berufung einlegen*.

Man geht von einer Zweiteilung der Kollokation in *Basis* und *Kollokator* aus, wobei die Basis der semantisch autonome Teil ist und der Kollokator die Basis modifiziert und seine Bedeutung erst in Kombination mit dieser erhält (vgl. Huber 1993:31). So wäre in *eine Entscheidung treffen* die *Entscheidung* die Basis und *treffen* der Kollokator; das Verb *treffen* ist in diesem Kontext nicht unabhängig semantisch interpretierbar, sondern erhält seine aktuelle Bedeutung erst durch den Kontext. Selbiges gilt für *einen Weg einschlagen, die Flucht ergreifen, Verdacht schöpfen*.

Können nun Kollokationen, also so alltägliche Ausdrücke wie *etwas von der Steuer absetzen* oder *jdm eine Rechnung ausstellen* ebenfalls zur Phraseologie gezählt werden? Diese Frage soll im nächsten Kapitel besprochen werden.

2.2 Zum Phraseologiebegriff

2.2.1 Weite oder enge Definition?

Die Frage, ob auch Kollokationen zur Phraseologie hinzugezählt werden sollen, wurde längere Zeit in der Phraseologieforschung kontrovers diskutiert. Einige Forscher bevorzugen eine Beschränkung auf idiomatische Ausdrücke, weil sie eine Überstrapazierung des Phraseologiebegriffs befürchten.⁷ Üblich ist die Unterscheidung zwischen **Phraseologie im engeren Sinne**, die Idiomatizität als notwendige Bedingung voraussetzt, und der **Phraseologie im weiteren Sinne**, die auch Kollokationen beinhaltet. Einige Überlegungen sprechen deutlich für die weitere Begriffsauflösung:

Erstens spielen in der translatorischen Praxis **fachsprachliche Phraseologismen** eine zentrale Rolle, bei denen es sich ja in der Regel um Kollokationen handelt. Auch Arntz (2001:113) äußert sich aus terminologischer Sicht „dafür, den Begriff der Phrase/Fachwendung mit Hinblick auf die fachsprachliche Realität sehr weit zu fassen“ und bezeichnet die Fachwendung als „unverzichtbare[s]“ „Bindeglied zwischen Terminus und Text“. Dies begründet er wie folgt:

„Von praktischem und theoretischem Interesse ist dabei insbesondere der Grad der Stabilität, die die Phrase in den verschiedenen Stadien ihrer Entwicklung aufweist. Vielfach stellt die Phrase nämlich lediglich eine Vorstufe auf dem Weg hin zum Terminus dar; es kann im Laufe der Zeit zu einer immer größeren Verdichtung auf der Benennungsebene kommen, ohne daß dies die zugrunde liegende Begriffsstruktur verändern müßte.“ (ebd)

Zweitens werden im mehrsprachigen Arbeiten **aktive Wörterbücher** benötigt, die Benutzern den richtigen Einsatz eines zuvor unbekannten Phraseologismus ermöglichen. Die aktive Bildung kann aber schon bei Kollokationen Schwierigkeiten bereiten, deren passives Verständnis kein Problem darstellt. So ist es für Nicht-Muttersprachler des Deutschen keineswegs intuitiv einsichtig, dass ein *Weg eingeschlagen* oder *die Flucht ergriffen* wird. Dabei spielen Kollokationen sogar eine noch grundlegendere Rolle als Idiome: Fremdsprachenlernende benötigen bereits auf einem recht frühen Niveau Kollokationen, um Sachverhalte richtig auszudrücken (*Geld abheben*, *eine Entscheidung treffen*), wohingegen der Gebrauch von blumigen Idiomen als sprachlicher Schmuck noch ein wenig verschoben werden kann. Informationen über Kollokationen

⁷ Z.B. Kjær 1990: „One might choose to consider these phraseological in a very broad sense, on the grounds of their reproduction as units, their lack of commutability [...] and because of their *Gebräuchlichkeit*. But [...] I reject the broad definitions of phraseology which have become increasingly popular in phraseology theory“ (ebd:24); "...I did not want to weaken the concept of phraseology any further“ (ebd:26).

können auch für Übersetzerinnen mit fundierten Sprachkenntnissen gefragt sein, sei es in der Fachsprache oder zur Behebung individueller Wissenslücken.

Drittens ist die Berücksichtigung von Kollokationen in der **maschinellen Übersetzung** relevant, bei der bekanntlich nicht von vorhandener sprachlicher Intuition ausgegangen werden kann. Sowohl bei maschineller Übersetzung als auch beim Fremdspracherwerb ist eine produktive Behandlung der Phraseologie erforderlich, sozusagen eine präzise „Bauanleitung“ wie Phraseologismen richtig eingesetzt werden können (s. 4.2), und diese Problematik schließt auch Kollokationen ein.

Dementsprechend zeichnet sich in der jüngeren Phraseologieforschung eine Tendenz hin zur weiteren Begriffsauffassung ab, wie unter anderem Ausführungen von Lubensky/McShane (2007:919) anschaulich machen:

„*The rationale for discussing phrasemes instead of idioms per se is that pedagogical and practical (e.g., in natural language processing) experience has shown that knowledge of a broad range of (semi-)fixed entities is advantageous. The intersection between phrases, idioms and formulaic language has been discussed [...] and numerous recent works have supported Cowie's⁸ prediction that 'studies of collocations have pushed the boundary that roughly demarcates the 'phraseological' more and more into the zone formerly thought of as free, and it should not surprise us if future dictionaries consist of a higher proportion than before of collocations, idioms, and formulae' [...]. Thus, our broad definition of phraseme is in keeping with a practically-oriented, more inclusive trend in phraseological description.*“

Abgesehen von diesen praktischen Erwägungen spricht noch ein Aspekt der **begrifflichen Abgrenzung** für die weite Phraseologiauffassung: Der Begriff *Idiomatik* bezeichnet bereits diejenigen festen Wortverbindungen, die irgendeine Form von übertragener Bedeutung haben. Beschränkt man auch die Bezeichnung *Phraseologie* auf bildliche Ausdrücke, so fehlt ein Oberbegriff für sämtliche feste Wortverbindungen unabhängig vom Grad ihrer Idiomatizität; andererseits wäre eine feste und zugleich idiomatische Wortverbindung mit den Termini *Idiom* und *Phraseologismus* doppelt besetzt. Angesichts der hohen Praxisrelevanz der Kollokationen ist es jedoch ratsam, diese nicht einfach in der Grauzone problematischer Grenzfälle zwischen Phraseologie und freiem Wortgebrauch zu belassen, sondern sie begrifflich einzuschließen.

Aus den beschriebenen Gründen wird im Rahmen dieser Arbeit von einer weiten Definition der Phraseologie ausgegangen, die die Idiomatizität nicht zur notwendigen Bedingung erklärt.

⁸ Die Autorinnen zitieren Cowie (1998:20).

2.2.2 Definition auf Basis der Grundmerkmale

Nun soll die Frage erörtert werden, in welchem Verhältnis die oben beschriebenen Grundmerkmale Festigkeit, Polylexikalität und Idiomatizität zueinander stehen.

Festigkeit scheint ein notwendiges Kriterium für Phraseologizität zu sein, sofern man berücksichtigt, dass es regionale und idiolektale Abweichungen geben kann. Nach bisherigem Forschungsstand ist auch die Polylexikalität ein notwendiges Kriterium, was im Fall der Einwortphraseologismen eingeschränkt wird. In diesem Ausnahmefall muss aber jedenfalls das dritte Kriterium, das der Idiomatizität, gelten. Die Idiomatizität ist ansonsten keine notwendige Bedingung für das Vorliegen eines Phraseologismus.

Dies kann zu folgender Definition zusammengefasst werden:

Die Phraseologie einer Sprache ist die Menge der Ausdrücke, für die Festigkeit als notwendiges Kriterium sowie mindestens eines der beiden Kriterien Polylexikalität und Idiomatizität erfüllt ist. Phraseologie bezeichnet außerdem das Fachgebiet, das sich mit diesen Ausdrücken befasst.⁹

Auch in der gängigen Phraseologiedefinition von Burger/Buhofer/Salm (1982:1) sind die drei Merkmale implizit enthalten:

„Phraseologisch ist eine Verbindung von zwei oder mehr Wörtern dann, wenn (1) die Wörter eine durch die syntaktischen und semantischen Regularitäten der Verknüpfung nicht voll erklärbare Einheit bilden und wenn (2) die Wortverbindung in der Sprachgemeinschaft, ähnlich wie ein Lexem, gebräuchlich ist.“

Die Polylexikalität wird eingangs zur Voraussetzung erklärt, in den Kriterien (1) und (2) kommen verschiedene Aspekte der Festigkeit zur Sprache. Die Idiomatizität geht aus der Formulierung „eine durch die [...] semantischen Regularitäten der Verknüpfung nicht voll erklärbare Einheit“ hervor, die impliziert, dass der Sinn nicht aus der Bedeutung der Einzelkomponenten erschließbar ist.

⁹ Sollen die Einwortphraseologismen nicht berücksichtigt werden, müsste die Definition folgendermaßen abgewandelt werden: Die Phraseologie einer Sprache ist die Menge der Ausdrücke, für die Festigkeit und Polylexikalität als notwendige Bedingungen und Idiomatizität als fakultative Bedingung gelten.

2.2.3 Impulse aus der Korpuslinguistik

Die Korpuslinguistik hat in den letzten Jahren wesentliche Beiträge zur Phraseologieforschung geleistet, die das Phraseologieverständnis der theoretischen Forschung erweitern und zum Teil in Frage stellen.

Große Textkorpora können auf zweierlei Weise für die Phraseologieforschung nutzbar gemacht werden (vgl. Bubenhofer/Ptashnyk 2010:11ff; Sailer 2007): Erstens dienen sie der **Überprüfung** bestimmter Phänomene bzw. deren Nachweis anhand realer sprachlicher Daten. Bereits bekannte Phraseologismen werden also im Korpus gesucht, um Aussagen über deren Vorkommenshäufigkeit, Distribution etc. zu treffen. Zweitens – und dies ist der interessantere, aber auch weitaus komplexere Aspekt – werden mit Korpora **Datenanalysen** mit dem Ziel durchgeführt, Mehrworteinheiten und musterhaften Sprachgebrauch mittels statistischer Methoden zu identifizieren und zu kategorisieren.

Befunde beider Formen der Korpusanalyse lassen auf eine äußerst **unscharfe Grenze zwischen Phraseologie, freiem Wortgebrauch und strukturell bedingten Sprachverwendungsmustern** schließen und werfen somit die „nicht gelöste Frage nach den Grenzen des Phraseologischen“ (Bubenhofer/Ptashnyk 2010:18) auf.

Die Untersuchung bekannter Phraseologismen in Korpora hat ergeben, dass diese „in weit aus höherem Grade variabel sind, als man das früher annahm“ (Burger 2003:27), also häufig durch eine **hohe Zahl an Varianten** vertreten sind. Entsprechende Ergebnisse beschreibt u.a. Moon (2007b:1054f.). Somit ergibt sich eine nicht eindeutig zuordenbare Grauzone zwischen variantenreichen Phraseologismen und freien Wortverbindungen mit sehr spezifischen kombinatorischen Restriktionen.

Weiterhin fördern Korpusanalysen häufige Mehrworteinheiten zutage, die außerhalb des klassischen Untersuchungsgegenstandes der Phraseologie liegen, insbesondere strukturell bedingte **Sprachverwendungsmuster** wie das von Moon (2007b: 1048) angeführte Beispiel „with a mixture of“, das in Kombination mit diversen Wortpaaren wie „admiration and horror“, „relief and disappointment“, „horror and disbelief“ und anderen auftritt. Heid (2007:1036) nennt auch Mehrwortadverbien (frz. *petit à petit* – dt. *Stück für Stück*) und -präpositionen (*with respect to*), Routineformeln (*there you go; let's see; who knows?*) sowie fachsprachliche Phraseme. Aus diesen Gründen schlussfolgert Moon (2007b:1054): „corpus evidence blurs the distinction between recurrent lexico-grammatical patterns and institutionalized set phrases“.

Angesichts dieser Ergebnisse gilt es nun, einen **praktischen Ansatz** zu finden, wie weit der Phraseologiebegriff bei der Konzeption eines elektronischen Wörterbuchs gefasst werden soll. Bubenhofer/Ptashnyk (2010:18) zufolge gibt es „gute Gründe, das Phänomen breiter zu fassen und durch statistische Verfahren eine breite Palette von musterhaftem Sprachgebrauch abzudecken“. Auch Lubensky/McShane (2007:919) ar-

gumentieren: „pedagogical and practical (e.g., in natural language processing) experience has shown that knowledge of a broad range of (semi-)fixed entities is advantageous“.

Laut Heid (2007:1036) entspricht eine **weiter gefasste Phraseologiedefinition** der computerlinguistischen Praxis, auch weiteren Forscherinnen zufolge ist dieser Ansatz **praxisnäher** und wird den Erfordernissen maschineller Sprachverarbeitung eher gerecht (z.B. Bubenhofer/Ptashnyk 2010:18). Lubensky/McShane (2007:919) sprechen von einem „practically-oriented, more inclusive trend in phraseological description“. Diesem Trend schließt sich die vorliegende Arbeit an.

3 Phraseologische Klassen

Wie in Kapitel 2 dargelegt, baut diese Arbeit auf einer weit gefassten Phraseologiedefinition auf. Ein derart weites Feld verlangt nach einer internen Strukturierung; daher werden im Folgenden einige phraseologische Klassen vorgestellt, die den phraseographischen Untersuchungen in Kapitel 5 als Basis dienen. Am Schluss wird kurz die Frage nach einer Gesamtsystematik angeschnitten.

Traditionell werden Phraseologismen in Klassen wie Sprichwörter, Gemeinplätze, Geflügelte Worte, Redewendungen, Zwillingsformeln und komparative Phraseologismen eingeteilt. Diese Bezeichnungen sind zwar – zum Teil auch außerwissenschaftlich – weit verbreitet, eignen sich aber nicht für eine umfassende und homogene Systematik, weil diese Einteilung nicht auf einheitlichen Merkmalen beruht. So unterscheiden sich etwa Sprichwörter und Redewendungen darin, dass erstere aus ganzen Sätzen bestehen (*Wo ein Wille ist, ist auch ein Weg.*), letztere nur aus einem Satzglied oder unvollständigem Satz (*jdm. den Wind aus den Segeln nehmen*). Die Unterscheidung wird also auf syntaktischer Basis getroffen. Gemeinplätze sind ebenso wie Sprichwörter ganze Sätze, enthalten aber keine neue Erkenntnis, sondern eine Tautologie (*Was man hat, das hat man.*). Somit ist die Unterscheidung hier semantisch motiviert. Geflügelte Worte wiederum gehen im Gegensatz zu Sprichwörtern und Redewendungen auf eine belegbare Quelle zurück (*Ich bin ein Berliner*, Zitat von John F. Kennedy). Hier sind also außersprachliche Faktoren maßgeblich.

Durch diese Heterogenität der Merkmale können sich die traditionellen Klassen auch überlappen. So kann etwa ein Sprichwort gleichzeitig eine Zwillingsformel enthalten (*Mitgefangen, mitgehängt.*), ein Geflügeltes Wort kann zur Redewendung werden (*aus seinem Herzen eine Mördergrube machen*, Zitat aus der Bibel).

Ferner sind die traditionellen Klassen hauptsächlich im literarischen und gemeinsprachlichen Bereich von Interesse, eignen sich aber nicht zur Klassifizierung fachsprachlicher Kollokationen.

Festzuhalten ist also, dass Phraseologismen anhand von Merkmalen unterschiedlicher Domänen wie Syntax oder Semantik kategorisiert werden können. Die im Rahmen dieser Arbeit relevanten Merkmale werden in 3.1 kurz vorgestellt¹⁰. Abschnitt 3.2 behandelt die Frage, ob und wie sich auf Basis der Merkmale eine einheitliche Systematik erstellen lässt.

¹⁰ Eine umfassende Darstellung von Merkmalen findet sich bei Burger (2003:33ff.).

3.1 Verschiedene Merkmale

3.1.1 Einteilung nach semiotischen Merkmalen

Der semiotische Aspekt schafft eine sinnvolle Grobeinteilung der Fülle von Phraseologismen. Es lassen sich vier unterschiedliche Zeichenfunktionen feststellen:

Referentielle Phraseologismen beziehen sich auf Objekte, Vorgänge oder Sachverhalte in der außersprachlichen Wirklichkeit oder einer fiktiven Welt. So beschreibt das Sprichwort *Der Apfel fällt nicht weit vom Stamm* den angenommenen Sachverhalt, dass Kinder ihren Eltern charakterlich ähneln, *der blaue Dunst* umschreibt ein Ding, nämlich Zigarettenrauch.

Strukturelle Phraseologismen hingegen beziehen sich auf nichts Außersprachliches, sondern stellen einen logischen Bezug zwischen verschiedenen sprachlichen Elementen her. Beispiele dafür sind *entweder – oder, weder – noch, einerseits – andererseits, sowohl – als auch, da ja, zum Beispiel*.

Kommunikative Phraseologismen oder **Routineformeln** beziehen sich ebenfalls auf nichts Äußeres, sondern stellen eine sprachliche Handlung dar oder markieren eine solche. Hierzu zählen alle Begrüßungsformeln, aber auch feste Formeln wie *meines Erachtens*.

Onymische Phraseologismen werden vielerorts nicht der Phraseologie zugeordnet. Will man es aber tun, müssen sie hier unter den semiotischen Klassen angeführt werden. Es handelt sich hierbei um Phraseologismen, die die Funktion von Eigennamen tragen, wie etwa *der Nahe Osten, das Rote Kreuz, die Vereinten Nationen*.

3.1.2 Einteilung nach semantischen Merkmalen

In semantischer Hinsicht ist besonders die in 2.1.3 beschriebene Einteilung in **Idiome**, **Teil-Idiome** und **Kollokationen** von Bedeutung, also die Unterscheidung, ob der Phraseologismus vollständig, teilweise, oder überhaupt nicht im übertragenen Sinne gemeint ist. Hierbei handelt es sich wie erwähnt nur um eine Grobeinteilung. Häufig anzutreffen sind etwa Phraseologismen, die zwar vollständig idiomatisch sind, jedoch nur über einen schwachen Idiomatizitätsgrad verfügen, wie *jdn. keines Blickes würdigen*, dessen phraseologische Bedeutung nicht weit entfernt vom gemeinten *jdn. ignorieren* liegt.

Bei Idiomen ließen sich weitere Einteilungen vornehmen nach Art der **Beziehung**, die zwischen beiden Lesarten besteht (Metapher, Metonymie etc.) sowie nach dem **Herkunftsreich des Bildes** (z.B. die Meeresmetaphorik in *hohe Wellen schlagen*; *Wir sitzen alle in einem Boot, das Ruder in die Hand nehmen*).

Gläser (1986) differenziert außerdem zwischen **nominativen** und **propositionalen Phraseologismen**. Nominativ sind Phraseologismen, die Objekte oder Vorgänge bezeichnen; propositional nennt man solche, die Aussagen über Objekte und Vorgänge treffen. Erstere entsprechen demnach einzelnen Satzgliedern, letztere ganzen Sätzen. Diese Unterscheidung deckt sich somit mit der syntaktischen Einteilung in satzwertige und satzgliedwertige Phraseologismen.

Im konnotativen Bereich kann zwischen verschiedenen **Stilschichten** unterschieden werden, also danach, ob ein Phraseologismus Slang, Umgangssprache, Standardsprache, gehobener Sprache etc. zuzuordnen ist.

3.1.3 Einteilung nach syntaktischen Merkmalen

Bei der syntaktischen Klassifikation muss unterschieden werden, ob vom inneren Aufbau des Phraseologismus die Rede ist (s. 3.1.3.4), oder von der Rolle des gesamten Ausdrucks im Satzgefüge (s. 3.1.3.1). Letztere ist für die Phraseographie weitaus wichtiger. An die syntaktische Funktion des Phraseologismus knüpfen sich auch Fraugen der syntaktischen Kontextanbindung an (s. 3.1.3.2 und 3.1.3.3).

3.1.3.1 Syntaktische Klasse

Ein Phraseologismus als Ganzes kann die Rolle unterschiedlicher **Satzglieder** einnehmen; diese werden im Folgenden besprochen. Als erste Grobunterteilung kann man **satzwertige** und **satzgliedwertige** Phraseologismen unterscheiden.

Satzwertig sind Phraseologismen, die einem ganzen Satz entsprechen. Dies ist beispielsweise bei Sprichwörtern und Gemeinplätzen normalerweise der Fall. Manche Phraseologismen enthalten zwar kein Verb, können aber dennoch als satzwertiges Gebilde betrachtet werden, weil sie in sich abgeschlossen sind und keiner weiteren Ergänzung bedürfen:

Auge um Auge, Zahn um Zahn.

Ende gut, alles gut.

Besser spät als nie.

Mitunter enthalten satzwertige Phraseologismen eine oder mehrere Leerstellen, die aktualisiert werden müssen, wie das Dativobjekt in *[Jdm] geht ein Licht auf* oder die adverbiale Ergänzung in *Der Zahn der Zeit nagt an [etw]*. Dennoch können diese Phraseologismen als satzwertig angesehen werden, da sie zumindest Subjekt und Prädikat enthalten. Andere satzwertige Phraseologismen – wie in der Regel Sprichwörter – bilden bereits eine geschlossene Einheit und müssen syntaktisch (und zum Teil auch semantisch) nicht an den Kontext angebunden werden. Die Frage der Kontextanbindung wird in 3.1.3.2 näher besprochen.

Die **satzgliedwertigen Phraseologismen** lassen sich in weitere Untergruppen gliedern, je nach dem, welchem Satzglied sie entsprechen. Hierbei wird bewusst nur nach Konstituententyp¹¹, nicht aber nach Satzgliedfunktion klassifiziert, da ein Phraseologismus unterschiedliche Funktionen im Satz einnehmen kann (vgl. hierzu die Funktion von *faule Ausreden* in *Faule Ausreden würden ihn nicht weiterbringen* vs. *Und komm mir nicht mit faulen Ausreden*).

Nomiale Phraseologismen stellen eine Nominalphrase dar, die im Satz Subjekt oder Objekt, Prädikativ und mitunter auch Attribut sein kann: *die Qual der Wahl, ein Blick über den Tellerrand, ein Fass ohne Boden, das Weiße Haus, die sterbliche Hülle*. Ein nominaler Phraseologismus kann zudem Bestandteil einer übergeordneten Konstituente, z.B. einer Adverbialphrase, sein.

Verbale Phraseologismen entsprechen einer Verbalphrase, enthalten also ein Prädikat und ggf. Objekte und Ergänzungen, wie z.B. *das Zeitliche segnen, ein Urteil fällen, aus einer Mücke einen Elefanten machen, das Gras wachsen hören, das Blaue vom Himmel herunter lügen, auf der Leitung stehen*.

Ferner gibt es noch die **adjektivischen Phraseologismen**, wobei zwischen der attributiven und der prädikativen Funktion zu unterscheiden ist. Die meisten auf den ersten Blick adjektivischen Phraseologismen sind nur als Prädikativ verwendbar, nicht aber als Attribut. Eigentliche adjektivische, d.h. attributiv verwendbare Phraseologismen sind im Deutschen selten; ein Beispiel wäre *frisch gebacken*, z.B. in *ein frisch gebackener Ehemann*. Nur prädikativ verwendbare Phraseologismen hingegen finden sich zahlreich:

Das Kind ist schlau. –

Das schlaue Kind geht in die Schule.

Das Kind ist schlau wie ein Fuchs. –

--

Weitere Beispiele sind *grün hinter den Ohren, mit allen Wassern gewaschen, gang und gäbe*. Diese Ausdrücke können nicht zu den adjektivischen Phraseologismen zählen, weil sie nicht in jeder Position eines Adjektivs verwendbar sind. Eine Möglichkeit wäre die Zuordnung zu den verbalen Phraseologismen in Verbindung mit dem Verb *sein*; noch geeigneter ist aber die Einordnung in einer eigenen Klasse von **prädiktiven Phraseologismen**, da für die Aktualisierung des Verbs unter Umständen mehrere Möglichkeiten bestehen (*gelten als, werden etc.*).

Schließlich gibt es noch **adverbiale Phraseologismen**, die die Funktion einer Ergänzung innehaben. Beispiele hierfür sind *mit allen Mitteln der Kunst, im Eifer des Gefechts*.

¹¹ Gemäß der Terminologie der Konstituentenstrukturgrammatik, vgl. z.B. Pelz (⁹2005), Lyons (⁸1995).

Präpositionale Phraseologismen wie *an Hand von* oder *in Bezug auf* erfüllen die Funktion einer Präposition.¹²

Weiteren Wortarten entsprechende Klassen, wie etwa konjunktionale oder interktionale Phraseologismen, sind denkbar. Auch können sprachspezifisch weitere Klassen auftreten, etwa die im Ungarischen häufigen Postpositionen.

Die Zuordnung eines Phraseologismus zu einer syntaktischen Klasse ist nicht in jedem Fall eindeutig; dazu muss ein weiterer Faktor berücksichtigt werden, nämlich die externen Leerstellen. Diese werden in 3.1.3.3 besprochen. Zunächst folgt ein allgemeiner Überblick über mögliche Formen der Kontextanbindung.

3.1.3.2 Kontextanbindung

Bei der Frage, wie Phraseologismen in den umgebenden Text eingebunden sind, kann zunächst zwischen zwei großen Gruppen unterschieden werden, nämlich einmal Phraseologismen, die grundsätzlich für sich alleine stehen und auch unabhängig von einem bestimmten Kontext einen Sinn ergeben. Zum anderen gibt es Phraseologismen, die in irgendeiner Form – sei es syntaktisch oder auch nur auf der logischen Ebene – an den umgebenden Text angebunden werden müssen. Auch Burger (2003:39f.) nimmt eine solche Einteilung vor, jedoch bezieht er diese nur auf satzwertige Phraseologismen; daher soll seine Einteilung hier modifiziert werden.

Bei der ersten Gruppe, die hier als **selbständige Phraseologismen** bezeichnet werden soll, kann es sich per definitionem nur um satzwertige Phraseologismen handeln. Burger (1998:40) spricht von **Topischen Formeln**, also „satzwertige Formulierungen [...], die durch kein lexikalisches Element an den Kontext angeschlossen werden müssen.“ Sie enthalten „generalisierende Aussagen, die auch ohne Verankerung in einem spezifischen Kontext [...] verständlich sind“ (ebd). Hierzu zählt Burger in erster Linie Sprichwörter und Gemeinplätze, wie *Morgenstund hat Gold im Mund* oder *Wo ein Wille ist, ist auch ein Weg*. Aber auch andere Klassen, vor allem manche Geflügelten Worte, fallen unter die selbständigen Phraseologismen.

Die zweite Gruppe, die hier analog **unselbständige Phraseologismen** genannt werden soll, umfasst alle Phraseologismen, bei denen irgendeine Form von Textanbindung stattfindet. Diese kann auf unterschiedliche Weise erfolgen. Burger spricht von **festen Phrasen**, bezieht sich damit jedoch nur auf satzwertige Phraseologismen, bei denen eine Anbindung nötig ist. Der Terminus *unselbständige Phraseologismen* wird

¹² Von einer *Präpositionalphrase* hingegen kann hier nicht gesprochen ist, da dieser Begriff in Konstituentenstrukturgrammatiken für Adverbialphrasen verwendet wird, die durch eine Präposition eingeleitet werden (vgl. Pelz⁹2005:155; Hutchins/Somers 1992:23f.)

hier als Überbegriff eingeführt für feste Phrasen sowie alle satzgliedwertigen Phraseologismen, die ja naturgemäß syntaktisch an den Kontext angebunden werden müssen. In Bezug auf die festen Phrasen beschreibt Burger drei Arten der Kontextanbindung, die ohne Weiteres auch für unselbständige Phraseologismen im Allgemeinen übernommen werden können.

Die meisten Phraseologismen werden über zu aktualisierende **Satzteile** an den Kontext angebunden. Dies ist grundsätzlich bei allen satzgliedwertigen Phraseologismen der Fall, aber auch bei einigen satzwertigen.

[Jdn] hat die Muse geküsst. → Akkussativobjekt

[Jdm] fällt die Decke auf den Kopf. → Dativobjekt

[Jd] stellt [jdm] [etw] in Aussicht. → Subjekt, Dativobjekt, Akkussativobjekt

In vielen Fällen erfolgt die Anbindung über ein **deiktisches Element**, das auf die Situation, einen Ort oder auf eine zuvor erwähnte Tatsache bezieht. Beispiele sind:

Schreib dir das hinter die Ohren!

Bis dahin fließt noch viel Wasser die Donau hinunter.

[Da] ist der Wurm drin.

Da ist was im Busch.

Da ist was faul.

Da geht die Post ab.

Dieses deiktische Pronomen ist jedoch nicht immer als fester Bestandteil des Phraseologismus zu sehen. Die eckigen Klammern zeigen an, dass an seiner Stelle auch eine zeitliche oder räumliche Ergänzung eingesetzt werden kann, z.B.

Also heute ist irgendwie der Wurm drin.

Gestern auf der Party ging ziemlich die Post ab!

An dem Angebot ist etwas faul.

Manche Fälle der deiktischen Anbindung sind daher ein Sonderfall der Anbindung durch einen Satzteil (Ergänzung).

Weiterhin kann ein Phraseologismus auch nur durch den **logischen Zusammenhang** in den Kontext eingebettet werden, ohne dass sich dies formal niederschlägt.

Die Luft ist rein.

Einmal ist keinmal.

Die Wände haben Ohren.

Der Schuss geht nach hinten los.

Es steht viel auf dem Spiel.

Grüner wird's nicht.

Der Zug ist schon abgefahren.

Die rein logische Anbindung kennzeichnet sich häufig dadurch, dass der Phraseologismus oder eine seiner Komponenten metaphorisch oder metonymisch auf einen äußereren Sachverhalt verweist und so den deiktischen Hinweis ersetzt. So steht *der Schuss*, der nach hinten losgeht, für eine Handlung, die das Gegenteil der gewünschten Wirkung zeitigt, der abgefahrene *Zug* für eine verpasste Gelegenheit.

3.1.3.3 Interne und externe Leerstellen

Wie oben beschrieben, kann ein Phraseologismus durch externe Leerstellen an den Kontext angebunden sein. Bei einem verbalen Phraseologismus spricht man hierbei von **externer Valenz** im Gegensatz zur **internen Valenz**, d.h. den bereits durch Komponenten des Phraseologismus selbst besetzten Leerstellen des Verbs.

| Extern: | Subjekt | | Objekt _{Dat} | |
|---------|-----------------------------|--------------|-----------------------|--|
| | [<i>Die Phraseologie</i>] | <i>macht</i> | [<i>Lernenden</i>] | <i>einen Strich</i> <i>durch die Rechnung.</i> |
| Intern: | Prädikat | | Objekt _{Akk} | Ergänzung |

Tabelle 1: Interne und externe Valenz

Intern sind diejenigen Elemente, die schon im Phraseologismus enthalten sind; extern sind solche, die entsprechend dem Kontext aktualisiert werden.

Der Begriff der externen Leerstellen kann auf nicht verbale Phraseologismen ausgeweitet werden. Auch bei anderen Phraseologismen erfordert die Einbettung in den syntaktischen Kontext die Ergänzung bestimmter Satzglieder:

[*Jdm*] fällt ein Stein vom Herzen → Dativobjekt
Mit [jdm] kann man Pferde stehlen → N_{Dat}
an Hand von [etw] → N_{Dat} (Subjekt, Prädikat, Akkussativobjekt)

Externe Leerstellen können **fakultativ** oder **obligatorisch** sein:

Obligatorisch sind Leerstellen, deren Aktualisierung zur syntaktischen Einbettung zwingend erforderlich ist. Hierbei muss jedoch unterschieden werden zwischen Leerstellen, die zur Erzeugung einer **vollständigen Konstituente**, und solchen, die zur Erzeugung eines **vollständigen Satzes** notwendig sind. Diese Abstufung ist z.B. bei Nominalphrasen wichtig, die unter Umständen auch alleine stehen können, z.B. als Beschriftung einer Graphik oder als Artikelüberschrift. Das Prädikat kann daher nicht pauschal als obligatorisch gesehen werden. Dasselbe gilt auch für andere Satzglieder, z.B. Adverbialphrasen. Die genannte Unterscheidung spielt noch aus weiteren Gründen eine Rolle, von denen weiter unten die Rede ist.

Fakultative externe Leerstellen sind solche, die an einer gegebenen Stelle ergänzt werden können, aber nicht müssen. Da die Liste möglicher Ergänzungen naturgemäß

endlos lang ist, sind hierbei nur diejenigen fakultativen Ergänzungen von Interesse, die in dem Kontext nicht zwingend, aber typischerweise anzutreffen sind.

Bei der Kategorisierung von Phraseologismen stellen die externen Leerstellen eine zusätzliche Schwierigkeit dar: Es muss entschieden werden, ob externe Leerstellen bei der syntaktischen Klassifikation der Phraseologismen mitberücksichtigt werden soll oder nicht (vgl. Burger 2003:42). Die Frage, ob man die externe Leerstelle mit zur Struktur des Phraseologismus rechnet oder nicht, ist in manchen Fällen ausschlaggebend für die syntaktische Klassenzuordnung, weil man jeweils ein unterschiedliches Ergebnis erhalten würde. So hat der strukturelle Phraseologismus *an Hand von [etwas]* ohne die externe Leerstelle *[etwas]* präpositionalen Charakter; rechnet man die Leerstelle bei der Klassenzuordnung jedoch mit, ergibt sich eine vollständige adverbiale Ergänzung, und man müsste somit von einem adverbialen Phraseologismus sprechen.

In anderen Fällen wäre eine Zuordnung unter Einbeziehung der externen Leerstelle sinnvoll, um eine komplizierte Umschreibung der syntaktischen Klasse zu vermeiden. So könnten Ausdrücke wie

[Jdm] fällt ein Stein vom Herzen.
In [jds] Haut möchte ich nicht stecken!
Mit [jdm] kann man Pferde stehlen.

als ganze Sätze aufgefasst werden, bei denen jeweils eine Leerstelle zu aktualisieren ist. Wäre eine Einbeziehung der externen Leerstelle nicht zulässig, müsste die syntaktische Klasse von *[Jdm] fällt ein Stein vom Herzen* etwa umschrieben werden als

{V + NP(Det + N) + PP(Präp + N_{Akk})}

Auf diese Weise müsste für jede erdenkliche Kombination von Satzgliedern eine eigene Klasse eröffnet werden, was nicht zweckmäßig erscheint. Dies spricht in solchen Fällen für eine Einbeziehung der externen Leerstellen in die syntaktische Klassenbildung, ebenso wie die Tatsache, dass die externe Leerstelle häufig das logische Subjekt enthält (*Ihm fällt ein Stein vom Herzen → Er ist erleichtert.*).

Betrachtet man dabei jedoch all jene Leerstellen als obligatorisch, die zur Erzeugung eines vollständigen Satzes aktualisiert werden müssen, ist deren Einbeziehung nicht zielführend, da die syntaktische Klassifikation auf diese Weise nur ganze Sätze ergeben würde. Hier kommt nun die oben beschriebene Unterscheidung zwischen der Aktualisierung zum vollständigen Satz und der zur nächst höheren Konstituente zum Tragen: nur letztere können bei der syntaktischen Klassifikation einbezogen werden. Es wird daher folgendes Modell vorgeschlagen, das als **Aufrundungsprinzip** bezeichnet werden soll:

All jene externen Leerstellen, die zur Vervollständigung der nächst höheren Konstituente benötigt werden, werden bei der syntaktischen Klassifikation hinzugerechnet, alle anderen nicht.

An Hand von [etw] würde somit als präpositionaler Phraseologismus gewertet werden, da auch ohne Hinzuziehung der externen Leerstelle [etw] schon eine komplette Konstituente vorliegt. [Jdm] fällt ein Stein vom Herzen hingegen zählt als Satz mit Leerstelle, um aus Einzelkonstituenten zusammengesetzte Klassenbezeichnungen zu vermeiden.

Wenn im Folgenden von der syntaktischen Klasse eines Phraseologismus die Rede ist, wurde nach dem beschriebenen Prinzip verfahren.

3.1.3.4 Innere morphosyntaktische Struktur

Der innere Aufbau eines Phraseologismus lässt sich über eine Konstituentenanalyse ermitteln, wie folgende Beispiele zeigen. Externe Leerstellen wurden durch eckige Klammern gekennzeichnet. Klammern im Index bedeuten, dass die betreffende Konstituente eine Leerstelle enthält.

[Jdm] platzt der Kragen.

- (1) $\Sigma_{()} \rightarrow NP + VP_{()}$
 - (2a) $NP \rightarrow Det + N$
 - (2b) $VP_{()} \rightarrow V + [NP_{Dat}]$
- $$\Rightarrow \Sigma_{()} \{VP_{()} ([NP_{Dat}] + V) + NP (Det + N)\}$$

[Jdn] auf frischer Tat ertappen

- (1) $VP_{()} \rightarrow [NP_{Akk}] + PP + V$
 - (2) $PP \rightarrow Präp + [NP_{Dat}]$
 - (3) $NP_{Dat} \rightarrow A + N$
- $$\Rightarrow VP_{()} \{[NP_{Akk}] + PP (Präp + NP_{Dat} (A + N)) + V\}$$

[Jdn] überfällt eine bleierne Müdigkeit.

- (1) $\Sigma_{()} \rightarrow NP + VP_{()}$
 - (2a) $NP \rightarrow Det + NP$
 - (2b) $VP_{()} \rightarrow V + [NP_{Akk}]$
 - (3) $NP \rightarrow A + N$
- $$\Rightarrow \Sigma_{()} \{VP_{()} ([NP_{Akk}] + V) + NP (Det + NP (A + N))\}$$

Als Klassifikationskriterium eignet sich der innere Aufbau nicht, da sich völlig unterschiedliche Kombinationen aus Konstituenten ergeben können.

Abgesehen von der formalen Analyse des inneren Aufbaus finden sich in Phraseologismen häufig **wiederkehrende Strukturen**, die ein Kriterium zur Klassenbildung

darstellen können. So weisen manche Phraseologismen eine paarige Struktur (Zwillingssformel; *gesund und munter*), in anderen finden sich rhetorische Figuren wie Reim (*Morgenstund hat Gold im Mund*), Vergleich (*reden wie ein Wasserfall*), Alliteration (*aus heiterem Himmel*), Gemination (*Morgen, morgen, nur nicht heute...*) oder Parallelismus (*Wie gewonnen, so zerronnen*).

Auch die Frage, ob ein **Einwortphraseologismus** vorliegt (s. 2.1.2), betrifft den inneren Aufbau. Hinsichtlich der syntaktischen Klasse sind Einwortphraseologismen zu meist Nominalphrasen; der Unterschied besteht allein im inneren Aufbau.

3.1.4 Weitere Merkmale

Für die Charakterisierung von Phraseologismen sind zahlreiche weitere Merkmale relevant, die sich aufgrund ihrer Differenziertheit nicht unbedingt zur Klassenbildung eignen. Zu nennen wären diesbezüglich vor allem **Stilschicht**, **areale Verbreitung**, **Gebrauchskontext** und weitere funktionale Aspekte, ebenso wie **semantische und morphosyntaktische Restriktionen** bei der Kontextanbindung. Eine detailliertere Beschreibung dieser Merkmale folgt in 5.4 im Zusammenhang mit Angaben, die in einem phraseologischen Wörterbuch benötigt werden.

3.2 Systematik

In 3.1 wurde eine Reihe von Merkmalen erarbeitet, anhand derer sich die Phraseologismen klassifizieren lassen. Neben den traditionellen, teilweise auch außerwissenschaftlich bekannten Klassen gibt es noch eine Vielzahl an semiotischen, semantischen und syntaktischen Klassifikationskriterien.

Nun soll die Frage aufgegriffen werden, wie diese Merkmale zueinander in Bezug gesetzt werden können und ob sich auf deren Basis eine übergreifende Systematik erstellen lässt. Burger (?2003:36ff.) schlägt eine Systematik auf Basis der beschriebenen Merkmale vor. Diese wird im Folgenden vorgestellt.

3.2.1 Basisklassifikation nach Burger (?2003)

Abb. 1 zeigt eine Übersicht über die von Burger (?2003) vorgeschlagene *Basisklassifikation*¹³.

¹³ Bei Burger wird die Klassifikation nur in reduzierter Form als Graphik dargestellt, daher wurde der Strukturbaum hier um Elemente erweitert, die Burger im Text beschreibt.

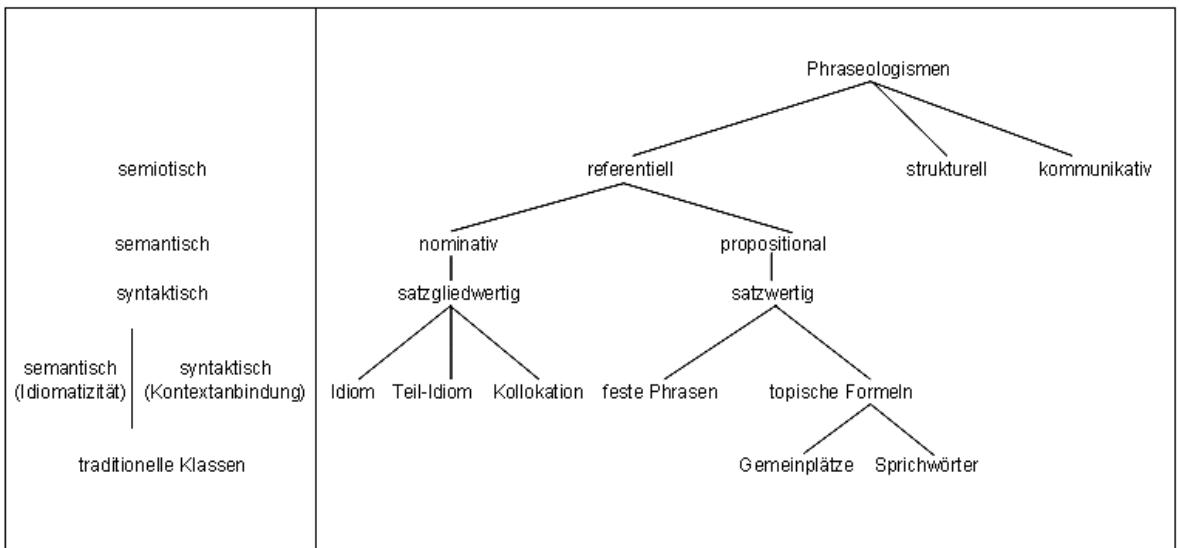


Abb. 1: Basisklassifikation nach Burger (2003)

Wie aus Abb. 1 ersichtlich ist, kombiniert Burger unterschiedliche Merkmale. In der Spalte links ist das für die jeweilige Gliederungsebene ausschlaggebende Merkmal angeführt.

Durch die erste Einteilungsebene nach semiotischen Gesichtspunkten werden zunächst die unübersichtlich große Gruppe der kommunikativen Phraseologismen sowie die strukturellen Phraseologismen abgespalten. Die semantische Einteilung in nominative und propositionale Phraseologismen deckt sich mit der syntaktischen Einteilung in *satzgliedwertig* und *satzwertig*. Die *satzwertigen* Phraseologismen subklassifiziert Burger nicht – wie die *satzgliedwertigen* – hinsichtlich ihrer Idiomatizität, sondern hinsichtlich der Art der Kontextanbindung, um die traditionelle Gruppe der Sprichwörter – die ja idiomatisch sein können oder auch nicht¹⁴ – nicht auseinander zu reißen.

Burgers Systematik liefert einen sehr guten ersten Überblick über die Materie. Es zeigt sich jedoch, dass eine hierarchische Anordnung generell nicht geeignet ist, das Gebiet der Phraseologie erschöpfend darzustellen, da durch die Inhomogenität der Kriterien einerseits Überlappungen unvermeidbar sind, während andererseits manche Klassen unberücksichtigt bleiben müssen.

Auf der untersten Ebene nimmt Burger die traditionellen Klassen Sprichwörter und Gemeinplätze in die Systematik auf. Jedoch lassen sich keineswegs alle traditionellen Klassen auf gleiche Weise einfügen. Bei der Klasse der Geflügelten Worte etwa ergäbe sich die Schwierigkeit, dass diese sowohl *satzwertig* (*Ich denke, also bin ich* von

¹⁴ Vgl. etwa Der Apfel fällt nicht weit vom Stamm (Idiom) vs. Kommt Zeit, kommt Rat (Kollokation).

Decartes¹⁵ oder *Jeder soll nach seiner Façon selig werden* von Friedrich dem II. von Preußen¹⁶ als auch satzgliedwertig (*aus seinem Herzen eine Mördergrube machen, Prophet im eigenen Land sein*, Bibelzitate) sein können. Unter den satzwertigen Geflügelten Worten gibt es sowohl kontextunabhängige topische Formeln als auch feste Phrasen, die syntaktisch oder logisch an den jeweiligen Kontext angebunden sind (*Das also ist des Pudels Kern!, Die Würfel sind gefallen*). Diese Klasse liegt also quer zur beschriebenen Systematik und wird von Burger daher, nebst einigen anderen, separat beschrieben.

Weiterhin stellt sich natürlich auch bei den satzwertigen Phraseologismen die Frage nach der Idiomatisität, und nicht nur bei den satzgliedwertigen. Die Einteilung in *satzwertig* und *satzgliedwertig* müsste ebenso bei den kommunikativen Phraseologismen vorgenommen werden (*Guten Abend vs. meines Erachtens*).

In solchen Unvollständigkeiten und Überschneidungen zeigt sich die Schwäche einer hierarchischen Darstellung der Phraseologie. Hier begegnet man dem von Isenberg (1978:578) beschriebenen *typologischen Dilemma*, nämlich der generellen Schwierigkeit, eine mehrdimensionale Materie in einem monohierarchischen System darzustellen¹⁷. Die Schwächen einer hierarchischen Klassifikation röhren daher, dass eine Hierarchie der Merkmale nicht inhärent gegeben ist, sondern stets einer gewissen Willkür unterliegt.

3.2.2 Darstellung im Begriffsfeld

Geeigneter scheint eine Darstellungsweise, die dem Nebeneinander der Klassen Raum gibt. Abb. 2 zeigt ein Begriffsfeld, in dem die auf unterschiedlichen Merkmalen beruhenden Klassen lose angeordnet sind. Es sei noch einmal erwähnt, dass die Aufzählung der Merkmale keineswegs vollständig ist, sondern bei Bedarf ergänzt werden muss.

Eine Hierarchie liegt nicht in der Natur der Merkmale, kann aber bei Bedarf je nach Forschungsinteresse eingeführt werden. Für das praktische Arbeiten scheint es daher sinnvoll einen Bestand in einer Datenbank zu ordnen, die nach unterschiedlichen Merkmalen abgefragt werden kann.

¹⁵ vgl. <http://books.google.de/books?id=XTBbrOvXlxMC&pg=PA7> (10. Oktober 2010)

¹⁶ Vgl. http://de.wikipedia.org/wiki/Liste_gefl%C3%BCgelter_Worte/J (10. Oktober 2010)

¹⁷ Genauer besteht das typologische Dilemma „darin, daß [...] die Forderungen nach Homogenität, Exhaustivität und Monotypie nicht gleichzeitig erfüllbar sind“ (Isenberg 1978:578).

| Bezugsrahmen | Phraseologische Klassen | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------------------------------------|--|--|---------------|-------|--|----------------|------------|-----------------|--|--|---------------------|--------|-------------------|---------------|--|-------------------|---------------|-----------------|--|--|-------------------|--|---------------------------------------|--|--|-------------------|--|--|-------|-------|-------------|--|--|--|--|------------------------------------|------------------------|-------------|--|--|--|--|--|--|--|
| semiotisch | kommunikativ strukturell referentiell onymisch | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| semantisch | <table border="0"> <tr> <td>Referenzart</td> <td>nominativ</td> <td>propositional</td> <td></td> <td></td> </tr> <tr> <td>Idiomatizität</td> <td>Idiome</td> <td>Teil-Idiome</td> <td>Kollokationen</td> <td></td> </tr> </table> | | | | | Referenzart | nominativ | propositional | | | Idiomatizität | Idiome | Teil-Idiome | Kollokationen | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Referenzart | nominativ | propositional | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Idiomatizität | Idiome | Teil-Idiome | Kollokationen | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| syntaktisch | <table border="0"> <tr> <td>Satzwertigkeit</td> <td>satzwertig</td> <td>satzgliedwertig</td> <td></td> <td></td> </tr> <tr> <td>syntaktische Klasse</td> <td></td> <td>NP VP AP PP (...)</td> <td></td> <td></td> </tr> <tr> <td>Kontext-anbindung</td> <td>selbstständig</td> <td>unselbstständig</td> <td></td> <td></td> </tr> <tr> <td>Selbstständigkeit</td> <td></td> <td>logisch deiktisch externe Leerstellen</td> <td></td> <td></td> </tr> <tr> <td>Art der Anbindung</td> <td></td> <td></td> <td>(...)</td> <td>(...)</td> </tr> <tr> <td>ext. Valenz</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>innere morphosyntaktische Struktur</td> <td>unigraph (Einwortphr.)</td> <td>polylexikal</td> <td></td> <td></td> </tr> <tr> <td></td> <td></td> <td>Zwillingssformel Parallellismus Reim (...)</td> <td></td> <td></td> </tr> </table> | | | | | Satzwertigkeit | satzwertig | satzgliedwertig | | | syntaktische Klasse | | NP VP AP PP (...) | | | Kontext-anbindung | selbstständig | unselbstständig | | | Selbstständigkeit | | logisch deiktisch externe Leerstellen | | | Art der Anbindung | | | (...) | (...) | ext. Valenz | | | | | innere morphosyntaktische Struktur | unigraph (Einwortphr.) | polylexikal | | | | | Zwillingssformel Parallellismus Reim (...) | | |
| Satzwertigkeit | satzwertig | satzgliedwertig | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| syntaktische Klasse | | NP VP AP PP (...) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Kontext-anbindung | selbstständig | unselbstständig | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Selbstständigkeit | | logisch deiktisch externe Leerstellen | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Art der Anbindung | | | (...) | (...) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ext. Valenz | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| innere morphosyntaktische Struktur | unigraph (Einwortphr.) | polylexikal | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | Zwillingssformel Parallellismus Reim (...) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Abb. 2: Phraseologische Klassen im Begriffsfeld

4 Phraseologie im interlingualen Arbeiten

In den vorangegangenen Kapiteln wurden Grundmerkmale sowie verschiedene Klassen der Phraseologie vorgestellt. Die Betrachtungen bezogen sich dabei nur auf den einsprachigen Bereich. Nun soll untersucht werden, welche Rolle der Phraseologie beim mehrsprachigen Arbeiten, also in der Translation oder beim Fremdspracherwerb, zukommt. Bevor wir uns hierbei auf die Mikroebene des Textes begeben, soll zunächst ein kurzer Blick auf den translatorischen Gesamtkontext geworfen werden.

4.1 Phraseologie im translatorischen Gesamtkontext

In der Translationswissenschaft besteht weitreichend Konsens darüber, dass für die Produktion eines geeigneten Zieltexes der **Skopos**, der Zweck der Translation, ausschlaggebend ist (vgl. z.B. Göpferich 2003). Es geht demnach nicht darum, einen Ausgangstext möglichst getreu zu übersetzen, sondern vielmehr um die Produktion eines Textes, der im zielkulturellen Kontext die gewünschte kommunikative Leistung erbringt, die sich mit der ursprünglichen Funktion des Ausgangstextes decken kann, aber nicht muss. Auch bei derselben Kommunikationsabsicht können kulturelle Unterschiede und abweichende Textsortenkonventionen die Wahl anderer sprachlicher Mittel erforderlich machen, so dass der Zieltext unter Umständen formal und inhaltlich völlig vom Ausgangstext abweicht. In einem solchen Fall wird auch die Übersetzung einzelner phraseologischer Wendungen des Ausgangstextes hinfällig. Erster Schritt einer translatorischen Handlung ist also zunächst die Beurteilung der kommunikativen Gesamtsituation.

Über einzelnen sprachlichen Elementen wie Phraseologismen steht zudem die übergeordnete Einheit **Text**, die u.a. durch den beschriebenen Gesamtkontext definiert ist und auf vielfache Weise mit Phraseologismen wechselwirkt. Beispielsweise beeinflussen Faktoren wie Textsortenkonventionen, einheitliche Textgestaltung, Stil etc. die Wahl sprachlicher Mittel wie Phraseologismen. Umgekehrt können Phraseologismen als Marker einer bestimmten Textsorte dienen, z.B. feste Wendungen in einem Vertrag oder einer Gebrauchsinformation, und somit auf die übergeordnete Einheit *Text* rückwirken. Eine Wechselwirkung zwischen Text und Phraseologie kann zudem dadurch stattfinden, dass ein Phraseologismus durch den Text, in den er eingebettet ist, neu geprägt und implizit oder explizit definiert wird. Ein Paradebeispiel für eine solche Neuprägung sind Gesetzestexte, die naturgemäß bestimmte Ausdrücke definieren. Prinzipiell kann jede Textsorte neue Wendungen und Lexeme prägen. Aus solchen textspezifisch und idiolektal entstandenen Wendungen werden häufig feste Phraseologismen, beispielsweise in Form von Geflügelten Worten, wie Zitate aus der Werbe-

sprache oder einer Politikerrede (*Ich bin ein Berliner*¹⁸, *L'état, c'est moi*¹⁹, *Soll jeder nach seiner Façon selig werden*²⁰).

Somit steht ein Phraseologismus nicht nur mit dem direkt umgebenden Text in Beziehung. Vielmehr kann er als Mikrotext aufgefasst werden, der ins intertextuelle Gesamtgefüge einer Kultur eingebettet ist. Das Beispiel der im Amerikanischen viel zitierten Wendung *That's what she said* mag dies veranschaulichen: Der Satz stammt aus der amerikanischen Fernsehkomödie *Saturday Night Live* und wurde vor allem durch die Serie *The Office* bekannt. Typischerweise wird die Wendung als Antwort auf eine unbeabsichtigt doppelbödige Aussage verwendet²¹. Ohne Kenntnis der zugrunde liegenden Texte (hier: Fernsehserien) ist die Aussage nicht verständlich.

Phraseologismen sind somit **Träger kultureller Information**, weil sie durch zahlreiche intertextuelle Bezüge an den Wissenshorizont der Mitglieder einer Kultur anknüpfen, die sich für Außenstehende nicht allein durch die sprachliche Ebene erschließen. Dies betrifft auch textsortenspezifische Wendungen (*Es war einmal...*), die bei Anhörigen dieser Kultur bestimmte Erwartungen an einen Text hervorrufen.

Die Faktoren *Translationsskopos*, *Text* und *Kultur* müssen also bei Übersetzungsentscheidungen berücksichtigt werden. Die Wahl einzelner sprachlicher Mittel folgt erst als letzter Schritt nach Absteckung dieser Rahmenbedingungen. Eine analoge Übertragung des Ausgangstextes kann in manchen Fällen eine adäquate Strategie darstellen, ist aber nicht immer die beste Lösung. Vor diesem Hintergrund ist die Übersetzung von Phraseologismen zu verstehen, von der im Folgenden die Rede ist. Wie alle anderen Wörterbücher ist auch ein phraseologisches Wörterbuch nur ein Werkzeug, das einen Impuls für kompetente translatorische Entscheidungen geben, diese aber nicht ersetzen kann.

4.2 Kontrastive Phraseologiedefinition

Nachdem die äußeren Rahmenbedingungen einer Translation abgesteckt wurden, soll nun die Rolle der Phraseologie auf der mikrostrukturellen Ebene des Ausgangs- und Zieltextes untersucht werden. Angenommen sei hierfür der Fall einer ausgangstextnahen Übersetzung. Phraseologie kann auf zweierlei Ebenen zu Schwierigkeiten führen, nämlich einmal beim Verständnis bzw. bei der **Analyse des Ausgangstextes** und zum anderen bei der **Produktion des Zieltextes**.

¹⁸ John F. Kennedy, vgl. http://de.wikipedia.org/wiki/Ich_bin_ein_Berliner (10.10.2010)

¹⁹ Louis XIV, vgl. <http://wissen.de/wde/generator/wissen/ressorts/geschichte/index.page=1178282.html>

²⁰ Friedrich II. von Preußen, vgl. http://de.wikipedia.org/wiki/Liste_gefl%C3%BCgelter_Worte/J (10.10. 2010)

²¹ vgl. http://www.enotes.com/topic/Double_entendre

Bei einer **erfahrenen Übersetzerin** mit überdurchschnittlichen Kenntnissen beider Sprachen treten Verständnisprobleme des Ausgangstextes vermutlich seltener auf, am ehesten wohl bei fachsprachlichen Phraseologismen oder bei einzelnen, weniger gebräuchlichen Idiomen. Auch was die zielsprachliche Textproduktion anbelangt, können diese beiden Bereiche am ehesten Probleme bereiten. Bei kreativeren Zieltexttypen beispielsweise können punktuell Ideen für stilistisch geeignete Idiome fehlen, auch wenn diese passiv bekannt sind. Die korrekte Verwendung alltäglicher Kollokationen hingegen ist bei Personen mit fundierten Sprachkenntnissen zu erwarten.

Anders verhält sich dies in Fällen, in denen nicht von einem vorhandenen „Sprachgefühl“ auszugehen ist, nämlich beim **Fremdspracherwerb** und bei **maschineller Übersetzung (MÜ)**. Bei der Textproduktion im Rahmen des Fremdspracherwerbs bereiten nicht nur Idiome, sondern auch Kollokationen Schwierigkeiten.

Regelbasierte MÜ-Methoden basieren auf einem Zusammenspiel von grammatischen Regelwerken und zweisprachigen Wörterbuchmodulen²². Der Übersetzungsprozess erfolgt in drei Schritten, nämlich der **Analyse** des Ausgangstextes (*Parsen*), der Übertragung in die Zielsprache, die auf unterschiedliche Arten erfolgen kann, sowie der **Generierung** des Zieltextes (vgl. Hutchins/Somers 1992).

Es liegt auf der Hand, dass Phraseologismen sowohl bei der Analyse als auch bei der Generierung zu Schwierigkeiten führen können, sofern sie nicht als Einheiten im Wörterbuch gespeichert sind. Bei der Analyse stellen zumindest Idiome mit übertrager Bedeutung eine mögliche Fehlerquelle dar; bei der Generierung sind für jede Art von Phraseologismen genaue Informationen erforderlich.

Bei Fehlen von intuitivem Sprachwissen ist demnach eine **präskriptive Sprachbe trachtung** gefragt, sozusagen ein „Bauplan“, der die Produktion korrekter zielsprachlicher Texte ermöglicht. Einen solchen Bauplan bildet das grammatische Regelwerk einer Sprache, das darüber Auskunft gibt, wie lexikalische Elemente morphologisch abgewandelt und aneinander gereiht werden müssen, um einen kohäsiven und kohärenten Text zu bilden.

Zur zielsprachlichen Textproduktion – sei es durch Lernende einer Fremdsprache oder im Rahmen maschineller Übersetzung – werden somit zunächst zwei Dinge benötigt, nämlich das **grammatische Regelwerk** und das **lexikalische Inventar** dieser

²² Im Gegensatz dazu stehen datenbasierte Systeme wie die statistische MÜ, die keinen theoretisch vorgefinierten Regelsatz enthalten, sondern anhand zweisprachiger Korpora die statistisch wahrscheinlichste Übersetzung herleiten (vgl. Carstensen et al. 2010:644ff.).

Sprache.²³ Wie sich unschwer erahnen lässt, genügt dies jedoch noch nicht, um einen kohäsiven und stilistisch einwandfreien Text produzieren zu können. Phraseologische Wortverbindungen stellen ein wesentliches Hindernis für ein derart regelhaftes Vorgehen dar – sei es, weil sich die Bedeutung eines Phraseologismus aufgrund seiner Idiomatizität nicht aus den Einzelkomponenten erschließt, oder sei es durch mangelnde Analogie zur Vergleichssprache bei der Bildung eines Ausdrucks wie *eine Entscheidung treffen*²⁴. Zudem beruhen viele Idiome auf kulturspezifischen Traditionen und Denkweisen.

Der Einsatz phraseologischer Wortverbindungen erfordert also über die Grammatikregeln hinaus zusätzliches Vorwissen. Demzufolge ist keine wörtliche Übersetzung möglich, es sei denn, es besteht zufällig eine Analogie zwischen den Vergleichssprachen. Vor diesem Hintergrund kann *Phraseologie* wie folgt definiert werden:

Der Begriff *Phraseologie* umfasst alle festen, polylexikalalen Ausdrücke, die sich nicht allein durch Anwendung des grammatischen Regelwerks und des lexikalischen Inventars einer Sprache aktiv bilden lassen, sondern zusätzliches Vorwissen erfordern.

Für den Fremdspracherwerb bedeutet diese Notwendigkeit von Vorwissen, dass Phraseologismen genau wie Lexeme in der ein- oder anderen Form auswendig gelernt werden müssen, um aktiv eingesetzt werden zu können.

Für die maschinelle Übersetzung bedeutet dies, dass Phraseologismen in einem Lexikon zur Verfügung stehen müssen. Es lässt sich bereits erahnen, dass Angaben zu einem solchen Eintrag sehr umfangreich sein müssen und sich vielschichtiger gestalten als bei einem einfachen Lexem.

In beiden Fällen wird ein phraseologisches Lexikon benötigt, das anwendungsbezogene Informationen auf mehreren Ebenen liefert. Mit Aufbau und Möglichkeiten eines solchen Nachschlagewerkes befasst sich Kapitel 5.

Wie die Definition des Phraseologismus als „eine durch die syntaktischen und semantischen Regularitäten der Verknüpfung nicht voll erklärbare Einheit“ (Burger/Buhober/Szialm 1982:1) zielt auch obige Definition auf die Nicht-Regelhaftigkeit phraseologischer Einheiten ab. Jedoch steht hier nicht die rein deskriptive Sprachbetrachtung im

²³ Am Rande sei hier bemerkt, dass Menschen im Gegensatz zu Maschinen natürlich nicht nur über Grammatikregeln, sondern vor allem über Kontextbeispiele und Analogiebildung lernen. Dieser Aspekt sollte in die Konzeption eines phraseologischen Wörterbuchs einfließen.

²⁴ Auf Portugiesisch „nimmt“ man eine Entscheidung; auf Ungarisch „bringt“ man sie.

Vordergrund, sondern vielmehr die Textproduktion, um dem präskriptiven Anspruch im mehrsprachigen Arbeiten gerecht zu werden.

Diese Auffassung von Phraseologie zieht jedoch gewisse Konsequenzen nach sich, die in den folgenden beiden Absätzen besprochen werden.

4.2.1 Mangelnde Analogie

Zum einen ist anzumerken, dass auf diese Weise die Bewertung eines Ausdrucks hinsichtlich seiner Phraseologizität nicht aus jedem Blickwinkel gleich ausfällt. So wäre der deutsche Ausdruck *die Nase putzen* aus portugiesischer Sicht nicht phraseologisch, da er regelmäßig und völlig analog zum Portugiesischen *limpar o nariz* gebildet werden kann. Das ungarische *kifújja az orrot*, wörtlich „sich die Nase ausblasen“, kann mangels Analogie nicht intuitiv richtig übersetzt werden; hier ist eine Zusatzinformation notwendig. Aus portugiesischer Sicht wäre der Ausdruck *sich die Nase putzen* also nicht unbedingt als phraseologisch zu werten, aus ungarischer Sicht aber sehr wohl.

Ob ein Ausdruck als phraseologisch gewertet werden kann oder nicht, hängt somit immer von den Gegebenheiten der jeweiligen Vergleichssprache ab. Dementsprechend definieren Hutchins/Somers (1992:105) in ihrem Werk zu maschineller Übersetzung:

From the point of view of translation, ‘idiom’ can be defined functionally. A phrase that can be translated compositionally, even if clearly idiomatic in meaning, need not be treated as such (e.g. to ask for the moon, French demander la lune). On the other hand, any phrase for which the translation is not an obvious combination of the translation of its components must be treated in a special manner (as an ‘idiom’).

Diese Definition ergibt sich als logische Konsequenz, sofern nur eine Vergleichssprache herangezogen wird. Sie erweist sich jedoch als zu eng, wenn mehrere mögliche Sprachkombinationen berücksichtigt werden sollen.

4.2.2 Schnittstelle zwischen Grammatik und Lexik

Obige Definition erklärt die Phraseologie zur Schnittstelle zwischen Grammatik und Lexik, eine Auffassung, die auch in der Computerlinguistik vertreten wird. So stellen Calzolari et al. (2002) fest: “MWEs [multiword expressions, HK] can be regarded as lying at the interface between grammar and lexicon”.

Daraus folgt, dass der Umfang der Phraseologie auch davon abhängt, wo genau die Grenzlinien zwischen Grammatik, Phraseologie und Lexik gezogen werden. Eine Grammatik könnte z.B. Regeln darüber enthalten, welche Zeitangaben mit welchen Präpositionen bzw. Suffixen zu bilden sind (*in einem Jahr* vs. *an einem Tag*) und diese

Regeln bei Bedarf um Ausnahmen ergänzen. Alternativ können solche Wortkombinationen auch der Phraseologie zugeordnet werden. Die Phraseologizität hängt also in einigen Fällen davon ab, wie engmaschig Grammatikregeln formuliert werden.

In anderen Fällen stellt sich die Frage, ob eine Wortkombination zur Phraseologie oder zur Lexik gerechnet werden soll, wie etwa Verb-Präposition- sowie Nomen-Präposition-Verbindungen (*bitten um, hinweisen auf, Lust auf, Verlangen nach, Bitte um, Aufforderung zu*). Diese könnten als feste, beim Fremdspracherwerb auswendig zu lernende Wortverbindungen zur Phraseologie gerechnet werden, oder aber in einem gewöhnlichen zweisprachigen Wörterbuch, also im Bereich der Lexik, erfasst werden.

Was dafür spricht, Verb-Präposition- und Nomen-Präposition-Verbindungen zur Lexik zu zählen, ist die Tatsache, dass ohne Angabe der entsprechenden Präposition das jeweilige Lexem nicht oder nur sehr eingeschränkt zum Einsatz kommen kann. Die Präposition ist also, ähnlich wie ein Artikel im Deutschen, eine zur korrekten Verwendung des Lexems notwendige Zusatzinformation.

Letztlich spielt es jedoch keine Rolle, wo genau die Grenzen zwischen Phraseologie, Lexik und Grammatik gezogen werden – letztlich ist dies eine reine Frage der Definition. Entscheidend ist die **lückenlose Erfassung** sämtlicher „problematischer“, sprich, nicht regelhafter Ausdrücke. Dafür ist es notwendig, dass ein Konsens besteht über die Zuordnung der verschiedenen sprachlichen Phänomene zu den drei Bereichen. Nur so kann das System einer Sprache vollständig und lückenlos erfasst werden, was eine wichtige Grundlage für maschinelle Übersetzungen darstellt.

4.3 Äquivalenzbeziehungen

4.3.1 Dimensionen der Äquivalenz

Bei der Übersetzung steht die kommunikative Äquivalenz des Gesamttextes im Vordergrund. Es kann daher auch für die Übersetzung von Phraseologismen keine Patentlösung geben.

Im Folgenden soll jedoch grob umrissen werden, auf welche Äquivalenzbeziehungen man bei der Gegenüberstellung zweier verschiedensprachiger Ausdrücke stoßen kann, und welche unterschiedlichen Übersetzungsverfahren sich daraus ergeben.

Grob kann zwischen drei Fällen unterschieden werden, nämlich dem Vorhandensein eines genauen Äquivalents, eines ähnlichen Ausdrucks, sowie dem gänzlichen Fehlen eines Äquivalents. Bei Koller (2007:605f.) und Wotjak (1982:115) finden sich die Bezeichnungen **totale Äquivalenz**, **partielle Äquivalenz** und **Nulläquivalenz**.

Koller (2007:605) beschreibt vier Klassen, mit denen er bisherige Klassifikationsvorschläge der Forschung zusammenfasst. Außer den drei genannten führt er auch die Substitutions-Äquivalenz an, die er wie folgt definiert:

„1:Substitutions-Entsprechung (*Ersetzung durch ein anderes Phrasem, Substitutions-Äquivalenz*)

Kriterien: semantische Äquivalenz, unterschiedliche lexikalische Besetzung, keine oder nur geringfügige konnotative Unterschiede. Beispiele: engl. hot potatoe – dt. heißes Eisen, fr. passer un savon à q. – dt. jm den Kopf waschen [...]“

Bei näherem Hinsehen jedoch wäre die Substitutionsäquivalenz eine Teilmenge der partiellen Äquivalenz, da Kollers Definition der partiellen Äquivalenz auch Unterschiede in der lexikalischen Besetzung beinhaltet:

„1:Teil-Entsprechung (partielle Äquivalenz)

Kriterien: semantische Äquivalenz, geringfügige Unterschiede in der lexikalischen Besetzung und/oder der syntaktischen Struktur und/oder konnotative Unterschiede.

Beispiele: [...] dt. im Laufe der Zeit – fr. avec le temps, engl. by [sic! HK] a pig in a poke – dt. die Katze im Sack kaufen [...]“ (Koller 2007:605)

Zudem besteht kein Grund, eine diesbezügliche Abweichung anders zu behandeln als andere, etwa die Stilschicht oder Syntax betreffende Abweichungen, da all diese Faktoren mitunter konnotative Unterschiede hervorrufen können. Daher soll hier zunächst nur von drei groben Äquivalenzklassen ausgegangen werden.

Aus diesen Definitionen geht hervor, dass Äquivalenz auf **mehreren Ebenen** vorhanden sein oder fehlen kann. Die bloße Übereinstimmung der denotativen Bedeutung erzeugt noch keine totale Äquivalenz, vielmehr müssen weitere Dimensionen wie etwa die syntaktische Einbettung oder die konnotative Bedeutung, die sich aus Faktoren wie Stilschicht, Sozio- und Regiolekten, Wertung etc. ergibt, berücksichtigt werden. Von totaler Äquivalenz kann nur gesprochen werden, wenn auf jeder dieser Ebenen Übereinstimmung vorliegt, was eher der Ausnahmefall sein dürfte. Koller (72004:215) postuliert daher, dass „die Verwendung des Äquivalenzbegriffs [...] die Angabe von Bezugsrahmen“ voraussetze. Einige dieser für die Phraseologie relevanten Äquivalenzebenen sollen ausführlicher dargestellt werden.

Ein wichtiger Aspekt bei Idiomen ist das **verwendete sprachliche Bild**, das bei bedeutungsgleichen Phraseologismen zweier Sprachen abweichend sein kann. So könnte die portugiesische saloppe Redewendung *gastar saliva* – wörtlich „Speichel verschwenden“ – auf Deutsch mit *gegen eine Wand reden* oder *sich den Mund fusselig reden* wiedergegeben werden. Gemeint ist in beiden Sprachen, dass man vergebens viele Worte aufwendet, um jemanden zu überzeugen. Jedoch bedienen sich beide

Sprachen eines anderen Bildes, um diese Situation zu beschreiben. Ein weiteres Beispiel ist engl. *to buy a pig in a poke* – dt. *die Katze im Sack kaufen*.

Eine Übereinstimmung der wörtlichen Bedeutung findet sich etwa bei der Redewendung *in jds. Fußstapfen treten*, die durch ung. *vki nyomdokain lépdel* auf allen Ebenen identisch wiedergegeben wird, ebenso *etwas für bare Münze nehmen* – *készpénznek vesz vmit*. Häufig ergibt sich dieser Fall bei gemeinsamen Quellen wie Werken von Goethe oder Shakespeare oder der Bibel.

Eine Übereinstimmung des Bildes kann bei ästhetisch-expressiven Textsorten eine Rolle spielen, wenn sich etwa eine bestimmte Metaphorik durch den gesamten Text zieht, oder aber bei Modifikationen, die bewusst mit beiden Lesarten des Phraseologismus spielen, z.B. bei journalistischen Texten oder in der Werbung.

Eine weitere wichtige Äquivalenzebene betrifft die **syntaktische Klasse**. Zielsprachliche Äquivalente können eine andere syntaktische Klasse oder Kontextanbindung aufweisen als der ausgangssprachliche Phraseologismus, beispielsweise wenn das deutsche *Jds. letztes Stündlein hat geschlagen* mit *to meet one's maker* ins Englische übersetzt wird. Beide Phraseologismen bezeichnen den Sachverhalt des Sterbens, jedoch handelt es sich beim englischen um eine Verbalphrase, bei der ein Subjekt zu ergänzen ist, während die deutsche Variante einen Satz mit Leerstelle (Genitivattribut) darstellt.

Auf ähnliche Weise wäre ein Kontext denkbar, in dem das englische *in a nutshell* mit *etwas kurz fassen* übersetzt wird, obwohl es sich beim englischen Phraseologismus um eine Adverbial-, beim deutschen hingegen um eine Verbalphrase handelt. Isoliert betrachtet sind die beiden Elemente nicht zwangsläufig äquivalent – so wäre es unangebracht, den Buchtitel *The Universe in a Nutshell* von Stephen Hawking²⁵ mit einer Verbalphrase zu übersetzen – sondern erst, wenn *in a nutshell* um ein Verb wie *to put sth. (in a nutshell)* ergänzt wird.

Für die Übersetzungspraxis besonders bedeutend ist die Ebene der **konnotativen Äquivalenz**, also unter anderem die Frage, ob zwei Phraseologismen dieselbe Stilschicht aufweisen. Dieser Aspekt ist umso wichtiger, da Phraseologie als Textgestaltungsmittel ein wichtiger Träger stilistischer Wirkung ist. Ebenso wichtig für die Praxis ist die **textnormative Äquivalenz**, die auf Anwendbarkeit innerhalb derselben Textsorte abzielt. Die Liste der verschiedenen Äquivalenzebenen ließe sich weiter verlängern. Festzuhalten ist jedenfalls, dass wenn zwei Vergleichsphraseologismen auf einer Ebe-

²⁵ Hawking, Stephen (2001): *The Universe in a Nutshell*. New York/ Toronto/ London/ Sydney/ Auckland: Bantam Books. Der Titel wurde mit der semantischen Entlehnung *Das Universum in der Nusschale* ins Deutsche übersetzt (s. 4.3.3).

ne voneinander abweichen, nicht mehr von totaler Äquivalenz die Rede sein kann. Die Untersuchung einiger praktischer Beispiele deutet jedoch darauf hin, dass eine derart kritische Bewertung der Äquivalenz nicht in allen Fällen zielführend ist.

Häufig finden sich leichte Abweichungen, die auf strukturellen Unterschieden basieren und mit keiner oder nur schwacher Idiomatizität einhergehen. In solchen Fällen ist ein Spiel mit der wörtlichen Bedeutung aufgrund des hohen Lexikalisierungsgrades (bzw. Phraseologisierungsgrades) eher unwahrscheinlich. Würde ein Lexikon solche Formulierungen als nur partiell äquivalent einstufen, wäre das unter Umständen irreführend, da diese Wendungen wohl in den meisten Kontexten völlig äquivalent einsetzbar sind. Tabelle 2 zeigt eine ungarisch-deutsche Gegenüberstellung solcher Beispiele.

| Ungarisch | Wörtliche Übersetzung | Deutsch |
|------------------------|---|--|
| vmiről nem is beszélve | „gar nicht erst zu reden von“ | ganz zu schweigen von |
| esik aus eső | „Es fällt der Regen“ | es regnet |
| ez nem érvényes | „das ist nicht gültig“ | Das gilt nicht! (im Spiel) |
| Régen láttam már! | „Ich habe ihn schon vor langem gesehen“ (= „Es ist schon lange her, dass ich ihn gesehen habe“) | Ich habe ihn schon lange <i>nicht mehr</i> gesehen |
| Van már programom. | „Ich habe schon Programm“ | Ich habe schon etwas vor. |
| Add át üdvözletemet. | „Überlege meinen Gruß“ | Grüße sie/ihn von mir! |
| sok hűhó semmiért | „viel Aufhebens um nichts“ (das Wort hűhó ist eine unikale Komponente und bedeutet nicht gleichzeitig auch Lärm iSv. Geräuschpegel.) | viel Lärm um nichts |
| eszébe jut | „es kommt in seinen/ihren Sinn“ | jdm. fällt etw. ein |
| kifújja az orrot | „die Nase ausblasen“ | sich die Nase putzen |
| szót vált | „ein Wort wechseln“ | Worte wechseln |
| Ezt nem élvezem. | „Das genieße ich nicht.“ | Das macht mir keinen Spaß. |

Tabelle 2: Strukturelle Abweichungen

Weitere Beispiele sind Funktionsverbgefüge und sonstige verbale Kollokationen, bei denen der Kollokator, d.h. das Verb anders als in seiner freien Bedeutung verstanden wird (*eine Entscheidung treffen*, *einen Weg einschlagen*, *Verdacht schöpfen*, *Brand stifteten*). Um solchen Fällen gerecht zu werden, soll eine vierte Äquivalenzklasse **pragmatische Äquivalenz** eingeführt werden, in der Phrasempaare zusammengefasst werden, die aufgrund struktureller Unterschiede nicht als total äquivalent gelten können, pragmatisch jedoch mit hoher Wahrscheinlichkeit völlig äquivalent einsetzbar sind. Eine Zuordnung zur totalen Äquivalenz wäre erstens methodisch unsauber, und

zweitens würde dann die Gefahr einer Ausweitung der totalen Äquivalenz auf weitere Fälle denotativer Äquivalenz bestehen, was aus weiter oben beschriebenen Gründen nicht ratsam wäre.

Im Folgenden seien die vier Äquivalenzklassen noch einmal zusammengefasst:

Totale Äquivalenz zwischen zwei Phraseologismen liegt vor, wenn Übereinstimmung auf allen Ebenen vorliegt. Dies ist wohl am ehesten bei Kollokationen der Fall, aber auch bei Idiomen, die aufgrund kultureller Kontakte oder einer gemeinsamen literarischen Quelle übereinzelnsprachlich verbreitet sind.

Pragmatische Äquivalenz liegt vor, wenn zwei Phraseologismen bedeutungsgleich sind, sich im selben Kontext einsetzen lassen und deren sprachliches Bild zumindest weitgehend identisch ist, die jedoch minimale strukturelle Unterschiede aufweisen.

Partiell äquivalent ist ein zielsprachlicher Phraseologismus, bei dem nur auf manchen der beschriebenen Ebenen Übereinstimmung besteht, der aber dennoch je nach Kontext eine geeignete Übersetzung darstellen kann. Unterschiede können z.B. das sprachliche Bild, die syntaktische Klasse oder die Stilschicht betreffen.

Bei der **Null-Äquivalenz** weist die Zielsprache überhaupt kein Äquivalent auf. Dieser Fall kann insbesondere bei stark kulturell gefärbten Idiomen auftreten sowie bei fachsprachlichen Phraseologismen, die sich auf einen in der Zielkultur noch unbekannten technischen Sachverhalt beziehen. Mögliche Übersetzungsverfahren sollen im nächsten Kapitel kurz besprochen werden.

4.3.2 Übersetzungsverfahren

Im Folgenden soll ein kurzer Überblick über mögliche Übersetzungsverfahren gegeben werden, die je nach Art der in der Zielsprache vorhandenen Äquivalente angewendet werden können. Dabei ist allerdings nicht von „Übersetzen“ des Phraseologismus im eigentlichen Sinne die Rede. Da Phraseologismen als kognitive Einheiten angesehen werden müssen, geht es vielmehr darum, aus einem bestehenden zielsprachigen Inventar ein passendes Äquivalent auszuwählen. Übersetzt wird ein Phraseologismus nur in spezifischen Fällen.

Wenn sich ein **totales Äquivalent** findet, kann dieses übernommen werden, sofern die Art des Übersetzungsauftrags keine Abwandlung des Textes auf einer höheren Ebene erforderlich machen.

Bei der **Null-Äquivalenz** weist die Zielsprache überhaupt kein Äquivalent auf. In diesem Fall wäre die **Paraphrase**, bei der die denotative Bedeutung durch freie Wortverbindung wiedergegeben wird, eine mögliche Übersetzungsstrategie. Bei einer

ersten Übersetzung kann dies eine passende Lösung sein; ein wiederkehrender fachsprachlicher Sachverhalt kann allerdings eine eigene Benennung notwendig machen.

Unter Umständen kann ein Phraseologismus wörtlich in die Zielsprache übertragen werden. Man spricht hier von **semantischer Entlehnung** (vgl. Schreiber 2003:152). Auf diese Weise kann es zur Neuprägung von Phraseologismen in der Zielsprache kommen; man denke dabei an Zitate aus der Bibel (*zur Salzsäule erstarren*) oder von Shakespeare (*viel Lärm um nichts*), die mit der Verbreitung dieser Werke in andere Sprachen eingingen. Auch bei idiomatischen Ausdrücken ist dies möglich, falls das Bild in der Zielsprache dennoch gut verständlich ist und - ohne Verfälschung des stilistischen Niveaus – die gleiche Wirkung erzielt wird. So wurde etwa Stephen Hawkings' Buch *The Universe in a Nutshell* mit *Das Universum in der Nusschale* übersetzt, obwohl die englische Wendung *in a nutshell* (in Kürze, kurz und bündig) zuvor im Deutschen nicht gebräuchlich war. Die **wörtliche Entlehnung** ist bei Mehrworteinheiten seltener anzutreffen, eher noch bei Nominal- als bei Verbalphrasen, z.B. *Joint venture*, *Know-how*, *Carsharing*. Wenn es sich um einen stark kulturell gefärbten, im Gesamtkontext weniger wichtigen Phraseologismus handelt, kommt auch eine **Auslassung** in Frage.

Wenn sich nur **partielle Äquivalente** finden, können nicht alle Ebenen zugleich übertragen werden, daher muss unter Abwägung des translatorischen Gesamtkontextes entschieden werden, welche Ebene für die Übersetzung am wichtigsten ist. So kann die Auswahl eines zielsprachlichen Phraseologismus mit abweichender syntaktischer Klasse zugunsten funktionaler Faktoren als Übersetzung geeignet sein. Etwa könnte die deutsche Verbalphrase *sich kurz fassen* bei entsprechender syntaktischer Einbettung gut mit der englischen Ergänzung *in a nutshell* wiedergegeben werden. Oder aber eine Wendung, die in der Zielsprache nicht typisch für die entsprechende Textsorte ist, wird aus dokumentarischen Gründen trotzdem beibehalten. So enthalten brasilianische standesamtliche Dokumente wie Heirats- oder Geburtsurkunde üblicherweise die floskelhafte Formulierung *O referido é verdade e dou fé* vor der Unterschrift, zumeist wiedergegeben mit *Die Wahrheit des Besagten bezeuge ich hiermit*. Deutschsprachige vergleichbare Dokumente werden nicht von Standesbeamten in der ersten Person verfasst.

Bei Idiomen ist es in manchen Fällen angezeigt, den Inhalt mit den in der Zielsprache zur Verfügung stehenden sprachlichen Mitteln möglichst getreu wiederzugeben – z.B. über eine **Paraphrase** mit freien Wortverbindungen oder einen zielsprachlichen Phraseologismus, der sich einer anderen Bildhaftigkeit bedient – und dafür einen Verlust des sprachlichen Bildes in Kauf zu nehmen. Dies empfiehlt sich besonders bei informativen Texttypen, bei denen die formale Ebene nicht unnötig Aufmerksamkeit auf sich ziehen sollte.

Bei anderen Idiomen kann es angebracht sein, die Denotation unter Beibehaltung der Ausdrucksebene leicht abzuwandeln. Dieses Verfahren wird als **Mutation** bezeichnet und kommt hauptsächlich bei expressiven, ästhetischen Texttypen zum Einsatz (vgl. Schreiber 2003:152).

Die ungarische Entsprechung des Ausdrucks *keinen blassen Schimmer haben* wäre im Allgemeinen (*halvány) sejtelme sincs róla*, wörtlich *keine (blasse) Ahnung haben*. Kommt es nun in einem Text speziell auf die Analogie *Helligkeit – Wissen/ Intelligenz* an, weil diese Metaphorik in einem größeren Textzusammenhang steht, können z.B. Wendungen wie *napfényre kerül* (*ans Tageslicht kommen*) in negierter Form oder *nem látja világosan* (*etwas nicht klar sehen*) verwendet werden, um die Metaphorik aufrecht zu erhalten, auch wenn der Sinn dadurch verzerrt wird.

Diese Beispiele machen deutlich, dass auch ein begrifflich nicht identischer Phraseologismus eine geeignete Übersetzungslösung darstellen kann. Die Auswahl eines Äquivalents hängt stark vom jeweiligen Kontext ab; daher wäre es wünschenswert, wenn ein phraseologisches Nachschlagewerk möglichst viele zielsprachliche Äquivalente – versehen mit genauen Anmerkungen zum Gebrauch – zur Auswahl stellen würde. Diese Ansicht vertritt auch Veisbergs (1996:243):

“...sometimes [...] the equivalent idiom does not present the best counterpart for the translation at all, while analogous idiom and even and [sic! HK] idiom with a completely different image may be of great use [...]. I would therefore argue for an extensive approach to idiom entries – including of several possible variants with the equivalent naturally being number one.”

Ob und wie dies umsetzbar ist, wird in Kapitel 5 untersucht.

4.4 Phraseologie in Wörterbüchern

Bevor wir uns mit elektronischer Phraseographie befassen, soll ein kurzer Seitenblick auf Phraseologie in gedruckten Wörterbüchern und damit verbundene Herausforderungen geworfen werden, um diesbezügliche Probleme bei der Konzeption eines elektronischen Wörterbuchs berücksichtigen zu können. Der Grundtenor einschlägiger Forschungsliteratur ist, dass Phraseologie vielen Werken noch nicht zufriedenstellend behandelt wird.

Grundsätzlich ist zwischen mono- und bilingualen sowie allgemeinen und speziell phraseologischen Wörterbüchern zu unterscheiden. Monolinguale sind im vorliegenden Kontext gleichermaßen interessant wie bilinguale, da auch differenzierte einsprachige Daten ein wichtiges Hilfsmittel bei Übersetzungsentscheidungen darstellen.

Ebenso scheinen allgemeine Wörterbücher einer Untersuchung wert zu sein, da diese von Personen ohne spezielles Fachwissen als erstes konsultiert werden (vgl. Moon 2007:909). Auch besteht nicht in allen Fällen Bewusstsein darüber, dass das gesuchte Lexem Teil einer phraseologischen Einheit ist. Gleichzeitig geht die Integration von Phraseologismen in ein auf Einwortlexeme zugeschnittenes „lexikographisches ‚Korsett‘“, wie Stantcheva (2002:39) sich ausdrückt, zwangsläufig mit größeren Schwierigkeiten einher. Im Folgenden werden auszugsweise einige wichtige Schwachstellen verschiedener Wörterbuchtypen beschrieben:

Moon (2007:912) weist auf die Fragestellung hin, unter welchem Lemma ein Phraseologismus eingeordnet werden soll. Burger (?2003:174) beschreibt die im *Duden. Deutsches Universalwörterbuch* verwendete Lösung, zuerst nach dem ersten vorkommenden Substantiv, und falls es keines gibt, nach dem ersten Autosemantikum im Phraseologismus zu sortieren. Er beurteilt dies als „relativ unproblematisches Verfahren“ (Burger ?2003:174), das auch für Laien nachvollziehbar sei. Zweifellos am benutzerfreundlichsten wäre die **Verzeichnung unter mehreren Lemmata**, entweder in Form von Querverweisen oder als volle Duplikate, denn, wie Moon (2007:912) feststellt, „users often expect to use dictionaries without studying front matter“. Jedoch beobachtet sie diesbezüglich inkonsequent geführte Querweise sowie widersprüchliche Angaben in doppelten Einträgen (vgl. Moon 2007:912), womit Benutzern wiederum wenig gedient ist.

Burger (?2003:170ff.) und Stantcheva (2002:177ff.) bemängeln bei allgemeinen, einsprachigen deutschen Wörterbüchern, dass **Phraseologismen nicht ausreichend gekennzeichnet** werden, was zu einer unzureichenden Abgrenzung von freien Wortverbindungen als Beispiel zu einem Eintrag führt. Stantcheva (2002:177ff.) stellt zudem fest, dass die zur Markierung von Phraseologismen verwendete Terminologie uneinheitlich sei und nicht dem aktuellen Forschungsstand entspreche. Auch hinsichtlich **der Nennform** von Phraseologismen stellen Burger (?2003:178), Moon (2007:913) und Stantcheva (2002:180f.) **Uneinheitlichkeit** sowie mangelnde Abgrenzung von aktualisierten Beispielen fest.

Ettinger (1989) untersucht verschiedene mono- und bilinguale Sammlungen französischer Phraseologismen für deutschsprachige Lernende. In den untersuchten Werken beobachtet er **Unzulänglichkeiten bei Registermarkierungen**, wodurch es Französischlernenden erschwert wird, neue Phraseologismen im richtigen Kontext einzusetzen. Insbesondere enthielten die von ihm untersuchten Werke zu viele veraltete Termini, die aber nicht entsprechend gekennzeichnet seien, außerdem fehlen Angaben zur Frequenz von Phraseologismen sowie diatopische, diastratische und diaphasische Markierungen.

Generell kritisiert er, dass „nicht zwischen **aktiven und passiven Wörterbuchtypen** unterschieden“ wird (Ettinger 1989:96). Dadurch werden Nutzern nicht genügend anwendungsbezogene Informationen an die Hand gegeben. Ähnliches beobachten Lubensky/McShane (2007:925) am Beispiel eines englisch-russischen Wörterbuchs:

„such coarse-grained bunching approaches do not [...] come close to the goal of bilingual phraseological dictionaries, which is to provide all the semantic contextual, stylistic and usage information that a non-native speaker needs in order to appropriately use phrasemes in real-world situations.“

Dies sind nur wenige Beispiele der Probleme der gedruckten Phraseographie; die Liste ließe sich fortsetzen. Ein Teil der beschriebenen Probleme lässt sich auf **mangelnde Umsetzung theoretischer Erkenntnisse in die lexikographische Praxis** (vgl. Burger 2003:170f.) zurückführen, wie etwa formale Uneinheitlichkeit und mangelnde Registermarkierung. Weitere Probleme, insbesondere die von Ettinger (1989) beklagte mangelnde Aktualität und Berücksichtigung der Frequenz, lassen sich durch **korpusgestützte Datengewinnung** aus der Welt schaffen, wodurch z.B. Frequenzanalysen und Erfassung verschiedener Varianten möglich werden. In der Tat geht der Trend in diese Richtung, wie Lubensky/McShane (2007:919) – zumindest im englischsprachigen Bereich – beobachten:

„As work continues on defining the optimal content and structure of entries in various kinds of lexicons, and as demands on the scope and quality of language-learning resources intensify, the availability of monolingual and parallel corpora put a new and valuable resource in the hands of lexicographers.“

Dabei ist anzumerken, dass der Erfolg der Datengewinnung aus Korpora auch von präziser und systematischer Weiterverarbeitung sowie von der Qualität der verwendeten Korpora abhängt, z.B. von der Qualität der Übersetzungen in Parallelkorpora.

Wieder andere Probleme sind durch die Platzbeschränkung und die lineare Struktur des Mediums *Buch* bedingt und würden sich durch den **Einsatz elektronischer Medien** automatisch lösen, wie Moon (2007:917) beschreibt:

„... books have naturally limits on what can be said about words and phrases. Many problems disappear with electronic dictionaries, and this includes indexing, placement, and accessibility. Since space constraints no longer apply, entries can be much fuller and provide the information that is currently missing or underreported.“

Elektronische Medien ermöglichen nicht nur die Aufnahme von Phraseologismen unter mehreren Lemmata sowie Einheitlichkeit dieser doppelten Einträge, sondern auch die Möglichkeit zu ausführlicheren gebrauchsbezogenen Angaben einschließlich Kontextbeispielen.

5 Entwurf einer mehrsprachigen elektronischen Phraseologiedatenbank

In den vorangehenden Kapiteln wurde die Vielfalt phraseologischer Klassen und mehrsprachiger Äquivalenzbeziehungen deutlich. Eine kurze Bestandsaufnahme der gedruckten Phraseographie hat gezeigt, dass ein elektronisches Nachschlagewerk dieser Komplexität viel besser gerecht werden kann, da Einträge über **verschiedene Lemmata und Suchfilter** gefunden und durch **Querverweise** miteinander verknüpft werden können. Zudem ist die Aufnahme einer **großen Auswahl zielsprachlicher Äquivalente** möglich, was von Vorteil ist angesichts der Tatsache, dass je nach Kontext auch stark abweichende Entsprechungen eine adäquate Übersetzung darstellen können (s. 4.3).

Ein weiterer Vorteil ist **Aktualität**, was sich angesichts rascher sprachlicher Entwicklungen in Fachsprache und Medien wesentlich auf die Qualität auswirkt. Am meisten kommt dieser Aspekt zum Tragen, wenn in Form von **Online-Zusammenarbeit** sehr viele Menschen an dem Nachschlagewerk mitarbeiten können. Diese Möglichkeit soll in 5.5.2 besprochen werden. Weiterhin erlaubt elektronische Phraseographie die Integration von **multimedialem Bild- und Tonmaterial**.

Dieses Kapitel beschreibt den konzeptuellen Entwurf einer elektronischen Phraseologie-Datenbank.²⁶ Zunächst werden ihre Grundeigenschaften (5.1) und Ordnungsprinzipien (5.2) festgelegt. Abschnitt 5.3 bietet einen Entwurf zur grundlegenden Strukturierung der Daten. Dieser wird in 5.4 durch die Analyse einsprachiger sowie übersetzungsbezogener Datenkategorien verfeinert. Der nächste Schritt ist die Planung der Contenterstellung und -Bearbeitung in 5.7, bei der u.a. die Usability ins Blickfeld rückt. Um die konkrete Implementation mit einem bestimmten Datenbanksystem geht es dabei nicht, sondern ausschließlich um einen konzeptuellen Entwurf, der als Grundstein für ein neues phraseographisches Projekt dienen kann.

²⁶ Da dem zu besprechenden Wörterbuch eine Datenbankstruktur zugrunde liegen muss, sind die Termini *Wörterbuch*, *Datenbank* und *Nachschlagewerk* im Folgenden synonym zu verstehen.

5.1 Grundeigenschaften

Konzipiert wird ein **mehrsprachiges phraseologisches Nachschlagewerk**, das dem Anspruch eines **produktiven Wörterbuchs** genügen soll, d.h., es soll die zur aktiven Verwendung eines zuvor unbekannten Phraseologismus nötigen Informationen bereitstellen, und zwar in unterschiedliche Übersetzungsrichtungen; es soll also **multidirektional** sein. Der produktive Anspruch schließt die Aufnahme veralteter oder weniger gebräuchlicher Phraseologismen keineswegs aus, jedoch müssen diese entsprechend gekennzeichnet werden. Gleichzeitig sollen, wie in 4.3.2 gefordert, zu einem Phraseologismus bei Bedarf **mehrere zielsprachliche Äquivalente** aufgenommen werden, darunter auch **partielle Äquivalente**, die je nach Kontext ebenfalls eine adäquate Übersetzungslösung darstellen können. Hierbei wird Übersetzungsbeispielen und Kommentaren zur Äquivalenz Raum gegeben, da diese die Qualität des Nachschlagewerks erheblich steigern.

Die Datenbank beschränkt sich weder auf einen bestimmten Typ von Phraseologismen noch auf ein **bestimmtes Fachgebiet**. Jedoch sollte die Möglichkeit bestehen, Phraseologismen einem Fachgebiet zuzuordnen und fachsprachlich relevante Information, etwa eine Graphik, sachgerecht einzubinden.

Bezüglich der Zielgruppe richtet sich das Wörterbuch vornehmlich **menschliche Benutzer**, also Sprachlernende und Übersetzerinnen. Dabei wird jedoch die Frage der **Maschinenlesbarkeit** ebenfalls berücksichtigt, das heißt, das Wörterbuch soll nach Möglichkeit so strukturiert sein, dass sein Inhalt zu einem späteren Zeitpunkt maschinell weiterverarbeitet und in einen maschinellen Übersetzungsprozess integriert werden kann. Hierfür ist es u.a. wichtig, die von Arntz et al. (2004:238) beschriebenen Prinzipien der **Granularität** und der **Elementarität** zu beachten. Dies bedeutet, dass Angaben zu einem Phraseologismus nicht einfach als Fließtext dargestellt werden, sondern dass jeder Informationstyp in einem eigenen Feld, einer gesonderten **Datenkategorie** gespeichert wird. Granularität bedeutet hierbei, dass „die Datenkategorien möglichst genau und fein definiert werden“ (Arntz et al. 2004:238), während Elementarität die Tatsache beschreibt, dass „jede Datenkategorie nur das zu ihr passende Datenelement enthält“ (Arntz et al. 2004:238). So soll etwa die Quelle einer Definition nicht im selben Feld wie die Definition selbst gespeichert werden.

Was die **Herkunft des Datenbestandes** betrifft, gibt es mehrere Möglichkeiten: Daten können aus Korpora gewonnen, durch ein kleineres Team von Expertinnen erstellt oder in Form von Online-Kollaboration von einem breiteren Publikum eingegeben werden (vgl. hierzu 5.5.2). Welche Eingabeform auch immer gewählt wird – es wird angestrebt, die Möglichkeit der **Online-Kollaboration** jedenfalls mit einzuschließen, sei es zur Ergänzung bestehender Daten oder zur kompletten Inhaltserstellung. Denn durch die Beteiligung mehrerer Personen kann ein realistischeres Bild des tatsächli-

chen Sprachgebrauchs und zudem höhere Aktualität gewährleistet werden. Eine solche Funktion, ob sie nun öffentlich oder nur für bestimmte Personen zugänglich ist, erfordert die Schaffung geeigneter Strukturen und eine verständliche Benutzeroberfläche (s. 5.5.5).

Anders als in der terminologischen Praxis ist nicht vorgesehen, dass ein Eintrag durch eine **Quelle** belegt werden muss; **Dateneingabe auf introspektiver Basis** soll möglich sein. Damit wird der Tatsache Rechnung getragen, dass „selbst die größten Korpora nur Sprachausschnitte darstellen und daher weite Teile der nicht schriftlich dokumentierten Sprachverwendung auch nicht per Korpusrecherche zugänglich sind“ (Dräger/ Juska-Bacher 2010:167). Dies kann z.B. umgangssprachliche oder literarische Wendungen betreffen. Zur Bestätigung des intuitiven Sprachwissens können Einträge über User-Befragungen auf ihre tatsächliche Gebräuchlichkeit und stilistische Wirkung hin geprüft werden

Aus dem produktiven Anspruch des Wörterbuchs ergibt sich die **Notwendigkeit umfassender Informationen zum Gebrauch** eines Phraseologismus wie etwa Stilschicht, areale Verbreitung, Bezogenheit auf bestimmte Textsorten und nicht zuletzt Kontextbeispiele, denn „ohne Beispiele und Belege ist es kaum möglich, dem Benutzer eine Vorstellung von realen Verwendungsbesonderheiten des Phraseologismus zu vermitteln“ (Dobrovolskij 1999:117).

Dem ist hinzuzufügen, dass je nach Verwendungszweck der Datenbank – menschliche Nutzer oder maschinelle Weiterverarbeitung – andere Daten relevant sind. Während für maschinelle Weiterverarbeitung systematisch kodierte Informationen über Stilschicht, syntaktische Anbindung etc. nötig ist, lernen Menschen vor allem aus anschaulichen Beispielen. Idealerweise sollte die Datenbank beiden Arten von Daten Raum geben.

Angesichts dieser Fülle von Daten stellt sich die Frage nach einer sinnvollen Struktur und Übersichtlichkeit in der Darstellung. In Bezug auf gedruckte Wörterbücher äußert Stantcheva (2002:38):

„All diese Datentypen möglichst vollständig zu erfassen sowie ökonomisch und dennoch benutzerfreundlich im Wörterverzeichnis zu präsentieren, ist eine schwierig zu lösende lexikografische Aufgabe.“

Diese Aussage ist für elektronische Nachschlagewerke nicht minder gültig, jedoch ist die Aufgabe hier durch flexiblere technische Möglichkeiten besser lösbar. Dazu ist einerseits eine logische Struktur erforderlich, in die alle beschriebenen Daten integriert werden können. Damit beschäftigen sich die Abschnitte 5.2 bis 5.4. Andererseits ist eine klar strukturierte Benutzeroberfläche notwendig, um trotz umfangreicher Daten Übersichtlichkeit zu wahren, etwa über Benutzerprofile, die die angezeigten Daten auf eine interessensspezifische Auswahl beschränken (s. 5.5.4, 5.5.5).

5.2 Ordnungs- und Suchprinzipen

Im Zusammenhang mit der Festigkeit von Phraseologismen wurde festgestellt, dass Phraseologismen kognitiv als Einheit gespeichert und abgerufen werden. Auch im Rahmen computerlinguistischer Bestrebungen nach Standards für Mehrworteinheiten in Lexika werden Phraseologismen beschrieben als „a sequence of words that acts as a single unit at some level of linguistic analysis“ (Calzolari et al. 2002)²⁷. Dies legt nahe, **Phraseologismen als Ganzes in einem eigenen Eintrag** zu speichern, der dann über die Eingabe bestimmter Suchbegriffe gefunden werden kann. Auch Budin (1990:68) stellt fest: „...there is no reason to treat the phraseology differently from the simple term“.

Ansätze, einzelne Komponenten getrennt abzuspeichern und einander über kombinatorische Vorschriften zuzuordnen, sind nur bei Kollokationen praktikabel, nicht aber bei undurchsichtigen Idiomen, deren Komponenten in freier Wortverbindung mitunter eine völlig andere Bedeutung tragen. Einen kombinatorischen Ansatz stellen Mel'čuk's *lexikalische Funktionen* dar, die die – paradigmatischen oder syntagmatischen – Beziehung zweier lexikalischer Einheiten formal beschreiben (vgl. Mel'čuk 2007:119ff.). Allerdings beschränkt Mel'čuk diese Methode auf Kollokationen; Idiome reiht er neben Lexemen unter die lexikalische Einheiten (ebd:120). Lexikalischen Funktionen stellen einen systematischen, aber hoch abstrakten Formalismus dar und machen zudem eine getrennte Behandlung von Kollokation und Idiomen erforderlich. Daher ist die Verwaltung von Phraseologismen als Ganzes in einem Wörterbuch für menschliche Nutzer übersichtlicher.

Grundsätzlich wird somit eine Datenbank für polylexikale Einträge konzipiert; diese sollte aber die Aufnahme einzelner Wörter nicht ausschließen, da das zielsprachliche Äquivalent eines polylexikalischen Ausdrucks ein unigraphes Lexem sein kann und umgekehrt.

Traditionell stellt sich beim Anlegen eines Wörterbuchs die Frage, ob der Inhalt **semasiologisch** („von den Wörtern zu den Bedeutungen“) oder **onomasiologisch** („von den Bedeutungen zu den Wörtern“) gegliedert werden soll (vgl. Arntz et al.⁵ 2004:189). Diese Frage wird im elektronischen Wörterbuch insofern relativiert, als dass die Speicherung ohnehin nicht an eine bestimmte Reihenfolge gebunden ist – technisch gesehen werden neue Einträge einfach in Eingabereihenfolge gespeichert und mit einer ID (Eintragsnummer) versehen.

²⁷ Der Aufsatz beschreibt das XMELLT-Projekt (Cross-lingual Multi-word Expression Lexicons for Language Technology) zur Erarbeitung internationaler Standards für Mehrworteinheiten in computergestützten Lexika, vgl. a. Ide (2000).

Von einer onomasiologischen Struktur könnte man dann sprechen, wenn mehrere begrifflich übereinstimmende Phraseologismen in einem Eintrag gespeichert würden. Dies scheint jedoch nicht ratsam, da zu jedem einzelnen Phraseologismus differenziertere Angaben hinsichtlich seines Gebrauchs und möglicher Konnotationen nötig sind.

Die Frage des Ordnungsprinzips reduziert sich also darauf, nach welchen **Suchkriterien** ein Phraseologismus gefunden werden kann. Sinnvoll ist es jedenfalls, wenn Phraseologismen über die Eingabe ihrer **Komponenten** – alle Autosemantika in verschiedenen Flexionsformen – auffindbar sind. Zusätzlich ist auch eine **semantische Beschriftungswortung** wünschenswert, die insbesondere unigraphe Synonyme einschließt (z.B. *sterben* bei *ins Gras beißen*). Aber auch eine weiterreichende Zuordnung verwandter Begriffe wäre denkbar, z.B. auf Basis eines semantischen Netzes oder Thesaurus, sofern dadurch die Trefferliste nicht unerwünscht groß wird.

5.3 Datenbankstruktur

Bei einer mehrsprachigen Datenbank ist es sinnvoll, zwischen zwei Arten von Datenkategorien zu unterscheiden:

Die erste Datengruppe bezieht sich zunächst nur auf den Phraseologismus einer Sprache. Dazu zählt etwa eine Definition in derselben Sprache, die Angabe des Fachgebiets und der Stilschicht, oder Angaben über Möglichkeiten der morphosyntaktischen Einbettung in den Text. Ein solcher **einsprachiger Datensatz** soll für Phraseologismen jeder Sprache unabhängig von der jeweiligen Vergleichssprache angelegt werden.

Die zweite Art von Daten umfasst Angaben über die Verknüpfung zweier bestimmter Phraseologismen in verschiedenen Sprachen, die hier als **Übersetzungsdaten** bezeichnet werden, und die als weiterer, übergeordneter Datensatz angelegt werden. Hierzu zählen translationsspezifische Daten, wie Äquivalenz und Übersetzungsbepiele.

Abb. 3 zeigt eine Datenbankstruktur nach diesem Aufbau.

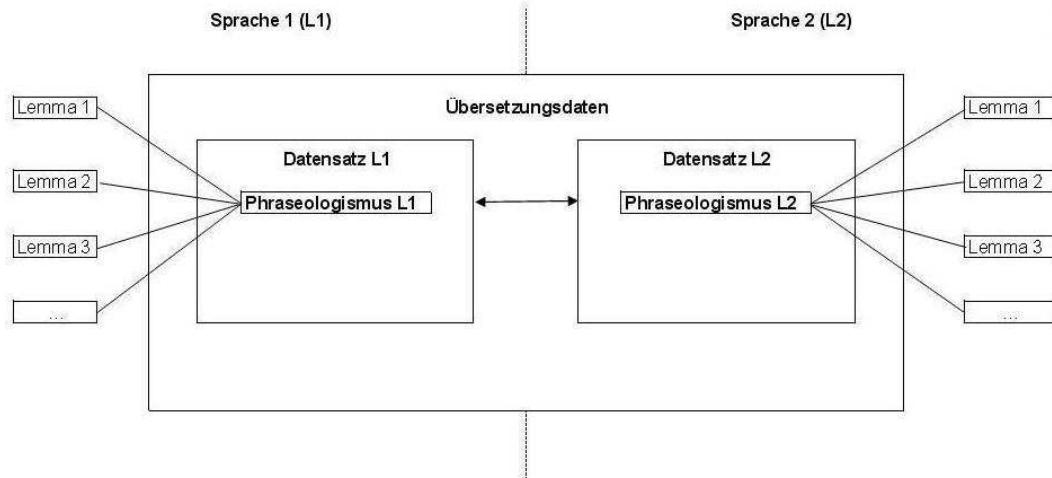


Abb. 3: Struktur eines phraseologischen Wörterbuchs

Ein einsprachiger Eintrag kann jeweils mit mehreren zielsprachlichen Phraseologismen, die mögliche Übersetzungen darstellen, verknüpft sein. Wie erwähnt ist es hilfreich, einen Phraseologismus mit möglichst vielen L2-Äquivalenten zu verknüpfen. Hierbei wird jede Verknüpfung als eigener Datensatz mit Angaben zur jeweiligen Äquivalenzbeziehung angelegt.

Für den Fall, dass für einen ausgangssprachlichen Phraseologismus kein genaues oder ungefähres zielsprachliches Äquivalent bekannt ist, wäre die Möglichkeit sinnvoll, trotzdem bestimmte Informationen in der Zielsprache zur Verfügung zu stellen – vornehmlich eine Definition und Paraphrase in der Zielsprache, die Impulse für eine Übersetzung liefern können. Diese Überlegung wird in 5.4.1.8 weiter ausgeführt.

In den folgenden Kapiteln werden die einzelnen Datenkategorien besprochen. Viele davon wurden von Arntz et al. (2004:231ff.) übernommen; die Kategorien werden hier jedoch in einsprachige und übersetzungsbezogene Daten eingeteilt und um weitere ergänzt. Die Darstellung erhebt keinen Anspruch auf Vollständigkeit. Je nach Interessenschwerpunkt können ggf. Ergänzungen vorgenommen oder beschriebene Kategorien weiter ausdifferenziert werden. Daher sollte eine Datenbank entsprechend ausbaufähig sein.

5.4 Datenkategorien

5.4.1 Einsprachige Datenkategorien

5.4.1.1 Verwaltungsbezogene Daten

Unter die von Arntz et al. (2004:231ff.) beschriebene Gruppe der verwaltungsbezogenen Daten fällt im phraseologischen Bereich das **Schlagwort**, auch Lemma oder

Stichwort genannt. Als Lemma sollten zumindest alle enthaltenen Autosemantika in ihrer Grundform, d.h. also Verben im Infinitiv und Nomen im Nominativ Singular gespeichert werden. Die Angabe weiterer Formen, z.B. Flexionsformen, alternative Schreibweisen und häufige Schreibfehler ist ebenfalls hilfreich.

Wichtig ist hierbei die Möglichkeit, Schlagworte **explizit definieren** zu können. Moderne Suchalgorithmen weisen zwar eine hohe Variationstoleranz auf, allerdings gibt es Grenzen, wie etwa graphisch und lautlich komplett abweichende Flexionsformen. Ein Beispiel hierfür ist die griechische Verbalflexion: Sie basiert auf zwei unterschiedlichen Stämmen, die bei einigen Verben komplett voneinander abweichen (z.B. τρώω [troo] „essen“, Imperfektstamm: τρω- (tro-), Perfektstamm: φα- (fa-)) Wenn ein solches Verb in einem Phraseologismus vorkommt, kann dieser nur durch Eingabe bestimmter Verbalformen in der Suche gefunden werden, es sei denn, es wurden zusätzliche Schlagworte definiert. Außerdem ist die Möglichkeit einer gezielten Definition von Schlagworten für eine semantische Beschlagwortung nötig (s. 5.2).

5.4.1.2 Kerninformationen zum Phraseologismus

Die wichtigste Angabe ist der Phraseologismus selbst in seiner **Nennform**, also in einer standardisierten Form, die nicht an einen konkreten Fall angepasst wurde. Die Notation der Nennform ist dabei, wie auch Burger (2003:178) betont, keineswegs „eine bloße Formalität“ da in ihr bereits Informationen über externe Leerstellen und morphosyntaktische Restriktionen des Phraseologismus enthalten sind.

Externe Leerstellen können durch Strukturformeln wie [etw] oder [jdm] in der Nennform enthalten sein. Dabei ist es wichtig, externe Leerstellen einheitlich zu notieren und diese klar von festen Komponenten abzugrenzen (vgl. die Funktion von *etwas* in *etwas im Silde führen* vs. [etw] *an die große Glocke hängen*). Im Folgenden werden externe Leerstellen daher durchgehend mit eckigen Klammern notiert.

Externe Leerstellen unterliegen morphosyntaktischen und semantischen Restriktionen, die bis zu einem gewissen Grad aus der Notation der Nennform hervorgehen können. Die Strukturformeln [jdn] und [jdm] geben bereits Aufschluss über den Kasus, [etw] hingegen nicht. Eine Möglichkeit wäre die Darstellung des Kasus als Index:

[etw Akk] *an die große Glocke hängen*

[etw Dat] *Ausdruck verleihen*

[etw Gen] *überdrüssig werden*

Morphosyntaktische Merkmale der externen Leerstellen, die aus der Nennform nicht hervorgehen, müssen in einer gesonderten Datenkategorie *Restriktionen* angeführt werden (s. 5.4.1.5/6).

Auch auf die semantische Besetzung der Leerstellen kann die Nennform einen Hinweis geben: Verbalphrasen werden in der Regel im Infinitiv angeführt; diese Form suggeriert jedoch, dass das zu ergänzende Subjekt menschlich sei, was nicht immer der Fall ist (vgl. *die Zügel in der Hand haben* vs. *glatt über die Bühne gehen*). Um Bildungen wie *Peter geht glatt über die Bühne zu vermeiden, sollen verbale Phraseologismen mit nicht-menschlichem Subjekt in flektierter Form dargestellt, also

[etw] schlägt hohe Wellen
[etw] geht glatt über die Bühne

Schließlich gibt die Nennform zum Teil auch Aufschluss über transformationelle Defekte des gesamten Phraseologismus, also Einschränkungen in der Abwandelbarkeit von Tempus, Numerus etc. So machen die jeweiligen Nennformen deutlich, dass *an [jdm] einen Narren gefressen haben* nur im Perfekt verwendet wird und *[etw] ist nicht das Gelbe vom Ei* immer in Negation steht. Auch feste Satzstrukturen ohne syntaktische Anbindung an den Kontext (*Stille Wasser sind tief*) erkennt man an der Satzform der Nennform und dem Fehlen von Leerstellen. Andere transformationellen Defekte gehen nicht aus der Nennform hervor und müssen in einer eigenen Datenkategorie explizit genannt werden (s. 5.4.1.6).

Die Darstellungsweise der Nennform sollte also mit den Datenkategorien *transformationelle Defekte* sowie *externe Leerstellen* und deren *Restriktionen* (s. 5.4.1.6) abgestimmt werden, um Einheitlichkeit und lückenlose Darstellung zu ermöglichen.

Unabdingbar ist auch eine **Definition** zur präzisen Eingrenzung der Bedeutung, insbesondere bei Idiomen und fachsprachlichen Wendungen. Dabei stellt sich die Frage, in welcher Sprache die Definition verfasst sein soll. Da der einsprachige Datensatz ja grundsätzlich unabhängig von möglichen Vergleichssprachen sein soll, ist klarerweise eine Definition in derselben Sprache unabdingbar. Jedoch wäre zusätzlich auch eine zielsprachliche Definition für eine Übersetzung hilfreich; dieser Aspekt wird in 5.1.4.8 näher erläutert.

Beim Eintragen zielsprachlicher Äquivalente gilt es, zwischen **phraseologischen** und **nicht-phraseologischen Einträgen** zu unterscheiden. Wie aus den Übersetzungsverfahren in 4.3.2 ersichtlich wurde, ist ein optimales zielsprachliches Äquivalent nicht zwangsläufig phraseologisch; beispielsweise könnte das ungarische *visszatér a régi kerékvágásba* („wieder ins alte Geleise zurückkehren“) mit *sich normalisieren* übersetzt werden. Die Datenbank sollte daher für unigraphe Einträge oder freie Wortverbindungen offen sein, dabei jedoch eine **entsprechende Kennzeichnung** ermöglichen. Erstens wird dadurch Aufschluss darüber gegeben, ob die Wortverbindung fest ist oder bei Bedarf abgewandelt werden kann. Zweitens wäre in der maschinellen Sprachverarbeitung die Untersuchung von Möglichkeiten interessant, mittels Verfahren der Mustererkennung auch Modifikationen von Phraseologismen in einem Text zu er-

kennen. So könnte die Artikelüberschrift *Lena kam, sang und siegte*²⁸ über den Eurovision-Song-Contest als Abwandlung von Caesars Ausspruch *Ich kam, sah und siegte* identifiziert werden. Dabei kann ein phraseologisches Lexikon als Vergleichsgrundlage dienen; bei eingetragenen freien Wortverbindungen hingegen wäre eine solche Prüfung auf Ähnlichkeit natürlich nicht zielführend.

5.4.1.3 Allgemeine benennungsbezogene Angaben

Zu einem Phraseologismus kann ein Synonym bzw. eine Paraphrase angeführt werden. Unigraphe und polylexikale Einträge können hier nebeneinander verzeichnet werden; es muss nicht zwischen Synonym und Paraphrase unterschieden werden. Wichtig ist, dass die Einträge der gleichen syntaktischen Klasse angehören wie der Phraseologismus selbst. Falls das Synonym als eigener Eintrag in der Datenbank steht, empfiehlt sich ein Querverweis.

Eine weitere benennungsbezogene Datenkategorie bilden die **Varianten**. Hier gilt es zu entscheiden, ob mehrere Varianten in einem Haupteintrag zusammengefasst, oder aber in jeweils eigenen Einträgen gespeichert werden sollen. Bei Varianten mit nur minimalen Abweichungen – z.B. gleicher lexikalischer Besetzung und leichten morphologischen Unterschieden wie in [jdm] *auf die Nerven gehen*/[jdm] *auf den Nerv gehen* – würde eine getrennte Speicherung unnötigen Aufwand bedeuten, da hier sämtliche gebrauchsbezogene Angaben vermutlich ohnehin identisch ausfallen. Ein Beispiel für eine solche Gruppe von Varianten wären die des ungarischen Phraseologismus *visszatér a régi kerékvágásba* (wörtlich „ins alte Geleise zurückkehren“, vgl. Anhang A):

visszatér a régi kerékvágásba
visszatér a rendes kerékvágásba
beáll a régi kerékvágásba
beáll a rendes kerékvágásba
visszazökken a régi kerékvágásba
visszazökken a rendes kerékvágásba

Anders liegt der Fall bei Varianten mit stärkerer Abweichung in der lexikalischen Besetzung, oder aber solche mit unterschiedlicher syntaktischer Funktion. Manche Phraseologismen treten in einer längeren und einer kürzeren Form auf, die sich in ihrer syntaktischen Funktion unterscheiden:

28

http://www.buecherhallen.de/aw/hoebu/community_hoebu/blog/~ifm/%5BMediennews%5D_Lena_kam_sang_und_siegte!/ (Stand: 2.10.2010)

| Kurzform | übliche Realisierung | Beispiel für andere Realisierung |
|--------------------------|--------------------------------------|--|
| durch die Blume (AdvP) | [etw] durch die Blume sagen (VP) | <i>Er ließ durch die Blume durchklingen, dass er dieser Frage kritisch gegenüberstand.</i> |
| [jds] sieben Sachen (NP) | [seine] sieben Sachen packen (VP) | <i>Er kam mit seinen sieben Sachen hier an.</i> |
| der Zahn der Zeit (NP) | der Zahn der Zeit nagt an [etw] (VP) | <i>Der Zahn der Zeit hatte in dem alten Gemäuer seine Spuren hinterlassen.</i> |

Tabelle 3: Lang- und Kurzform von Phraseologismen

Sowohl die lange als auch die kurze Form sind als phraseologisch zu werten. Die Beispiele zeigen, dass die Kurzform auch anders eingesetzt werden kann als in Verbindung mit dem langen Phraseologismus und somit ein eigenständiger Phraseologismus ist. Auch oben genanntes Beispiel ist häufig in der Kurzform *vissza a régi kerékvágásba* (wörtlich „zurück ins alte Geleise“) anzutreffen.

Aufgrund der abweichenden syntaktischen Klasse sind beide Varianten nicht mehr paradigmatisch austauschbar und erfordern unterschiedliche syntaktische Angaben wie externe Leerstellen. Daher ist es sinnvoll, beide Formen in getrennten Einträgen zu speichern und im Feld *Varianten* jeweils einen Querverweis zur anderen Form zu legen.

Antonyme sind für Übersetzer dann interessant, wenn sich für den gesuchten Ausdruck kein Äquivalent findet, wohl aber das Gegenteil verneint werden könnte. Allerdings stehen nur die wenigsten Phraseologismen in eindeutig antymischer Beziehung zueinander; zudem lassen sich viele Phraseologismen überhaupt nicht negieren. Daher käme eher eine weitergefasste Klasse von **verwandten Phraseologismen** in Betracht, unabhängig von der Art ihrer Beziehung. Hier können dann z.B. auch Phraseologismen verknüpft werden, die über eine ähnliche Bildhaftigkeit verfügen (z.B. Licht-Intelligenz-Analogie wie bei *jdm geht ein Licht auf, es dämmert, keinen blassen Schimmer haben, in geistiger Umnachtung, Geistesblitz, einen Schatten haben*, etc.), oder andere Zusammenhänge, die keine Nennung unter Varianten rechtfertigen. Eine solche offene Lösung wird der Komplexität der Materie am ehesten gerecht.

Eine phraseologische Datenbank sollte außerdem Raum für eine veranschaulichende **Graphik** geben, die insbesondere bei fachsprachlichen Ausdrücken von großem Nutzen sein kann. Bei gemeinsprachlichen Phraseologismen wird dies kaum vonnöten sein.

5.4.1.4 Funktionale Angaben

Die funktionalen, d.h. gebrauchsbezogenen Informationen sind für die praktische Einsetzbarkeit eines Phraseologismus in der translatorischen Praxis und beim Fremdspracherwerb von besonderer Bedeutung.

Eine wichtige Angabe ist die **stilistische Konnotation**, also die Frage, ob ein Phraseologismus besonders literarisch, formell, vulgär etc. ist. Stilistische Unterschiede können sich auf verschiedenen Ebenen²⁹ ergeben:

Stilschicht: gehoben, gemeinsprachlich, umgangssprachlich, Slang, vulgär

Wirkung: literarisch, formell, journalistisch³⁰

Aktualität: veraltet, veraltend, Modeerscheinung

Gruppenspezifitk: Dialekt, Jugendsprache, Gaunersprache

Medium: mündlich, schriftlich

Anwendungsbereich: fachsprachlich, gemeinsprachlich

Bewertung: euphemistisch, pejorativ, ironisch

Diese Auflistung könnte natürlich weiter ergänzt und ausdifferenziert werden.

Für die Dateneingabe im Wörterbuch ist es von Vorteil, alle obigen Attribute in ein Feld zusammenzulegen, dabei allerdings vordefinierte Werte zur Auswahl zu stellen (s. 5..3). Eine Zusammenlegung ist angezeigt, weil eine Überdifferenzierung für Anwender unübersichtlich sein kann und sich die Ebenen zudem vielfach überlappen und einander bedingen. So sind *fachsprachliche* Wendungen in der Regel auch *formell*, *literarische* auch *schriftlich* und *vulgäre* dem *mündlichen* Sprachgebrauch zuzuordnen. Dadurch, dass die Angaben nicht in Textform, sondern mittels Auswahl aus vordefinierten Optionen erfolgen, ist dennoch Granularität gewährleistet.

Anzumerken ist außerdem, dass mehrere der stilistischen Merkmale zutreffen können und Mehrfachnennungen daher möglich sein sollen. Die Kombination mehrerer Merkmale liefert ein realistisches Bild vom Gebrauch des Phraseologismus; so ließe eine gleichzeitige Nennung von *fachsprachlich* und *gemeinsprachlich* darauf schließen, dass es sich um einen weithin bekannten fachsprachlichen Ausdruck bzw. fachsprachnahe Gemeinsprache handelt (*ein gleichschenkliges Dreieck, Gewinn abwerfen*).

²⁹ Einige Bezeichnungen für stilistische Ebenen wurden Koller (2004:243ff.) entnommen.

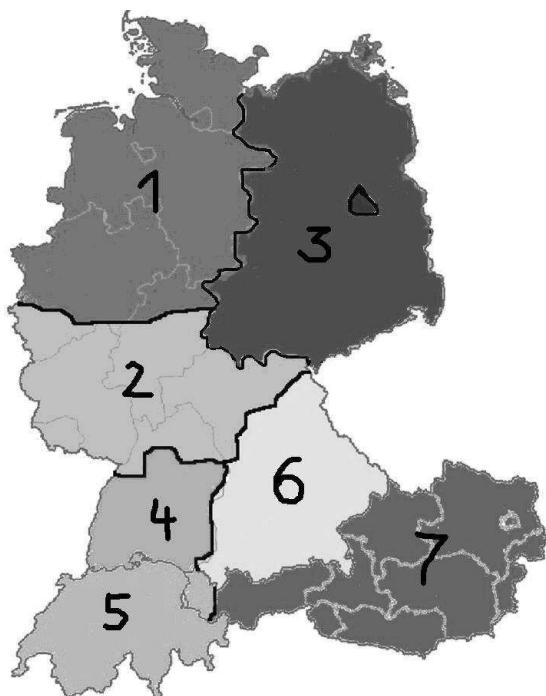
³⁰ Damit sind ausdrucksstarke Phraseologismen gemeint, die häufig in journalistischen Texten verwendet werden, um dem Text eine gewisse „Würze“ zu verleihen ([etw] erhitzt die Gemüter, [etw] kommt unter den Hammer etc.)

Eine weitere wichtige Angabe ist das **Fachgebiet**. Hier bietet sich die Integration einer bestehenden Fachgebietssystematik, wie die Lenoch-Klassifikation an (vgl. Arntz et al.⁵ 2004:234, 241).

Interessant sind auch Angaben zur **arealen Verbreitung** eines Phraseologismus. Für Translationszwecke ist hierbei eine Grobgliederung eines Sprachraums am übersichtlichsten, die einen Eindruck vermittelt, ob ein Phraseologismus in der Zielregion bekannt ist oder nicht, ohne allzu sehr ins Detail zu gehen. In die Unterteilung sollten sowohl Dialektgruppen als auch Landesgrenzen einfließen, da der Phraseologismenbestand innerhalb eines Landes durch nationale Medien geprägt sein kann. Eine mögliche Einteilung des deutschsprachigen Raums zeigt Abb. 4. Hierbei handelt es sich um einen Vorschlag zur Veranschaulichung der Idee; dieser wäre aus varietätenlinguistischer Sicht genauer auszuarbeiten.

1. Niederdeutsch + Friesisch
2. Mitteldeutsch
3. Nieder- und Mitteldeutsch
(neue Bundesländer, Berlin)
4. Alemannisch (Deutschland und Elsaß)
5. Alemannisch (Schweiz und Westösterreich)
6. Bairisch (Deutschland)
7. Bairisch (Österreich)

Abb. 4: Areale Untergliederung des deutschsprachigen Raums



Zur Ermittlung der relativen Gebräuchlichkeit eines Phraseologismus in einer bestimmten Region ist die User-Befragung mit anschließender statistischer Auswertung ein vielversprechender Ansatz, der in 5.7.3.1 näher untersucht wird. Wenn diese Möglichkeit in einer Datenbank jedoch nicht besteht, ist die areale Verbreitung kaum objektiv feststellbar. In diesem Fall ist es daher ratsamer, dieses Feld nur für offensichtliche Regionalismen (*[etw] is g'hupft wie g'hatscht*) zu nützen und es in allgemeineren Fällen freizulassen.

Besonders wichtig ist die Angabe von **Kontextbeispielen** in Form von ein bis drei Sätzen. Gerade für linguistisch weniger versierte Anwender dürften solche Beispiele weitaus aussagekräftiger sein als ein abstrakter Formalismus über syntaktische Klasse

und Kontextanbindung. Zu den Kontextbeispielen gehört jeweils eine **Quellenangabe**; hierfür sollte unter diese Datenkategorie eine eigene Unterdatenbank mit diversen Datenkategorien für genaue bibliographische Angaben diverser Medien (Literatur, Filme, Internet, etc.) gekoppelt werden. Durch das Anfügen mehrerer Kontextbeispiele ergibt sich im Laufe der Zeit ein Bild über die Textsorten, in denen dieser Phraseologismus typischerweise auftritt.

Solche **typische Textsorten** können auch in einer eigenen Kategorie von vornherein explizit genannt werden. Wird etwa ein Ausdruck eingetragen, der aus einem Vertrag oder Rechtstext stammt, so dürfte schnell klar sein, ob ein Ausdruck vielleicht spezifisch für genau diese Textsorte ist. In weniger offensichtlichen Fällen kann diese Angabe freigelassen werden.

In vielen Fällen ist auch der **Ursprung eines Phraseologismus** von Interesse, etwa dann, wenn es sich um ein Geflügeltes Wort aus der Literatur (Bibel, Goethe, Schiller, Shakespeare etc.) handelt. Auch multimediale Quellen können angegeben werden, etwa bei Zitaten aus Film und Werbung (*Geschüttelt, nicht gerührt* aus den James-Bond-Filmen). Bei nicht belegbaren Phraseologismen wird diese Rubrik freigeben. Der Ursprung bzw. der ursprüngliche Verwendungskontext liefert eine mögliche Übersetzungsquelle, etwa bei einem in beiden Sprachen vorliegenden Werk.

Weitere spezifischere Angaben zum Gebrauchskontext, die durch die beschriebenen Datenkategorien noch nicht erfasst wurden, können in einem Feld **Bemerkungen zum Gebrauch** erfasst werden.

5.4.1.5 Semantische Angaben

Außer der Definition, die weiter oben beschrieben wurde, fallen weitere Datenkategorien in den Bereich der Semantik, insbesondere der **Grad der Idiomatizität**, der mittels der in 2.1.3 beschriebenen Einteilung in Idiome, Teil-Idiome und Kollokationen zum Ausdruck gebracht wird. Dies dient als erste Orientierung, ob ein Phraseologismus in ungefähr wörtlich verstanden werden kann oder nicht. Eine nähere Spezifizierung der Bedeutung ist nicht nötig, da diese der Definition entnommen werden kann. Genauere Angaben über den Zusammenhang zwischen wörtlicher und phraseologischer Bedeutung (Metapher, Metonymie, Vergleich etc.) sind für Linguisten interessant, nicht aber für Übersetzer und Fremdsprachenlernende.

Interessant wäre auch eine Suche nach **Bedeutungsfeldern**, um Phraseologismen mit ähnlicher Bedeutung anzeigen zu lassen. So würde eine Suche nach dem Begriff *aufgeben* die Phraseologismen *die Flinte ins Korn werfen*, *den Kopf in den Sand stecken* und *das Handtuch werfen* ergeben. Dies könnte über die **Integration eines semantischen Netzes oder eines Thesaurus** realisiert werden. Solange dies nicht geleistet ist, kann eine semantische Suche über eine entsprechende Beschlag-

wortung erfüllt werden (s. 5.2), und/oder durch eine Suche innerhalb des Feldes *Synonym/Paraphrase*.

Auch eine komplette formale **Repräsentation der Bedeutung** durch ein integriertes semantisches Netz wäre denkbar. Auf dessen Basis lässt sich auch die **semantische und syntaktische Valenz**, also externe Leerstellen einer phraseologischen Einheit und deren Merkmale, formal beschreiben. Diese Ansätze sind insbesondere für maschinelle Sprachverarbeitung von Interesse und werden daher im Rahmen dieser Arbeit nicht ausführlich behandelt. Zur weiteren Information über semantische Netze sei auf Carstensen et al. (³2010:377ff., 504ff.) verwiesen. Zentral für die formale Beschreibung semantischer Relationen sind auch die Arbeiten von Mel'čuk zur *Bedeutung–Text–Theorie* und *Lexikalischen Funktionen* (z.B. Mel'čuk 1998; Mel'čuk 2007).

5.4.1.6 Syntaktische Angaben

In den Bereich der Syntax fallen drei Datenkategorien, nämlich die syntaktische Klasse, die Art der Kontextanbindung und externe Leerstellen. In diesen Datenfeldern bietet sich die Vorgabe möglicher Optionen an, unter denen eine Auswahl getroffen wird. Aus den in 3.1.3.1 beschriebenen Klassen ergeben sich folgende Werte:

Syntaktische Klasse: Satz, satzwertiges Gebilde, Satz mit Leerstelle, Nebensatz, Nominalphrase (NP), Verbalphrase (VP), Prädikativ, Adjektivphrase (AdjP), Adverbialphrase bzw. Ergänzung (AdvP), Präpositionalphrase (PP), Postposition, Interjektion, Konjunktion, Subjunktion.

Kontextanbindung: keine, logisch, deiktisch, durch Satzglied

Externe Leerstellen: Subjekt, Prädikat, Akkusativobjekt, Dativobjekt, Genitivobjekt, Ergänzung, Attribut; NP, AdvP etc.

Die Zuordnung zu einem bestimmten Konstituententyp erfolgt nach dem in 3.1.4 beschriebenen Aufrundungsprinzip: Externe Leerstellen, die zur Bildung der nächsthöheren vollständigen Konstituente benötigt werden, werden bei der Klassifizierung zum Phraseologismus hinzugerechnet, andere nicht. Demgemäß können die genannten Klassen auch mit Leerstellen auftreten, z.B. *NP mit Leerstelle*, *VP mit Leerstelle* usw.

Die syntaktische Klassifizierung des Phraseologismus beruht nur auf dem Konstituententyp, da sich bzgl. der Funktion im Satz mehrere Möglichkeiten ergeben können (s. 3.1.3.1). Bei externen Leerstellen hingegen ist die Satzgliedfunktion i.d.R. schon festgelegt (so wird *[jdn] hinter's Licht führen* nicht durch eine beliebige NP, sondern durch ein Akkusativobjekt ergänzt). Daher enthält die Kategorie *externe Leerstellen* nicht nur Konstituentenbezeichnungen, sondern auch Satzglieder. Alternativ kann man

die jeweilige Konstituente mit Merkmalen versehen, z.B. statt *Dativobjekt NP_{fun=obj, case=dat}* (wobei *fun* für „Funktion“ und *case* für den Fall steht).

In anderen Fällen hat die externe Leerstelle keine eigenständige Satzgliedfunktion inne – in *an [jdm] einen Narren gefressen haben* muss eine NP ergänzt werden, die erst in Kombination mit *an* eine Ergänzung darstellt – daher enthält die Liste der externen Leerstellen auch einfache Konstituenten.

Wie in 3.1.3.3 beschrieben, wird zwischen **obligatorischen** und **fakultativen externen Leerstellen** unterschieden. Die Nennung einer fakultativen Leerstelle ist nur dann angebracht, wenn deren Besetzung besonders typisch ist, wie etwa bei der Erweiterung auf die Langform eines Phraseologismus (s. 5.4.1.3), z.B. *[seine] sieben Sachen → [seine] sieben Sachen packen*. Eine Liste sämtlicher fakultativer Leerstellen wäre endlos erweiterbar und daher sinnlos.

Den externen Leerstellen soll eine Datenkategorie **Restriktionen** untergeordnet werden, die Einschränkungen in der Besetzung der Leerstellen enthält. Dabei ist zwischen morphosyntaktischen und semantischen Restriktionen zu unterscheiden. Wie in 5.1.4.5 erwähnt, ist bei maschineller Verwendung des Wörterbuchs eine formale Beschreibung solcher Restriktionen durch ein integriertes semantisches Netz hilfreich. Solange diese Möglichkeit nicht besteht, werden syntaktische Restriktionen durch eine Auswahl vordefinierter Werte dargestellt; semantische Restriktionen werden hingegen in Textform eingegeben – eine Darstellungsweise, die natürlich nur für menschliche Benutzer verwertbar ist.

Zu den **syntaktischen Restriktionen** zählen Angaben wie syntaktische Funktion, Kasus, oder – bei einer Ergänzung – die Präposition. Je nach Notationsweise können syntaktische Restriktionen bereits aus der Bezeichnung der externen Leerstelle selbst hervorgehen.

Die **semantischen Restriktionen** bestehen in bestimmten Merkmalen wie <belebt>, <unbelebt>, <konkret>, <abstrakt> etc., die das zu ergänzende Satzglied erfüllen muss. Die Angaben können auch spezifischer sein, z.B. muss das Subjekt bei *in Kraft treten* ein Gesetz, eine Vorschrift oder ähnliches sein. Die Restriktion kann so weit gehen, dass eine Leerstelle nur mit wenigen, bestimmten Lexemen besetzt werden kann. (In einem solchen Fall stellt sich die Frage, ob man es mit einem weiteren Phraseologismus zu tun hat; die Übergänge zwischen Kookkurrenz und Phraseologizität sind hier fließend.)

Aufgrund der Festigkeit von Phraseologismen kann die Möglichkeit zu morphosyntaktischen Abwandlungen eingeschränkt sein, man spricht von **transformationellen Defekten**. Rothkegel (2007:1034) unterscheidet Einschränkungen bei folgenden Transformationen: „Passivtransformation, Relativsatztransformation, Imperativsatz, Nominalisierung, Satznegation, Fragesatztransformation, Konstituentennegation, Ex-

pansion durch Attribute, Quantifizierung“. Zu ergänzen wäre hierbei noch die Abwandlung des Tempus (beispielsweise kann *einen Narren an jdm gefressen haben* nur im Perfekt stehen). Einige dieser Restriktionen gehen bereits aus der Nennform hervor (z.B. die Konstituentennegation in *auf keinen grünen Zweig kommen*; s. 5.1.4.2); andere müssen explizit genannt werden, wie etwa die Zulässigkeit oder Unzulässigkeit der Pluralbildung bei Nominalphrasen (*ein kluger Kopf/ kluge Köpfe* vs. *Stein des Anstoßes*) Zugunsten der Einheitlichkeit ist eine vollständige Auflistung in dieser Kategorie jedenfalls angezeigt.

5.4.1.7 Weitere interessensspezifische Angaben

In den obigen Abschnitten wurden einige für die Translation relevante Basisinformationen aufgezählt. Für weitere Interessensgruppen wie Linguistinnen, Sprachdidaktinnen, Terminologinnen etc. können noch präzisere Angaben von Interesse sein. Diese werden im Folgenden kurz genannt.

Für die **Linguistik** kann die **semiotische Unterscheidung** zwischen strukturellen, referentiellen, kommunikativen und onymischen Phraseologismen (s. 3.1.2) oder der **innere morphosyntaktische Aufbau** einschließlich Auffälligkeiten wie Zwillingsformel, Reim, etc. (3.1.4.4) von Interesse sein. Auch die Art der **Beziehung zwischen phraseologischer und wörtlicher Lesart** sowie die **Etymologie** von Phraseologismen sind für die Phraseologieforschung relevant. Eine eigene Kategorie könnte auch für **sprachliche Auffälligkeiten** aller Art eröffnet werden, wie unikale Komponente (*etwas auf dem Kerbholz haben*), Irregularität bei Valenz (*Löcher in die Luft starren*), auffällige Syntax (*in Teufels Küche kommen*) oder zwei parallel realisierbare Lesarten (*sich die Zunge verbrennen*).

Weiterhin kann die **traditionelle phraseologische Klassenzuordnung** wissenswert sein, also Angaben wie Sprichwort, Geflügeltes Wort, Gemeinplatz, Redewendung, komparativer Phraseologismus, Zwillingsformel. Diese Klassen eignen sich zwar nicht für eine konsistente Systematik, sind aber aufgrund ihrer weiten Verbreitung nennenswert.

Für den **Spracherwerb** ist die **Aussprache**, idealerweise als Tondatei, eine wichtige Information.

Auch im **fachsprachlichen Bereich** können weitere Details von Interesse sein. Außer den bereits beschriebenen Kategorien führen Budin/Galinski (1992) unter anderem **Symbol**, **Phrasembestimmung (Prozessdefinition)**, **Inhouse-Phraseme**; **Trademark und Warenzeichen** sowie **genormte Phraseme** an.

5.4.1.8 Zielsprachliche Angaben zum Phraseologismus

Wie bereits zu Beginn dieses Kapitels erwähnt, sind aus Anwendersicht auch eine **zielsprachliche Definition und Paraphrase** hilfreich. Auch Dobrovolskij (1999:116) zieht dies Möglichkeit in Betracht. Den Nutzen dieser zielsprachlichen Angaben zeigt folgende Überlegung: Angenommen, eine zweisprachige Phraseologiedatenbank wird in der Sprachrichtung Ungarisch-Deutsch abgefragt. Diese Abfrage ist für zwei Gruppen von Anwendern von Interesse, nämlich solche, die

- a) passiv einen ungarischen Text rezipieren wollen und dabei eine Verständnishilfe benötigen (z.B. Deutsch-Muttersprachler, Ungarisch-Deutsch-Übersetzer);
- b) aktiv einen deutschen Text produzieren wollen (Ungarisch-Muttersprachler mit Deutsch als Fremdsprache; Ungarisch-Deutsch-Übersetzer).

Für beide Gruppen hat eine deutsche Definition Vorteile. Gruppe a) bemüht sich um passives Verständnis, und dies gelingt am besten mit Hilfe einer möglichst präzisen Erläuterung in der eigenen Muttersprache (die auch auf eventuelle begriffliche Asymmetrien zum vorgeschlagenen Äquivalent eingeht). Gruppe b) hingegen möchte aktiv in die Fremdsprache Deutsch übersetzen. Sollte für den ausgangssprachlichen Phraseologismus kein genaues Äquivalent in der Zielsprache vorliegen, so ist es für Benutzer b) von Interesse, eine Umschreibung des Phraseologismus aus zielkultureller Sicht zu finden, die eine Anregung für die zielsprachliche Formulierung liefern kann.

Eine weitere hilfreiche Angabe kann ein der Verweis auf einen **zielsprachlichen verwandten Phraseologismus** darstellen, z.B. auf ein Antonym oder einen Phraseologismus mit ähnlicher Metaphorik, insbesondere dann, wenn kein zielsprachliches Äquivalent vorliegt.

Auf den ersten Blick scheinen diese zielsprachlichen Datenfelder besser zu den übersetzungsbezogenen Daten zu passen. Da sie sich jedoch ausschließlich auf den Ausgangsphraseologismus beziehen und noch an keine spezifische Übersetzung gekoppelt sind, ist eine Speicherung unter den einsprachigen Daten sinnvoller. Zudem sind die genannten Angaben dann auch sichtbar, wenn noch kein zielsprachliches Äquivalent bekannt ist (Null-Äquivalenz) und somit auch kein Übersetzungsdatensatz angelegt wird. Ein phraseologischer Eintrag kann Definition, Paraphrase und einen Verweis auf verwandte Phraseologismen in vielen Zielsprachen enthalten, die bei Bedarf ausgeblendet werden.

5.4.2 Übersetzungsbezogene Datenkategorien

Im Folgenden wird dargestellt, welche Angaben in einem übersetzungsbezogenen Datensatz gemacht werden müssen, der zwei Phraseologismen unterschiedlicher Sprachen miteinander verbindet.

Zum einen ist die Angabe notwendig, ob zwischen den verknüpften Ausdrücken **totale, pragmatische** oder **partielle Äquivalenz** vorliegt. Bei Null-Äquivalenz wird gar kein übersetzungsbezogener Datensatz angelegt. Die Äquivalenz muss jeweils für beide Übersetzungsrichtungen getrennt angegeben werden, da sich diesbezüglich Asymmetrien ergeben können. Unter dieser allgemeinen Angabe zur Äquivalenz soll ein Textfeld für einen **Kommentar zur Äquivalenz** angelegt werden, in dem möglichst genau darüber Aufschluss gegeben wird, wann diese Übersetzung angezeigt ist und wann nicht. Hier kann eine genaue Beschreibung des erforderlichen Kontextes und stilistischer Markierung erfolgen. Zudem besteht Raum für Beobachtungen, die sich speziell aus dem eigenen sprachlich-kulturellen Hintergrund ergeben und die für einen Muttersprachler nicht relevant wären. Dieser Kommentar kann je nach Präferenz in der einen oder in der anderen Sprache verfasst werden; es sollte daher für beide Übersetzungsrichtungen jeweils ein Feld für jede Sprache zur Verfügung stehen.

Außerdem muss ein Feld für **Übersetzungsbeispiele** vorhanden sein, um darzustellen, in welchen Fällen genau diese Übersetzung gewählt wurde. Auch dieses Feld muss für beide Übersetzungsrichtungen zur Verfügung stehen und zudem durch Quellenangaben belegt werden.

5.4.3 Übersicht Datenkategorien

In diesem Kapitel wurden sowohl einsprachige als auch translationsspezifische Datenkategorien zusammengestellt, die für mehrsprachig Arbeitende von Bedeutung sind. Diese Ergebnisse sind in Abb. 5 zusammengefasst. Hierbei handelt es sich um eine Erweiterung der Abb. 3 (s. 5.3), angereichert um die besprochenen Datenkategorien im einsprachigen sowie im übersetzungsbezogenen Datensatz.

5.4.4 ER-Modell

Abb. 6 zeigt ein Schema der Datenbank nach dem Entity-Relationship-Modell (vgl. Kemper/Eickler 2006). Die Verknüpfung zwischen zwei Phraseologismen wurde als Beziehung (Raute) dargestellt; die Übersetzungsdaten sind durch Attribute (Ellipsen) dieser Beziehung vertreten. Die einsprachigen Datenkategorien bilden Attribute zu den jeweiligen Phraseologismen; diese wurden vereinfachend dargestellt. Jeder L1-Eintrag kann mit mehreren L2-Einträgen verknüpft werden und umgekehrt (n:m-Relation).

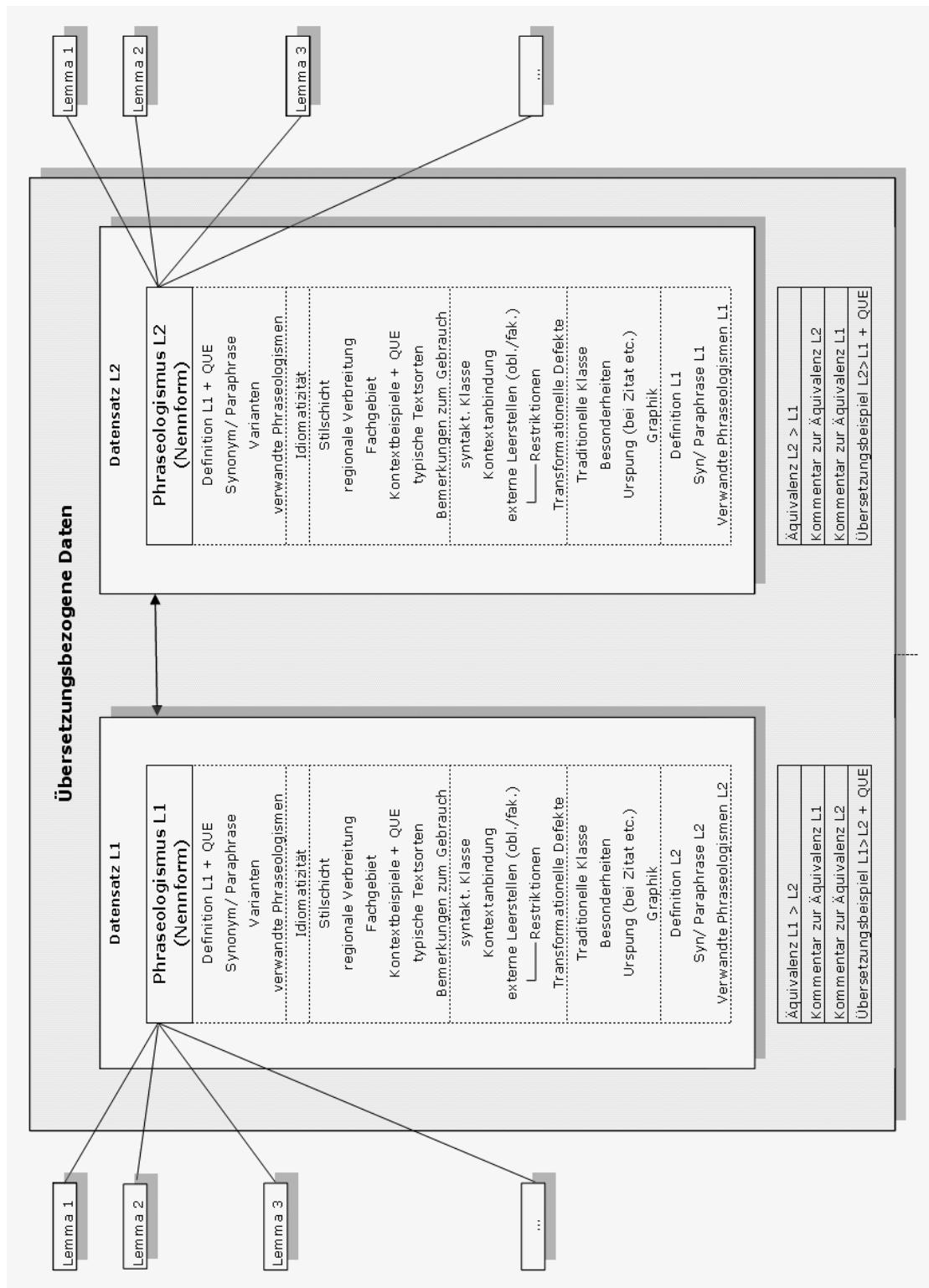


Abb. 5: Vollständige Eintragsstruktur des phraseologischen Wörterbuchs

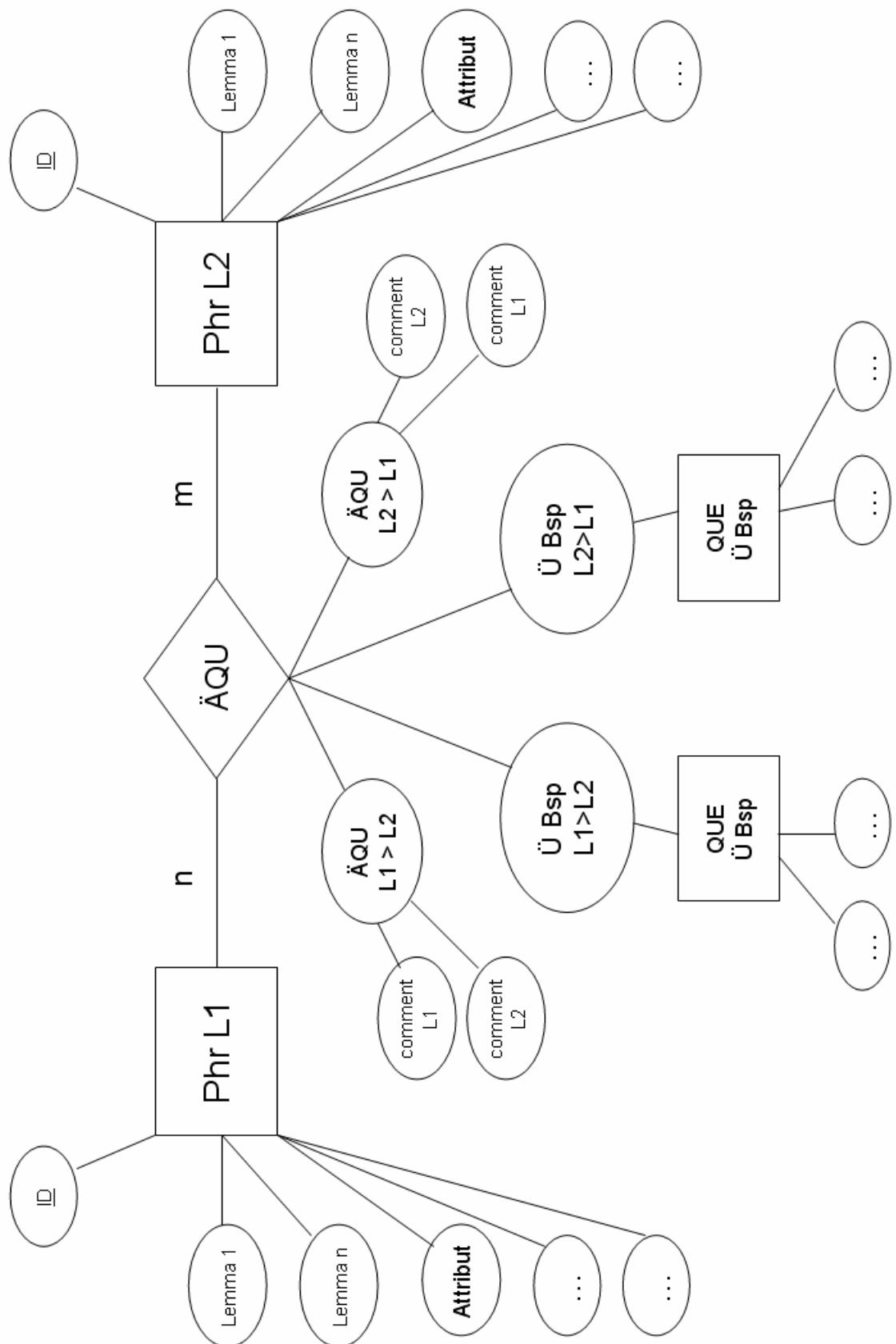


Abb. 6: ER-Diagramm

5.5 Contenterstellung und Bearbeitung

5.5.1 Datengewinnung

Nachdem die Art und Anordnung der Daten besprochen wurde, wird nun diskutiert, woher der phraseologische Datenbestand überhaupt kommt, der sich aus den Phraseologismen selbst, den gebrauchsbezogenen Angaben sowie möglichen Übersetzungsbereihungen zusammensetzt. In allen Bereichen gibt es mehrere Möglichkeiten: die Datenerstellung durch ein **Redaktionsteam**, die zumeist ‚Recycling‘ aus Vorgängerwerken beinhaltet, **Datenextraktion aus Korpora** sowie das **Mitwirken von Userinnen**. Im letzten Fall gibt es entweder die Möglichkeit, durch Online-Befragungen Informationen zu gewinnen und diese dann zu übernehmen, oder aber Userinnen direkt interaktiv an der Contenterstellung zu beteiligen.

Vergleicht man diese Methoden, liegt es auf der Hand, dass Korpusanalysen und (sinnvoll gestaltete) User-Beteiligung ein repräsentativeres Bild des aktuellen Sprachgebrauchs liefern als die bloße sprachliche Intuition eines Expertenteams. Die introspektive Methode birgt die Gefahr, dass Angaben zum tatsächlichen Gebrauch – die Frequenz, häufige Varianten und funktionale Aspekte – zu wenig berücksichtigt werden. Zudem können sich durch die gängige Praxis, Daten aus Vorgängerwerken zu übernehmen, mittlerweile veraltete Phraseologismen einschleichen, ohne entsprechend gekennzeichnet zu werden.

Der **Einsatz von Korpora** bewirkt nicht nur eine erhebliche Verbesserung der inhaltlichen Qualität, sondern spart auch Zeit bei der Sammlung von Daten, die dann direkt in elektronischer Form zur Weiterverarbeitung bereit liegen. Die Qualitätsverbesserung ergibt sich im einsprachigen Bereich durch das Auffinden von gebräuchlichen Varianten, von typischen Kontexten sowie Aussagen über die Frequenz und bevorzugtes Vorkommen in bestimmten Textsorten (vgl. Dräger/Juska-Bacher 2010).

Demnach ist die Korpusanalyse gut geeignet, um einen Datenbestand von Grund auf aufzubauen, jedoch kann sie alleine noch nicht alle Anforderungen an phraseographische Arbeit abdecken. Vielmehr ist die Korpusanalyse, wie Lubensky/McShane (2007:920) sich ausdrücken, „just a tool like any other“ und bedarf in mehrfacher Hinsicht weiterer Bearbeitung: Neben der Selektion von Treffern ist eine Strukturierung der Daten und Ergänzung weiterer Detailinformation erforderlich, um eine differenziertere Darstellung von Daten im Sinne der in 5.1 beschriebenen Granularität zu erzielen.

Die Qualität der gewonnenen Daten steht und fällt mit der Qualität der verwendeten Korpora, insbesondere der Qualität der Übersetzungen in Parallelkorpora. Die Qualität einsprachiger Korpora wächst unter anderem mit deren Größe und einer repräsentativen Auswahl an Textsorten. Selbst in großen Korpora fehlen manche relevanten Phra-

seologismen, insbesondere „weite Teile der nicht schriftlich dokumentierten Sprachverwendung“ (Dräger/Juska-Bacher 2010:167), die daher manuell nachgetragen werden müssen.

Die Eingabe durch Experten und die Korpusanalyse schließen einander daher keinesfalls aus, sondern können ergänzend eingesetzt werden, wie auch Lubensky/McShane (2007:920) ausführen:

„there is no reason to put dictionaries and parallel corpora in competition since they have different strengths and weaknesses. Whereas it is true that parallel corpora can show more contexts than are possible in dictionaries, they are also full of noise, including incorrect and imprecise translations, and they do not provide the detailed description possible through the introspection of a highly-trained lexicographer.“

Eine weitere, bislang noch „nicht systematisch nutzbar gemachte“ Form der Datengewinnung (Dräger/Juska-Bacher 2010:175) besteht in der **Online-Befragung von Benutzern**. Diese bietet die Möglichkeit, immanentes Sprachwissen einer breiteren Bevölkerungsschicht einzufangen und zu kanalieren. Die Einbeziehung von Benutzerinnen kann jedoch nicht nur durch Umfragen im Vorfeld erfolgen, sondern auch in Form von **interaktiver Mitarbeit am Wörterbuch** (s. 5.5.2). Dabei können Nutzerinnen über eine vordefinierte Eingabemaske Phraseologismen um verschiedene Angaben ergänzen oder gar selbst Einträge angelegen. Hierbei ist jedoch die Einplanung von Mechanismen zur gegenseitigen Bewertung sowie die Administration durch Expertinnen erforderlich.

Somit wäre eine **Hybridlösung** optimal, bei der die Vorteile verschiedener Methoden kombiniert werden: Korpusanalysen liefern einen Basisbestand an Daten, der durch Expertinnen und Userinnen manuell nachbearbeitet und ergänzt wird, wobei die Administration und ggf. Anpassung der Datenbankstruktur ebenfalls durch ein Expertenteam erfolgt.

In jedem Fall muss die Datenbank eine **Möglichkeit der manuellen Eingabe** beinhalten, unabhängig, ob diese für Neueinträge oder nur zur Bearbeitung von Daten genutzt wird. Dafür sollte eine Eingabemaske angelegt werden, ein Formular, das bestimmte Felder enthält. Wie diese im Einzelnen aussehen, wird in 5.5.3 erörtert.

Weiterhin muss die **Möglichkeit der kollaborativen Bearbeitung** gegeben sein, da eine Datenbearbeitung dieses Ausmaßes kaum von einer Person bzw. von einem Arbeitsplatz aus zu bewältigen ist. Näheres hierzu findet sich in 5.5.2.

Bei kollaborativer Bearbeitung ist eine **klar strukturierte, verständliche Gestaltung der Benutzeroberfläche** unabdingbare Voraussetzung für die korrekte Eingabe und damit für die Qualität der Daten. In 5.5.5 werden daher Aspekte der Usability besprochen.

5.5.2 Online-Kollaboration

Wie erwähnt, müssen bei einer Datenbank dieses Umfangs die nötigen Strukturen geschaffen werden, um Web-Kollaboration zu ermöglichen, einmal zur Bewältigung des hohen Arbeitsaufwands, zum anderen zur Gewährleistung von Aktualität. Die Möglichkeit der kollaborativen Bearbeitung sollte zumindest einem Team von Expertinnen zur Verfügung stehen. Zusätzlich kann auch die Einbeziehung von Userinnen großen Nutzen bringen. Diese ist in verschiedene Abstufungen denkbar:

Entweder das Wörterbuch enthält bereits einen Bestand, der nur von Experten bearbeitet werden kann, und Benutzer haben die Möglichkeiten, in Umfragen **gebrauchsbezogene Angaben zu Einträgen** wie Bewertung der Stilschicht und Textsorte, oder areale Verbreitung zu machen, die dann statistisch ausgewertet werden. Diese Möglichkeit wird in 5.5.3.1 näher besprochen.

Auch ein weiterreichendes **Ausfüllen von Datenfeldern** durch User wäre denkbar, wie etwa Definition, Fachgebiet, Anführung von Kontext- und Übersetzungsbeispielen. Bei einigen Feldern, insbesondere der Definition, können dabei bestehende Einträge überarbeitet werden, so dass nach dem Wiki-Prinzip der Konsens über den endgültigen Inhalt entscheidet. Bei anderen Daten sind parallel bestehende Mehrfachangaben sinnvoll; beispielsweise könnten zu einem Eintrag von verschiedenen Usern mehrere Kontext- oder Übersetzungsbeispiele hinzugefügt werden (zum Eingabemodus vgl. 5.5.3).

Schließlich wäre bei entsprechender Administration auch **eine komplette Dateneingabe durch Userinnen** möglich, wobei phraseologische Daten in ein leeres Formular mit vordefinierten Feldern eingetragen werden. Hierfür ist eine klar verständliche Benutzeroberfläche mit Hilfe-Funktionen und Online-Tutorials eine wesentliche Voraussetzung (s. 5.5.5).

Eine gangbare Möglichkeit kann eine **gestaffelte Vergabe von Zugriffsrechten** darstellen, die Usern mit längerer Erfahrung oder Nachweis einer sprachbezogenen Tätigkeit erweiterten Zugriff oder ein stärkeres Stimmengewicht einräumt als Neueinsteigern.

Generell ist zu sagen, dass eine weit reichende Einbeziehung von Nutzerinnen großen Gewinn bringt, was Aktualität und Repräsentativität der Daten betrifft. Umso mehr Rechte Nutzerinnen jedoch eingeräumt werden, desto intensivere Planung und Administration ist für eine konsistente Führung der Datenbank erforderlich. Beispielsweise muss die Oberfläche so gestaltet sein, dass Inhalte auch bei mangelndem Fachwissen verständlich sind; weiterhin müssen Online-Hilfen und Support zur Verfügung stehen. Die Entscheidung, wie weit User in die Dateneingabe und -bearbeitung einbezogen werden, ist also nicht zuletzt vom veranschlagten Budget und der Anzahl fachlicher Mitarbeiter abhängig.

5.5.3 Eingabe und Repräsentation der Daten

Nachdem feststeht, welche Datenkategorien aufgenommen werden, muss bestimmt werden, durch welchen Datentyp diese in der Datenbank repräsentiert sind. Damit geht auch die Frage nach dem jeweiligen Eingabemodus bei der Bearbeitung einher. Daher werden diese beiden Faktoren gemeinsam behandelt.

Grundsätzlich kann man verschiedene **Datentypen** unterscheiden, je nach dem, wie die Information strukturiert ist. Manche Daten werden in Form von **Freitext** eingegeben, wie etwa die Definition oder die Nennform des Phraseologismus selbst. In anderen Fällen ist es sinnvoll, eine endliche Zahl von Vorschlägen zur Auswahl zu stellen. Ein Beispiel hierfür ist die Stilschicht, bei der Begriffe wie *umgangssprachlich*, *literarisch*, *formell* etc. zur Verfügung stehen, oder auch die Wahl des Fachgebietes.

Bei der Auswahl aus vordefinierten Optionen besteht weiterhin die Möglichkeit, entweder nur eine **einfache Nennung**, oder aber eine **Mehrfachnennung** zuzulassen. So ist es bei den externen Leerstellen unbedingt erforderlich, dass mehrere Angaben gemacht werden können, da ein Phraseologismus meist über mehrere externe Leerstelle verfügt. Einfachnennungen bieten sich an, wenn verschiedene Möglichkeiten einander ausschließen – beispielsweise kann ein Phraseologismus nicht gleichzeitig Idiom und Kollokation sein.

Im Sinne der Granularität und somit der Maschinenlesbarkeit (s. 5.1) wäre es wünschenswert, **möglichst viele Felder als Optionsfelder** mit definierten Werten darzustellen. Dies setzt aber eine erschöpfende Typologie möglicher Inhalte voraus, die im Rahmen dieser Arbeit nur exemplarisch geleistet werden kann. Einige Felder erfordern die Integration weiterer Subsysteme, z.B. eines semantischen Netzes, einer Fachsprachenklassifikation oder Textsortentypologie.

Weiterhin kann ein Datenfeld auch eine **Verknüpfung** mit einem anderen Eintrag (z.B. bei Varianten) sowie eine **Multimedia-Datei** (Aussprache, Graphik) enthalten.

Je nach Datentyp bietet sich eine andere Art von Feld in der Eingabemaske an. Für Freitext wird ein **Textfeld** angelegt, Mehrfachnennungen werden über **Kontrollkästchen** (Checkboxes) realisiert. Einfachnennungen können über ein **Drop-Down-Menü** oder **Radiobuttons** erfolgen (vgl. Abb. 7). Radiobuttons haben den Vorteil, dass daneben noch Platz ist für eine Erläuterung der gefragten Information.

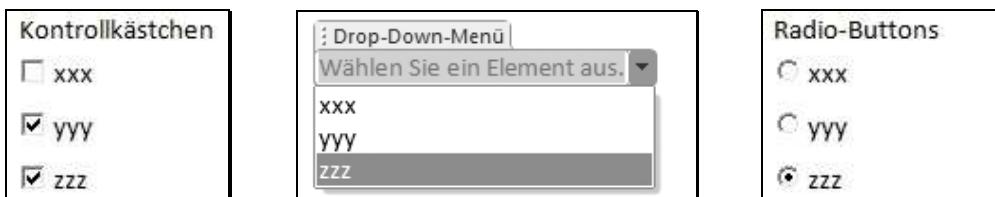


Abb. 7: Feldtypen in der Eingabemaske

In der Anfangsphase eines phraseographischen Projekts sollten Felder mit Einfachnennungen sparsam eingesetzt werden, da vielleicht noch nicht alle Fälle bedacht wurden, die im weiteren Verlauf auftreten können. Einfachnennungen führen zu Verzerrungen, wenn mehrere Möglichkeiten gleichzeitig bestehen – etwa, wenn ein Phrasenologismus gleichzeitig mehreren Stilschichten zugeordnet werden kann – das System diese aber nicht vorsieht. Aus demselben Grund empfiehlt sich auch bei Kontrollkästchen die Option *Sonstiges*, einerseits weil mit großer Wahrscheinlichkeit irgendein unvorhergesehener Fall eintritt, und zweitens, um keine Entscheidung zu erzwingen, wenn User mit der Auswahl (z.B. der syntaktischen Klasse) unsicher sind. In Fällen, in denen noch keine Typologie möglicher Einträge erstellt wurde, kann zunächst ein Freitextfeld angelegt werden, das später durch ein Auswahlfeld ersetzt wird, wenn genügend Erfahrungswerte über gängige Einträge vorliegen. Dies wäre z.B. denkbar bei den *typischen Textsorten*, von denen es eine unüberschaubar große Anzahl gibt.

Ob ein phraseologisches Wörterbuch von einem kleinen Redaktionsteam erstellt wird oder als breit angelegtes Wiki-Projekt von der Öffentlichkeit bearbeitet werden kann: Es ist jedenfalls davon auszugehen, dass mehrere Personen an der Dateneingabe beteiligt sind. Daraus ergibt sich die Frage, ob **ein Feld mehrfach angelegt** werden kann oder nicht, d.h., ob bei der Bearbeitung eines Feldes der vorherige Eintrag überschrieben wird, oder ob das betreffende Feld durch eine Neueingabe dupliziert wird, so dass beide Einträge für die Abfrage noch zur Verfügung stehen.

Diese Entscheidung hängt von der Art der Daten ab: Bei der Definition ist die Eingang auf einen Konsens von Vorteil, da mehrere unterschiedliche Definitionen nicht unbedingt zu mehr Eindeutigkeit verhelfen. In anderen Fällen, z.B. bei Schlagwort, externen Leerstellen oder Kontextbeispielen, dienen mehrere parallele Eingaben der Präzisierung.

Bezüglich der Parallelität gibt es noch eine dritte Möglichkeit, die besonders bei einem öffentlichen Kollaborationsprojekt interessant wäre: Bei den Feldern *Areale Verbreitung* und *Stilschicht* wäre es besonders elegant, eine **statistische Auswertung** der Angaben einzubauen, die z.B. anzeigt, in welchen Regionen eines Sprachraums ein Ausdruck wie bekannt ist – eine nützliche Entscheidungshilfe für eine zielgruppen-spezifische Übersetzung. Diese Möglichkeit wird in 5.5.3.1 näher beschrieben.

Tabelle 4 zeigt eine Übersicht über alle Datenkategorien mit Vorschlägen zu Datentyp, Parallelität und möglichen Einträgen. Die Spalte *Datentyp* gibt Aufschluss darüber, ob Informationen als Text, Einfach- oder Mehrfachauswahl eingegeben werden. Die Spalte *Mögliche Einträge* bietet eine Übersicht über die Auswahloptionen bzw. über möglichen Text, wobei die meisten Angaben exemplarisch zu verstehen sind und keine vollständige Auflistung aller Optionen enthalten. Aus der rechten Spalte ist ersichtlich, ob parallele Felder angelegt werden sollten oder nicht.

| | Gebiet | Kategorie | Eingabe-modus | mögliche Einträge | Parallele Einträge? |
|----------------------|--------------------------------------|---|-------------------|---|---------------------|
| Einsprachige Angaben | Allgemein | Schlagwort | Freitext | semantische Schlagworte und ggf. stark abweichende Flexionsformen | ja |
| | | Phraseologismus (Nennform) | Freitext | | nein |
| | | Definition | Freitext | | nein |
| | | Synonym/ Paraphrase | Freitext | | ja |
| | | Varianten | Freitext/ Verweis | | ja |
| | | verwandte Phraseologismen | Verweis | Verweis auf Antonyme, Phraseologismen mit ähnlicher Metaphorik etc. | ja |
| | Semantik | Graphik | Multimedia | | ja |
| | | Grad der Idiomatizität | einfach | Kollokation, Idiom, Teil-Idiom | nein |
| | | Stilschicht | mehrfach | z.B. formell, literarisch, Slang | Statistik |
| | | regionale Verbreitung | mehrfach | Auflistung der Bundesländer/ Kantone von D, A und CH, evtl Regionen deutscher Minderheiten | Statistik |
| Sprachliche Angaben | Pragmatik | Fachgebiet | einfach | (eigene Untersystematik) | ja |
| | | Kontextbeispiele | Freitext | | ja |
| | | Quelle Kontext | Freitext | genaue Quellenangaben zu Buch, Zeitschrift, Website, Film, TV etc. | nein |
| | | typische Textsorten | Freitext | z.B. wissenschaftl. Text | ja |
| | | Bemerkungen zum Gebrauch | Freitext | Zusatzangaben zum Gebrauchs-kontext | nein |
| | Syntax | syntakt. Klasse | einfach | Satz, satzhähnliches Gebilde, Satz mit Leerstelle, Nebensatz, NP, VP, Adj.P, Adv.P, Prädikativ | nein |
| | | Kontextanbindung | einfach | keine, durch Satzglied, logisch, deiktisch | nein |
| | | Externe Leerstellen (oblig./ fakultativ) | mehrfach | z.B. Akkussativobjekt, Dativobjekt, Ergänzung | nein |
| | | Restriktionen | mehrfach | z.B. <menschlich>, <abstrakt> | nein |
| | | Transformationelle Defekte | mehrfach | Passiv- / Frage- / Relativsatztransformation, Imperativsatz, Nominalisierung, Satz- / Konstituentennegation, Expansion durch Attribute, Quantifizierung | nein |
| | Sonstiges | Traditionelle Klasse | mehrfach | Sprichwort, Gemeinplatz, Geflügeltes Wort, Redewendung, Zwillingssformel, komparativer Phraseologismus etc. | nein |
| | | Besonderheiten | mehrfach | unikale Komponente, vorangestellter Genitiv, untypische Valenz etc. | nein |
| | | Ursprüngl. Quelle | Freitext | Verweis auf literarisches Werk Politikerrede, Film o.ä. | nein |
| Übersetzung | Daten in ZS | Paraphrase/ Syn. ZS | Freitext | | nein |
| | | Definition ZS | Freitext | | nein |
| | | verwandte Phraseologismen ZS | Verweis | | ja |
| | | Äquivalenz | einfach | total, partiell, keine | nein |
| | Übersetzungsbsp | Freitext | | | ja |
| | Kommentar zur Äquivalenz (L1) | Freitext | | z.B. "nur verwendbar in der und der Situation..." | ja |
| | Kommentar zur Äquivalenz (L2) | Freitext | | z.B. "nur verwendbar in der und der Situation..." | ja |

Tabelle 4: Datenkategorien mit Eingabemodus

5.5.3.1 Statistische Datenauswertung

Bei einer großen Teilnehmerzahl ist in manchen Fällen eine statistische Auswertung von User-Angaben eine verlässlichere Informationsquelle als Expertenintrospektion (vgl. Dräger/Juska-Bacher 2010:175). Dies betrifft folgende Daten:

- (1) Gebräuchlichkeit des Phraseologismus und seiner Varianten
- (2) Beurteilung stilistischer Faktoren
- (3) Areale Verbreitung
- (4) Bekanntheit je Altersklasse

Als erster Schritt empfiehlt sich das Anlegen eines Benutzerprofils, in dem Teilnehmer einmalig Angaben zu Region und Alter machen, damit diese Daten nicht für jeden zu bewertenden Phraseologismus neu gemacht werden müssen.

Bezüglich der **Region** ist eine Grobeinteilung des jeweiligen Sprachraums sinnvoll, wie sie in 5.4.4 besprochen wurde. Für linguistische Zwecke kann bei Bedarf eine feinere Untergliederung vorgenommen werden. Userinnen werden gebeten, all jene Regionen anzugeben, durch die sie nach eigener Einschätzung sprachlich geprägt wurden (z.B. Wohnsitz für länger als 3 Jahre).

Auch die Angabe der **Altersklasse** kann in funktionaler Hinsicht aufschlussreich sein: Wenn ein Phraseologismus vorwiegend in älteren Bevölkerungsschichten bekannt ist, ist er möglicherweise veraltet; vorwiegender Gebrauch in jüngeren Altersklassen kann ein Hinweis auf – mitunter kurzlebige – Jugendsprache sein.

Zur Ermittlung der **Gebräuchlichkeit** eines Phraseologismus (bzw. einer Variante desselben) kann über folgende Fragekonstellation ermittelt werden:

Wie bekannt ist Ihnen der Phraseologismus?

- Aktiv:** Er ist mir bekannt und ich verwende ihn in einem passenden Kontext aktiv.
- Passiv:** Ich kenne und verstehe ihn zwar, würde ihn aber nicht verwenden.
- Variiert:** Ich kenne/verwende ihn in einer etwas anderen Form, nämlich _____
- Vage:** Ich habe den Ausdruck schon gehört, weiß aber nicht genau, was er bedeutet.
- Unbekannt:** Sagt mir überhaupt nichts.

Bei Auswahl einer der ersten drei Antworten folgen weitere Fragen zur **stilistischen und kontextuellen Bewertung** des Phraseologismus gemäß den in 5.4.1.4 beschriebenen Merkmalen, bei Angabe der Optionen *vage* oder *unbekannt* jedoch nicht, um eine Verfälschung der Statistik durch bloße Vermutungen zu vermeiden.

Auf Basis dieser Angaben lässt sich z.B. eine interaktive Landkarte erstellen, auf der Regionen je nach Prozentsatz der aktiven Verbreitung unterschiedlich schattiert

sind. Bei Anklicken einer Region erscheinen differenziertere Diagramme, die z.B. die prozentuale Verteilung der Angaben zum Bekanntheitsgrad oder die stilistische Bewertung in der jeweiligen Region anzeigen. Auch weitere Kombinationen von Angaben sind interessant, z.B. die prozentuale Bevorzugung bestimmter Varianten etc.

Die Anzeige solcher statistischer Ergebnisse bei der Datenabfrage ist jedoch erst ab einem gewissen Schwellenwert von Teilnehmerinnen sinnvoll. Eine nicht repräsentative Verteilung der Altersklassen und Herkunftsgebiete unter den Befragten muss ggf. durch Methoden der Gewichtung ausgeglichen werden.

5.5.4 Datenabfrage

Im Folgenden wird kurz skizziert werden, wie eine Datenabfrage für Userinnen aussehen kann. Hierbei soll im Sinne von Lubensky/McShane's Idee verfahren werden, Information stufenweise anzuseigen:

„...one could design an on-line bilingual phraseological dictionary using methods of 'progressive disclosure', such that a basic entry were provided first with links to more information of specific kinds for interested users“ (Lubensky/McShane 2007:927).

Nach Auswahl der gewünschten Sprachen wird man zunächst aufgefordert, einen Suchbegriff einzugeben. Daraufhin erscheint eine Reihe von Phraseologismen, die diesen Suchbegriff enthalten. Man hat nun die Möglichkeit, durch Anklicken eines Phraseologismus mehr einsprachige Details anzeigen zu lassen, oder man begibt sich über einen weiteren Link direkt zu Übersetzungsvorschlägen für den gewählten Phraseologismus. Dort scheint eine Reihe von Zielsprachlichen Phraseologismen auf, sortiert nach Äquivalenzgrad. Klickt man einen Begriff an, werden sowohl übersetzungsbzogene Daten angezeigt, als auch eine Selektion einsprachiger Daten zu beiden Phraseologismen in Gegenüberstellung.

Wurde zunächst die Detailansicht zum ausgangssprachlichen Phraseologismus gewählt, so besteht von hier aus die Möglichkeit, den Begriff direkt übersetzen zu lassen, ohne noch einmal zurück zur Übersicht zu gehen.

Als weiteres Feature kann man auch nach Phraseologismen suchen, die eine dem eingegebenen Suchbegriff ähnliche Bedeutung haben. Wie in 5.1.4.5 erwähnt, kann diese Suche auf einem integrierten semantischen Netz, semantischer Beschlagwörter oder auf einer Suche im Feld *Synonym/Paraphrase* basieren. Bei Eingabe des Begriffes *sterben* können so z.B. Phraseologismen wie *ins Gras beißen*, *den Löffel abgeben* und *Jds letztes Stündlein hat geschlagen* aufscheinen.

5.5.5 Usability

Unabhängig davon, ob die Datenbank von einem überschaubaren Expertinnenteam oder von einem breiteren Publikum bearbeitet wird – eine **benutzerfreundliche Oberfläche** ist unabdingbare Voraussetzung für die Datenqualität und somit für den Erfolg des Projekts. Nur über eine klar strukturierte und verständliche Benutzeroberfläche kann sichergestellt werden, dass in jedem Feld die jeweils gefragten Daten (und nur diese) eingegeben werden.

Usability ist dabei keineswegs nur bei einem öffentlich zugänglichen Kollaborationsprojekt, sondern auch bei ausschließlicher Bearbeitung durch Experten relevant. Dies hat verschiedene Gründe: Erstens kann das erforderliche IT-Verständnis bzw. die Motivation zur Einarbeitung nicht pauschal vorausgesetzt werden. Zweitens erschließen sich viele phraseologischen Datenkategorien nicht intuitiv, sondern erst durch eingehende Beschäftigung mit der Materie (als Beispiel sei hier an die Berücksichtigung externer Leerstellen bei der syntaktischen Klassenzuordnung erinnert, s. 3.1.3.3). Erläuterungen, Online-Tutorials und sonstige Hilfen verhelfen zu Eindeutigkeit und können eine ansonsten erforderliche Einarbeitung ersetzen. Somit wirkt sich eine verständliche Oberfläche nicht nur positiv auf die Datenqualität, sondern auch auf die Motivation zur Teilnahme aus.

Wie aber kann dieses Ziel erreicht werden? Neben der Berücksichtigung gängiger Usability-Kriterien (z.B. Nielsen/Mack 1994) werden insbesondere zwei Aspekte vorgeschlagen:

Aus den genannten Gründen empfehlen sich **interaktive Hilfen und Tutorials**, um die Natur des gefragten Feldinhalts und weitere Prinzipien bei der Eingabe verständlich zu machen. Hierbei ist es wichtig, neben einem Tutorial zum Einstieg vor allem **begleitend zur Eingabe** Hilfestellungen einzuplanen, da ein Teil der User vermutlich den Aufwand scheut, einen separat zugänglichen theoretischen Text vorher durchzulesen, bzw. bei Auftreten von Zweifeln dorthin zurück zu navigieren. Besser ist es, wenn jeweils neben einem Eingabefeld ein kleiner Erläuterungstext mit Informationen und Beispielen zum gefragten Inhalt erscheint. Eine ausführlichere Version dieser Erläuterung kann auch erst bei Anklicken oder Überschweben mit der Maus (*Hovern*) angezeigt werden.

Ein weiterer wichtiger Faktor in puncto Usability ist die **selektive Anzeige von Information**. Userinnen sollen nicht auf Anhieb mit der gesamten Komplexität phraseologischer Daten überwältigt werden. Besser ist es, bei der Dateneingabe standardmäßig nur eine reduzierte Auswahl an Datenkategorien anzuzeigen, und spezifischere Felder erst durch Anklicken eines minimierten Bereichs sichtbar zu machen. Spezifischere Datenfelder können auch von vornherein nur fortgeschrittenen Usern zugänglich gemacht werden. (vgl. 5.5.2).

Auch für die **Datenabfrage** ist eine etappenweise Anzeige der Information hilfreich: Neben einem reduzierten Standard-Anzeigeprofil können weitere vordefinierte **Profile für bestimmte Interessensgruppen** (Linguisten, Translatoren, Sprachdidaktiken etc.) zur Auswahl stehen. Zusätzlich soll die Möglichkeit bestehen, ein individuelles Anzeigeprofil mit gewünschten Datenkategorien anzulegen.

5.6 Resümee: Kriterienkatalog für ein elektronisches phraseologisches Wörterbuch

In diesem Kapitel wurden grundlegende Eigenschaften und eine mögliche Datenbankstruktur eines multilingualen allgemeinen phraseologischen Wörterbuchs erarbeitet, zudem wurden Möglichkeiten der Dateneingabe und -bearbeitung besprochen. Auf Basis des Beschriebenen lässt sich zusammenfassend eine Liste von Anforderungen zusammenstellen, die ein solches phraseographisches Projekt im Idealfall erfüllen sollte:

1. eine differenzierte Datenbankstruktur, in der jede Informationseinheit nach den Prinzipien der Granularität und Elementarität separat gespeichert wird;
2. einheitliche Repräsentation der Daten;
3. Ausbaufähigkeit der Datenbank;
4. Trennung einsprachiger und übersetzungsbezogener Daten;
5. Möglichkeit der Zuordnung mehrerer zielsprachlicher Äquivalente zu einem Eintrag, darunter auch partielle Äquivalente;
6. Möglichkeit der Beschlagwortung;
7. Möglichkeit der Verlinkung verwandter Phraseologismen;
8. umfassende funktionale Angaben, die eine aktive Verwendung von Phraseologismen ermöglichen;
9. Integration eines Statistik-Tools für interaktive User-Befragungen;
10. Möglichkeit der Web-Kollaboration bei der Eingabe, um aktuellen und breit gefächerten Datenbestand zu gewährleisten;
11. gestaffelte Vergabe von Bearbeitungsrechten; Bewertungsmechanismen;
12. Erstellung interessensbezogener und individueller Abfrageprofile;
13. allgemeine Benutzerfreundlichkeit durch Online-Tutorials, interaktive Hilfsfunktionen und Staffelung der Information;
14. insbesondere Direkthilfe bei der Eingabe (ohne zurücknavigieren zu müssen)

Es ist nicht davon auszugehen, dass eine bestehende Software den genannten Kriterien in vollem Umfang entspricht. Die Neuprogrammierung eines mehrsprachigen phraseologischen Wörterbuchs scheint daher ein lohnenswertes Projekt. Der Projektumfang muss dabei die Aufstellung eines multidisziplinären Mitarbeiterteams ermöglichen, das Fachwissen aus verschiedenen sprachbezogenen und informationstechnischen Disziplinen in sich vereint. Gewinnbringend wäre zudem eine enge Zusammenarbeit mit Herstellern sprachverarbeitender Systeme in der Frühphase, um die Austauschbarkeit von Daten durch standardisierte Formate und lexikalische Repräsentationen zu erzielen.

In bestehenden Systemen kann obiger Kriterienkatalog als Evaluationsbasis hinsichtlich des Umgangs mit Phraseologismen dienen. Das folgende Kapitel beschäftigt sich mit phraseographischen Möglichkeiten der Terminologieverwaltungssoftware *MultiTerm*, wobei die beschriebenen Kriterien als Leitlinie dienen.

6 Phraseographie am Beispiel von *SDL MultiTerm*

In diesem Kapitel sollen Möglichkeiten untersucht werden, Phraseologie mit der Terminologieverwaltungssoftware *SDL MultiTerm 2007* (im Folgenden „*MultiTerm*“) zu verwalten, wobei das in Kapitel 5 entworfene Datenmodell als Leitlinie dient. *MultiTerm* erlaubt eine flexible Strukturierung von Termbanken, ist jedoch nicht von Haus aus auf phraseographischen Gebrauch zugeschnitten. Es werden daher drei verschiedene Eintragsmodelle entworfen, um zu überprüfen, auf welche Weise die in Kapitel 5 erarbeitete Struktur am ehesten umgesetzt werden kann. Hierfür wurde ein kleines Beispielkorpus verwendet, das in 6.1 kurz vorgestellt wird. In 6.2 wird die allgemeine Funktionsweise von *MultiTerm* erläutert, 6.3 befasst sich mit Möglichkeiten der phraseographischen Nutzung.

6.1 Verwendetes Korpus

Zur praktischen Umsetzung der in diesem Kapitel vorgestellten Datenmodelle wurde ein kleines Beispielkorpus von 25 ungarischen Phraseologismen verwendet, die zum Teil über mehrere deutsche Äquivalente verfügen. Als Ausgangsbasis wurde die Sprachkombination Ungarisch-Deutsch verwendet, zum Testen der Mehrsprachigkeit wurden einzelne englische und portugiesische Beispiele hinzugezogen.

Die Beispiele entstammen keinem bestimmten Textkorpus, sondern wurden über Jahre hinweg aus Texten, Übersetzungen und mündlichem Sprachgebrauch gesammelt und aufgezeichnet. Zum Zweck dieser Arbeit wurden die Beispiele anhand von Quellen und Lexika belegt und von ungarischen Muttersprachlerinnen mit ausgezeichneten Deutschkenntnissen überprüft.

Eine ausführliche Liste des ungarisch-deutschen Korpus mit Definitionen, Kontextbeispielen und Kommentaren zur Äquivalenz befindet sich in Anhang A. Bei der Zusammenstellung des Korpus dienten größtenteils ungarische Phraseologismen als Grundlage, für die dann deutsche Äquivalente gefunden wurden. Daher wurde die Übersicht in Anhang A auch in dieser Übersetzungsrichtung angelegt. In den Datenmodellen kommen jedoch beide Übersetzungsrichtungen, jeweils gestützt durch Übersetzungsbeispiele, vor. Bei manchen Beispielen war die Phraseologizität des deutschen Äquivalents ausschlaggebend für die Aufnahme. So handelt es sich bei *szabadlábon* nur um ein Wort, die deutsche Entsprechung *auf freiem Fuß* ist hingegen phraseologisch.

Varianten werden in der Übersicht aus Platzgründen in einem Feld angeführt und durch Schrägstriche bzw. Klammern für optionale Zusätze kenntlich gemacht. In der Datenbank erhält jede Variante ein eigenes Feld. Dasselbe gilt für Quellenangaben, die in der Übersicht in Klammern im selben Feld verzeichnet wurden.

Bei der Auswahl wurde darauf geachtet, dass verschiedene phraseologische Klassen ausreichend vertreten sind, um durch eine hohe Diversität der Daten die Modellierung möglichst allgemeingültig zu gestalten. Folgende phraseologische Klassen sollten mehrfach im Korpus vertreten sein:

| | |
|--------------------|---|
| Äquivalenz: | total partiell pragmatisch |
| Idiomatizität: | Kollokation Teil-idiom Idiom |
| Stilschicht: | Allgemeinsprache Fachsprache gehoben journalistisch umgangssprachlich |
| Satzgliedfunktion: | Satz oder satzwertiges Gebilde Satz mit Leerstelle Nominalphrase Verbalphrase Adverbialphrase |

Tabelle 5 zeigt eine Übersicht über das Korpus mit quantitativer Verteilung der relevanten Kriterien. Die deutschen Äquivalente werden ebenfalls genannt, jedoch beziehen sich alle Angaben außer der Äquivalenz nur auf den ungarischen Ausgangsphraseologismus. Phraseologismen mit mehreren deutschen Äquivalenten wurden mehrfach eingetragen; ihre Eigenschaften jedoch nur einmal gewertet; die aufgrund von Duplikaten nicht gewerteten Felder sind schraffiert. Der Äquivalenzgrad wurde für jede Äquivalenzbeziehung einmal gewertet.

Alle drei in diesem Kapitel vorgestellten Datenmodelle wurden anhand der beschriebenen Daten getestet. Der Inhalt der Datenbanken wird auszugsweise zur Veranschaulichung der Modelle durch Graphiken dargestellt.

Tabelle 5 (umseitig): Übersicht über das Korpus mit quantitativer Verteilung der relevanten Kriterien

| Phraseologismus HU | Phraseologismus DE | Idiomatizität HU | Stilskichtigkeit HU | syntakt. Klasse HU | Aquivalenz HU-DE |
|--|--|----------------------|-----------------------|---------------------------------------|--|
| | | Koll. Teil- idiom | Idiom Alg. Spr. | Satz, satz- wertig, Satz mit LS | NP VP Prädikativ Adv. P. total pragmatisch |
| árvít ételemben | im übertragenen Sinn / im übertragenen Sinne | x | x | ugs. literarisch journalistisch | x x |
| szabadlábon | auf freiem Fuß | x | x | x x | x x |
| [vnjl] színe-java | die Elite [Gen Attr.] | x | x | x | x |
| [vnjl] színe-java | die Crème-de-la-crème [Gen Attr.] | | | | |
| [vnjl] színe-java | das Beste von etw/ [Gen Attr.] | | | | |
| [vnjl] színe-java | (die) Spitzen- [N] [Gen Attr.] | | | | |
| [vnjl] színe-java | die erlesenen [NP] | | | | |
| Sok hihő semmiért | Viel Lärm um nichts | x | x | x | x |
| visszaér a régi kerékvágásba | wieder seinen gewohnten Lauf nehmen | x | x | x x | x x |
| visszaér a régi kerékvágásba | wieder ins alte Fahrwasser geraten | | | | |
| nem látja a fától az erdőt | vor lauter Bäumen den Wald nicht mehr sehen | | | | |
| [vkinek] nyomordokain lépdel | in [ids] Fußstapfen treten | | | | |
| felkapja a vizet | sich aufreigen | x | x | x | x |
| hegyesszög | ein spitzer Winkel | x | x | x | x |
| csödői jelent | Konkurs anmelden | x | x | x | x |
| Az érték műltbeli alakulása nem jelent meghibásodást, kováriketést, a befektetési alap értékek jóvöröbeli alakulására nézve. | Die Wertentwicklung der Vergangenheit lässt keine verlässlichen Rückschlüsse auf die zukünftige Wertentwicklung eines Kapitalanlagefonds zu. | | | | |
| gazdáig fellendülés | wirtschaftlicher Aufschwung | x | x | x | x |
| a söjtőképző | das finstere Mittelalter | x | x | x | x |
| [vnjl] felpezsdti a kedélyeket | für gute Laune sorgen | x | x | x | x |
| felkorhácsolja a kedélyeket | die Stimmung einheizen | | | | |
| felkorhácsolja a kedélyeket | die Gemüter erhitzten | x | x | x | x |

| Phraseologismus HU | Phraseologismus DE | Syntaktik, Klasse HU | | | | Äquivalenz HU-DE |
|---------------------------------------|--|----------------------|-----------------|----|----|------------------|
| | | Idiomatisität HU | Stilrichtung HU | | | |
| Sok hába között elvész a gyerek. | Viele Köche verderben den Brei. | x | x | | x | x |
| nyáron | im Sommer | x | x | | | x |
| beáll az alkony | die Dämmerung bricht herein | x | x | x | | x |
| hatályba lép | in Kraft treten | x | x | x | | x |
| annak idején | zu jener Zeit, danach | x | x | | | x |
| annak idején | dann (Zukunft) | | | | | |
| á-tól cettig | von A bis Z | | | | | |
| szeplőteleken fogantás | unbefleckte Empfängnis | x | x | x | x | x |
| hitelbe | auf Kredit | x | x | | | x |
| eltünik, mint s zurke szamár a ködben | spurlos verschwinden | x | x | | | x |
| eltünik, mint s zurke szamár a ködben | (völlig) von der Bildfläche verschwinden | | | | | |
| eltünik, mint s zurke szamár a ködben | wie vom Erdboden verschluckt sein | | | | | x |
| fizetést teljesít | eine Zahlung leisten | x | x | x | | x |
| [vminek] fele se tréfa | mit [etw] ist nicht zu spaßen | x | x | x | | x |
| [vminek] fele se tréfa | [etw] ist eine ernste Angelegenheit/ ein ernstes Problem | | | | | x |
| [Vm] bázlik itt/ [Vm]iben | Etwas ist faul an [etw]/ Hier/ Darauf ist etwas faul. | x | | x | | x |
| | | 12 | 7 | 8 | 18 | 7 |
| | | | 2 | 7 | 6 | 5 |
| | | | 5 | 10 | 1 | 5 |
| | | | 5 | 13 | 5 | 18 |

6.2 Übliche Verwendung von MultiTerm

6.2.1 Allgemeines

SDL MultiTerm 2007 ist eine Terminologieverwaltungssoftware, die eine flexible Termbankdefinition ermöglicht. Die Software ist in zwei Ausführungen erhältlich, nämlich in einer Einzelplatzversion für die Bearbeitung durch nur eine Benutzerin zur selben Zeit (*MultiTerm Desktop*³¹) und einer serverbasierten Version für die gleichzeitige Bearbeitung durch mehrere Personen (*MultiTerm Server*). Die Benutzeroberfläche und verfügbare Optionen unterscheiden sich in beiden Versionen nicht voneinander. Die vorliegenden phraseographischen Untersuchungen wurden mit der lokalen Version der Software durchgeführt, die Ergebnisse sind jedoch eins zu eins auf die serverbasierte Version übertragbar.

MultiTerm Server lässt sich zusätzlich mit der Webserveranwendung *MultiTerm Online* kombinieren, über die Termbanken im Internet veröffentlicht und im Webbrowser eingesehen und bearbeitet werden können, ohne dass Anwender *MultiTerm* lokal installieren müssen. Diese Möglichkeit ist für ein kollaboratives phraseographisches Projekt von Interesse.

Mittlerweile existiert eine neue Version von *MultiTerm*, die für die Durchführung dieser Arbeit nicht zur Verfügung stand. Sie beinhaltet insbesondere Verbesserungen der Benutzerfreundlichkeit beim Anlegen von Termbanken; in Bezug auf die grundlegende Struktur und hier diskutierte phraseographierelevante Aspekte hat sich jedoch nichts Grundlegendes verändert.

6.2.2 Eintragsstruktur

MultiTerm ermöglicht die Erstellung begriffsorientierter Termbanken, in denen sämtliche Benennungen für einen Begriff in verschiedenen Sprachen – also Synonyme und Äquivalente – in einem Eintrag zusammengefasst werden, Homonyme dagegen in getrennten Einträgen.³² Ein Eintrag wird dabei durch drei Ebenen strukturiert, wie Abb. 8 zeigt:

³¹ Früher *MultiTerm Workstation*

³² Durch seine hohe Flexibilität erlaubt *MultiTerm* auch die Erstellung benennungsorientierter Termbank, jedoch ist dies nicht die klassische Verwendung von *MultiTerm*.

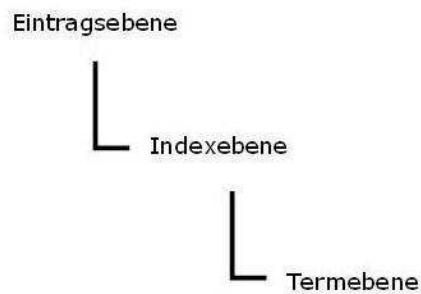


Abb. 8: Eintragsstruktur bei *MultiTerm*

Auf der *Eintragsebene* liegen sämtliche Angaben, die sich auf den gesamten Begriff unabhängig von dessen sprachlicher Realisierung beziehen, also z.B. Fachgebiet oder eine sprachunabhängige Graphik. Hier sind auch die verschiedenen Sprachen (als so genannte *Indexfelder*) angesiedelt, unter denen die nächst tiefere Ebene, die *Indexebene* hängt. Auf der Indexebene befinden sich sprachbezogene Daten, so zum Beispiel eine Definition in der jeweiligen Sprache, einsprachig beschriftete Graphiken und verschiedene synonyme Benennungen, die *Terme*. Unterhalb der Terme, auf der *Termebene*, sind dann Angaben zu den jeweiligen Benennungen verzeichnet, zum Beispiel grammatische Angaben, konnotative Bedeutung, Zugehörigkeit des Terms zu einer bestimmten Firmenterminologie etc.

Abb. 9 zeigt ein Beispiel eines klassischen Eintrags. Dieser wurde einer Termbank über EU-Institutionen entnommen, die im Rahmen von Lehrveranstaltungen am Zentrum für Translationswissenschaft der Universität Wien entstanden ist³³.

³³ <http://mtsrv.trans.univie.ac.at:8090/TermbaseFinderApp/>, Termbank *EU-Inst*, geleitet von Irmgard Soukup-Unterweger und Klaus Fleischmann.

Entry level

Entry number: 28

Fachgebiet: Politik

Unterfachgebiet: EU

Illustration:

Quelle: <http://images.google.at/imgres?imgurl=http://www.spiegel.de/img/0,1020,887242,00.jpg&imgrefurl=http://www.spiegel.de/w=180&sz=9&hl=de&start=68&tbin=1!mx3!VcgbthM:&tbnh=101&tbnw=101&prev=/images%3Fq%3Ddie%2BEurop%25C3%25A4sche%2BZentralbank%26start%3D54%26gbv%3D2%26ndsp%3D18%26a%26rls%3Dorg.mozilla:en-US:official%26sa%3DN> [18.01.2008]

English

Definition: [...] the Central Bank is responsible for managing the euro — for example, by setting interest rates. Its prime concern is ensuring price stability so that the European economy is not damaged by inflation. The Bank takes its decisions independently of governments and other bodies.

Quelle: http://europa.eu/abc/panorama/howorganised/index_en.htm#central [25.01.2008]

Term: European Central Bank

Quelle: http://europa.eu/abc/panorama/howorganised/index_en.htm#central [25.01.2008]

Workflow-Status: freigegeben

Benennungstyp: Hauptbenennung

Sprachebene: Fachsprache

Grammatik: sg.t

Kontext: Since 1 January 1999 the European Central Bank (ECB) has been responsible for conducting monetary policy for the euro area - the world's largest economy after the United States.

Quelle: <http://www.ecb.int/ecb/orga/escb/html/index.en.html> [25.01.2008]

Term: ECB

Quelle: <http://www.ecb.int/ecb/html/index.en.html> [25.01.2008]

Workflow-Status: freigegeben

Benennungstyp: Kurzform, Synonym

Sprachebene: Fachsprache

Grammatik: sg.t

Kontext: In order to best serve its objective of maintaining price stability, the ECB, like any other central bank, needs to thoroughly analyse economic developments.

Quelle: <http://www.ecb.int/mopo/strategy/html/index.en.html> [25.01.2008]

German

Term: Europäische Zentralbank

Quelle: http://europa.eu/abc/panorama/howorganised/index_de.htm#central [11.01.2008]

Benennungstyp: Hauptbenennung

Workflow-Status: freigegeben

Sprachebene: Fachsprache

Grammatik: f., sg.t

Kontext: Die Europäische Zentralbank, deren Präsident Jean-Claude Trichet ist, hat ihren Sitz in Frankfurt am Main.

Quelle: http://europa.eu/abc/panorama/howorganised/index_de.htm#central [11.01.2008]

Term: EZB

Quelle: <http://www.ecb.eu/home/html/lingua.de.html> [11.01.2008]

Benennungstyp: Kurzform, Synonym

Sprachebene: Fachsprache

Grammatik: f., sg.t

Workflow-Status: freigegeben

Kontext: Trichet sah in mehr Transparenz des Marktes eine Möglichkeit, Vertrauen wieder herzustellen. „Die EZB selbst ist ein Vorbild an Transparenz. Wir sind die einzige Zentralbank, die in Echtzeit erklärt, was sie tut und warum.“

Quelle: http://www.europarl.europa.eu/news/public/story_page_042-10173-254-09-37-907-20070906STO10164-2007-11-09-2007/default_de.htm [11.01.2008]

Abb. 9: Klassischer Eintrag in MultiTerm

Die drei genannten Ebenen sind hier graphisch durch Einzüge kenntlich gemacht. Auf der Eintragsebene wurden Eintragsnummer, Fachgebiet und der Link zu einer Graphik vermerkt, ebenfalls auf dieser Ebene liegen die Indices, d.h. die Sprachen. Darunter, auf der Indexebene wurde eine Definition in der jeweiligen Sprache eingetragen (die auf alle Terme zutrifft), auf dieser Ebene liegen auch verschiedene synonyme Termini (hier: Vollform und Abkürzung), denen jeweils termbezogene Angaben wie Sprachebene, Grammatik, Kontext etc. untergeordnet sind. In *MultiTerm* lassen sich beliebig viele weitere Sub-Ebenen anhängen.

Das Feld *Term* ist automatisch in jedem Index vorhanden und ist ein Pflichtfeld; es muss also für jede angezeigte Sprache mindestens einmal ausgefüllt werden. Die an-

deren Felder lassen sich frei definieren und in beliebiger Hierarchie anordnen. Feldname, Eingabemodus und weitere Eigenschaften werden vom Datenbankersteller festgelegt. Dies ermöglicht eine flexible Anpassung der Datenbank an die jeweiligen Bedürfnisse, erfordert aber auch eine gründliche Planungsarbeit.

Eine weitere Möglichkeit von *MultiTerm* ist das Anlegen sogenannter *Cross-References*, also Querverweisen zu anderen Einträgen der Datenbank, etwa verwandten Begriffen, Antonymen etc. Diese Möglichkeit kann bei der phraseographischen Nutzung der Software hilfreich sein. Ebenso können Hyperlinks zu Internetquellen gelegt werden.

Eine begriffsorientierte Datenbank bietet den Vorteil einer semantischen Logik im Aufbau, die bei ausschließlich alphabetischer Gliederung fehlt, verlangt aber auch, dass verschiedene Benennungen eindeutigen Begriffen zugeordnet und scharf von einander abgegrenzt werden können. Dies ist schon im unigraphen Bereich nicht immer der Fall: es ist bekannt, dass im Sprachvergleich begriffliche Überlappungen und Inklusionen vorkommen (vgl. Arntz et al. ⁵2004:153ff.). Bei der Phraseologie rückt dieser Aspekt noch mehr in den Vordergrund, da, wie in 4.3 ausgeführt wurde, auch ein partielles Äquivalent eine passende Übersetzung darstellen kann. Auch spielen in der Phraseologie neben der denotativen Bedeutung die feinen Unterschiede zwischen verschiedenen Konnotationen, die bei einem begriffsorientierten Ansatz in einem Eintrag zusammengefasst werden, eine noch zentralere Rolle.

6.2.3 Anlegen der Datenbank

Zum Anlegen einer neuen Termbank werden der Reihe nach drei Ebenen erstellt:

- Termbankdefinition
- Eintragsmodell
- Layout

Zunächst wird in der **Termbankdefinition** (*Termbase Definition*) das eigentliche **Datenmodell** festgelegt, d.h. es werden sämtliche Felder, die später benötigt werden, angelegt und hierarchisch geordnet.

Durch ein **Eintragsmodell** (*Input Model*), wird festgelegt, welcher Ausschnitt aus der gesamten Datenbankstruktur für einen bestimmten Zweck zur Bearbeitung angezeigt werden soll. So können z.B. nur bestimmte Sprachen angezeigt werden, ein Feld für grammatische Angaben nur unter der Sprache, für die die Angabe relevant ist etc.

Im **Layout** schließlich können – neben graphischer Gestaltung – Benutzerprofile für bestimmte Zwecke angelegt werden, in denen nur die jeweils relevanten Felder angezeigt werden. Bezogen auf die Phraseologie könnten hier etwa für Linguistinnen,

Übersetzerinnen und Sprachlernende jeweils unterschiedliche Felder angezeigt werden, die für den jeweiligen Zweck von besonderem Interesse sind.

Die Festlegung der *Termbankdefinition* erfolgt in drei Schritten, wie auf den nachfolgenden Abbildungen ersichtlich ist: Zuerst werden die Sprachen bestimmt (Abb. 10); im zweiten Schritt werden sämtliche später benötigten Felder angelegt und deren Eigenschaften definiert, und zwar zunächst noch ohne jegliche hierarchische Gliederung (Abb. 11). Im dritten Schritt werden die angelegten Felder dann hierarchisch geordnet (Abb. 12).

Beim Anlegen der Felder (Abb. 11) können auch Feldeigenschaften, d.h., ob es sich um ein Textfeld, ein Optionsfeld oder sonstiges handelt, unter *Properties* festgelegt werden. Optionsfelder werden in *MultiTerm* in Form von Drop-Down-Menüs, die hier *Picklist* heißen, realisiert. Die Picklist kann sowohl für Felder mit Mehrfach- als auch Einfachnennung verwendet werden. Werden für eine Picklist Mehrfachnennungen zugelassen – diese Eigenschaft wird im dritten Schritt (Abb. 12) zugewiesen – können später bei der Dateneingabe über Betätigung der Steuerungstaste mehrere Elemente der Liste ausgewählt werden.

Hier wird deutlich, dass eine gründliche Vorausplanung der Datenbank notwendig ist – es muss zur Durchführung dieses Schrittes bereits eine genaue Übersicht darüber bestehen, welche Felder später benötigt werden und welche Eigenschaften diese besitzen sollen.

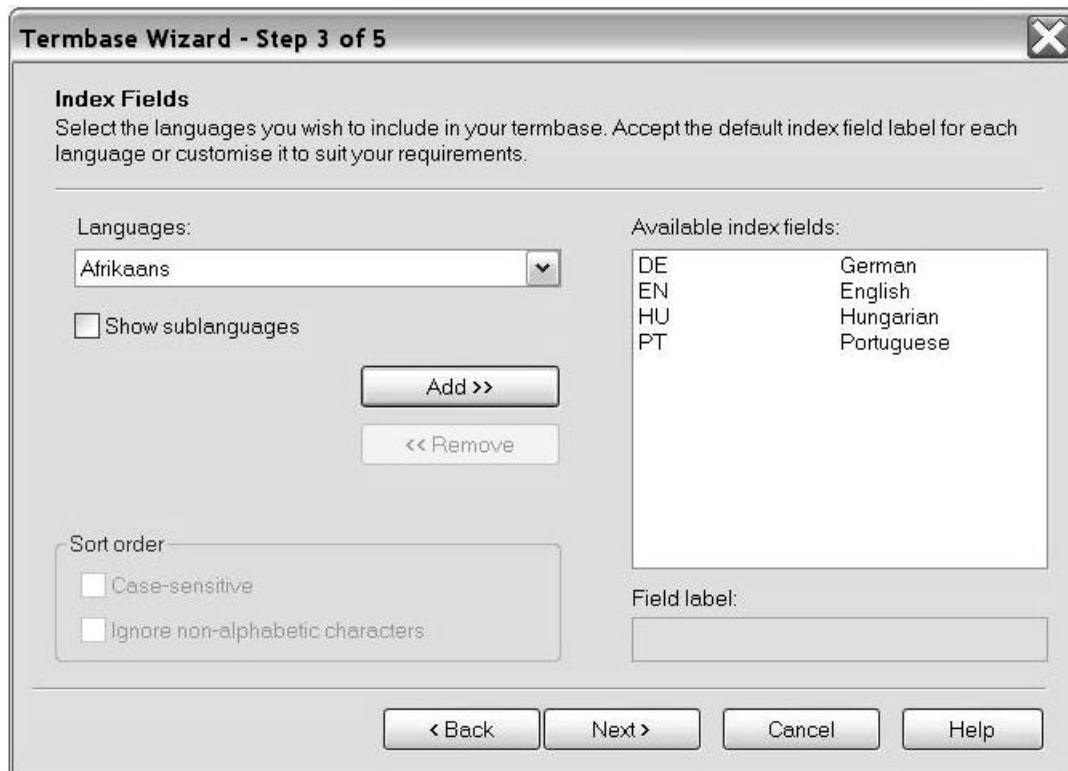


Abb. 10: Termbankdefinition – Indexfelder



Abb. 11: Termbankdefinition – beschreibende Felder

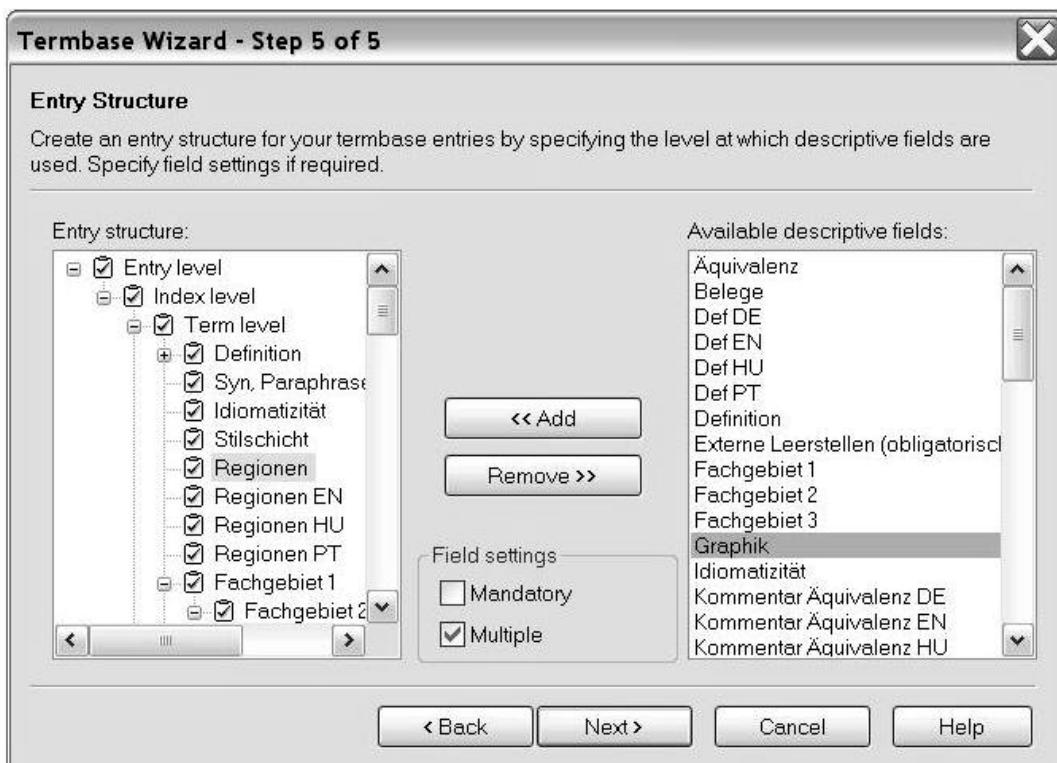


Abb. 12: Termbankdefinition – Eintragsstruktur

Im dritten Schritt werden die Felder hierarchisch angeordnet (Abb. 12), indem man Elemente aus der Liste rechts an die gewünschte Stelle im Strukturabaum links einfügt. Hierbei müssen sämtliche Felder auf die entsprechende Stufe der Hierarchie – Eintrags-, Index-, Termebene oder tiefer – gelegt werden. Sämtliche Felder einer Ebene stehen hier zunächst nebeneinander, unabhängig davon, in welchem „Ast“ der Termbank sie später benötigt werden. Beispielsweise wurde für jede Sprache ein eigenes Optionsfeld *Regionen* (areale Verbreitung) kreiert, in dem verschiedene Regionen des jeweiligen Sprachraums zur Auswahl stehen. Da die Hierarchie innerhalb der Termbankdefinition noch nicht nach Indices (Sprachen) gegliedert ist, liegen hier sämtliche *Regionen*-Felder nebeneinander.

Wichtig ist bei diesem Schritt, *sämtliche* Datenfelder ihrem Platz in der Hierarchie zuzuordnen, wenn nötig auch an mehreren Stellen, damit sie später dort angezeigt werden können. Einzelne Felder können später im Eingabemodell an bestimmten Stellen zwar weggelassen, jedoch nicht hinzugefügt werden.

Der nächste Schritt beim Erstellen der Termbank ist das Anlegen eines **Eingabemodells**. Hier kann festgelegt werden, welcher Ausschnitt aus der gesamten Datenbankstruktur für einen bestimmten Zweck angezeigt werden soll – so können z.B. nicht benötigte Sprachen für die Bearbeitung ausgeblendet werden oder nicht anwendbare Felder – z.B. für eine Sprache irrelevante Regionen oder grammatische Angaben – weggelassen werden. Abb. 13 zeigt in vereinfachter Form, wie die Hierarchie eines Eingabemodells aussehen könnte.

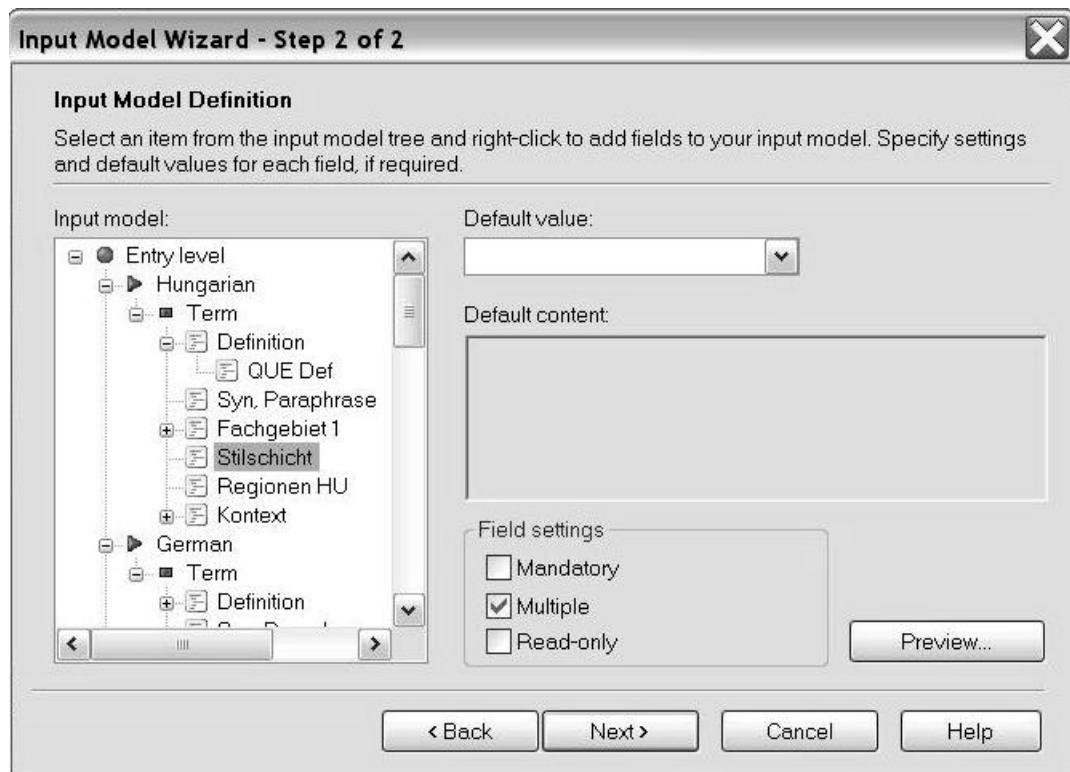


Abb. 13: Eintragsmodell – hierarchische Struktur

Im Gegensatz zur Termbankdefinition werden hier also nicht mehr alle Felder der Indexebene nebeneinander angezeigt, sondern die gewünschte Auswahl wird bereits der jeweiligen Sprache untergeordnet. Auch können hier Standard-Feldinhalte festgelegt werden und bestimmt werden, ob ein Feld mehrfach vorkommen darf und ob es ein Pflichtfeld ist.

Im **Layout** schließlich kann das graphische Erscheinungsbild eines Eintrags, z.B. Schriftgröße und -farbe definiert werden, außerdem können Felder dort umbenannt werden, z.B. in eine andere Sprache.

Sowohl im Eingabemodell als auch im Layout besteht die Möglichkeit, Felder der ursprünglich angelegten Hierarchie wegzulassen. Eine Termbank kann über mehrere Eingabemodelle und mehrere Layouts verfügen, so dass auf diese Weise interessens- und kompetenzspezifische Bearbeitungsprofile angelegt werden können. So könnte für Linguisten, Übersetzer oder Sprachlernende die Ansicht auf diejenigen Felder beschränkt werden, die für die jeweilige Tätigkeit von besonderem Interesse sind.

6.2.4 Datenabfrage

Abb. 14 zeigt die Benutzeroberfläche von *MultiTerm*. Der Bereich links oben, das **Suchfenster**, zeigt die Einträge in alphabetischer Reihung gemäß der Sprache, die in der Navigationsleiste oben als Ausgangssprache ausgewählt wurde. Zum Auffinden eines Terms gibt es nun entweder die Möglichkeit der **Suche** oder des Setzens eines **Filters**, der die Auswahl der im Suchfenster angezeigten Einträge nach bestimmten Kriterien einschränkt.

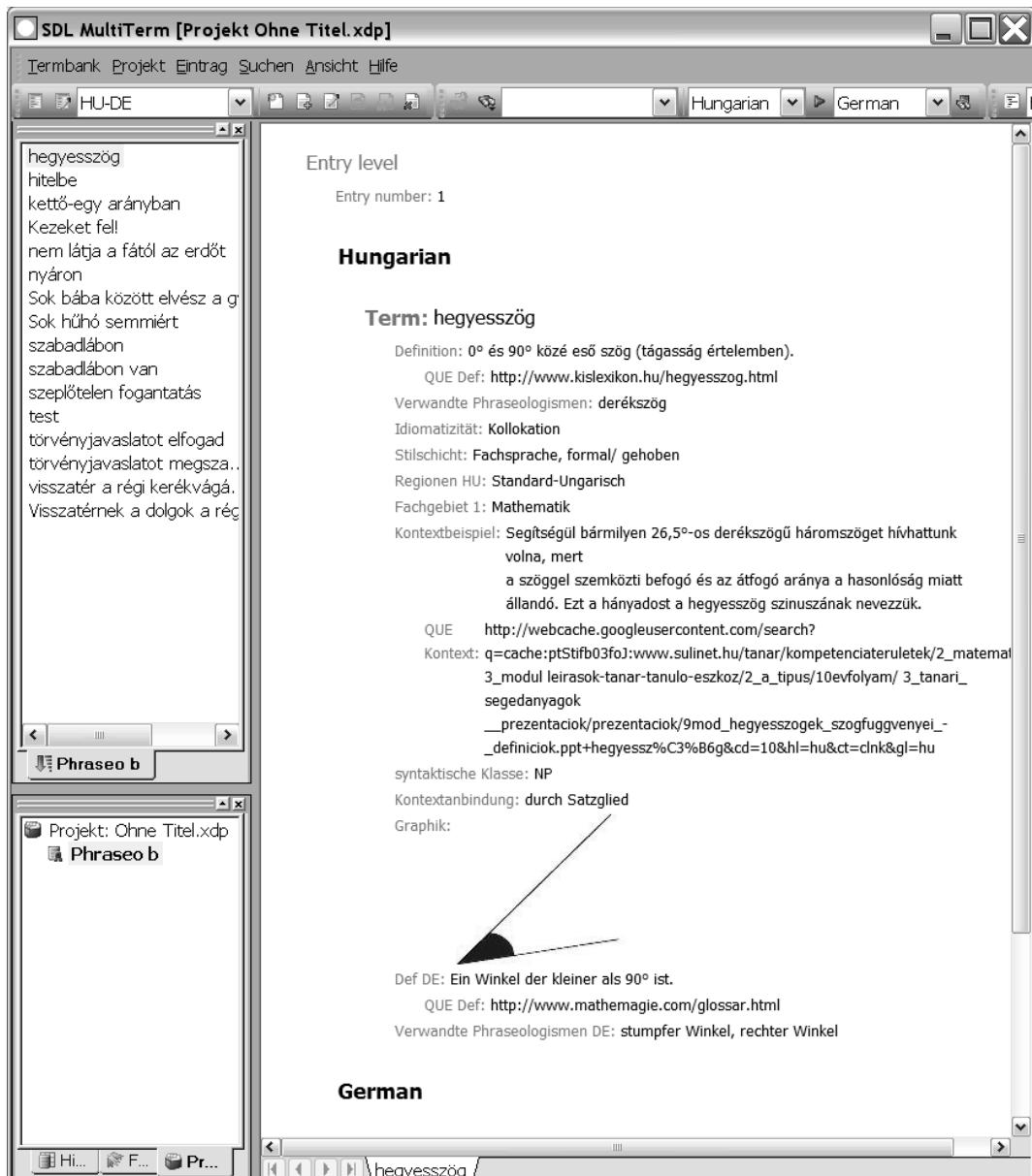


Abb. 14: Benutzeroberfläche von *MultiTerm* mit Suchfenster (links oben), Projektfenster (links unten) und Eintragsfenster (Mitte)

Suchergebnisse werden im sogenannten *Projektfenster* links unten angezeigt. Bei der Suche gibt es drei Möglichkeiten: Die **normale Suche** sucht nach der genauen eingegebenen Zeichenfolge, die auch über Wortgrenzen hinausgehen oder nur mit Wortteilen übereinstimmen kann. Die **unscharfe Suche** (*fuzzy search*) ist sehr fehlertolerant, findet Einträge auch bei leichten Abweichungen, Vertauschung von Zeichen oder Ausschreibungen. So wird der ungarische Phraseologismus *Kezeket fell!* („Hände hoch!“), in dem das Wort „Hand“ im Akkusativ Plural *kezeket* vorkommt, auch bei Eingabe der Grundform *kéz* gefunden. Auch der Plural von *kehely* („Kelch“), *kelyhek*, würde noch gefunden, obwohl hier ein Tausch von Konsonanten stattfindet. Vollständig abweichende Flexionsformen wie *sein* vs. *waren* oder stark von einander abweichende

Stämme in der griechischen Verbalflexion (z.B. bei $\beta\lambda\epsilon\pi\omega$ [*vlepo*] „sehen“, Imperfektstamm $\beta\lambda\epsilon\pi-$ (*vlep-*), Perfektstamm $(\varepsilon)\delta-$ ((*i*)*d-*)) werden hingegen nicht erkannt. In den verwendeten Sprachen traten keine derartigen Fälle auf (nach „waren“ würde vermutlich niemand suchen), aber im Hinblick auf multilinguale Erweiterbarkeit sollten derlei Fälle berücksichtigt werden.

Sowohl die normale als auch die unscharfe Suche suchen jeweils nur im Termfeld der gewählten Sprache. Die dritte Form der Suche, die **Volltextsuche**, sucht in sämtlichen Feldern eines Eintrags. Man kann die Suche aber nicht auf ein bestimmtes Feld begrenzen, wodurch sie bei großen Termbanken unter Umständen sehr lange dauert. Auch die Kombination von unscharfer Suche und Volltextsuche ist nicht möglich.

In 5.2 und 5.4.1.1 wurde die Zuordnung von **Suchbegriffen** zu Einträgen vorgeschlagen. Nötig wäre dies aufgrund der toleranten unscharfen Suche nur zum Zweck einer semantischen Beschlagwortung sowie bei stark abweichenden Flexionsformen; leichtere Abweichungen werden erkannt. *MultiTerm* sieht die Zuordnung von Suchbegriffen zu einem Eintrag jedoch nicht explizit vor; möglich wäre dies nur innerhalb eines normalen Textfeldes *Suchbegriffe*, dessen Inhalt bei einer Volltextsuche gefunden würde.

Die Möglichkeit, innerhalb von bestimmten Feldern zu suchen, besteht indirekt über das Setzen eines **Filters**. Dabei wird im Suchfenster nicht die komplette Eintragsliste, sondern nur eine bestimmten Kriterien entsprechende Auswahl angezeigt. Beim Anlegen eines Filters im Filterassistenten (Abb. 15) markiert man zunächst ein Feld und wählt dann aus einer Liste eine Bedingung aus, die auf dieses Feld zutreffen muss. Unter anderem kann definiert werden, dass ein Feld ein bestimmtes Wort enthalten muss. Verlangt man z.B., dass das Feld *Suchbegriff* das Wort *Aufregung* enthalten muss, findet die Anwendung des Filters den Eintrag *Viel Lärm um nichts*, sofern dieser entsprechend beschlagwortet wurde. Innerhalb eines Filters können mehrere Bedingungen – bezüglich desselben oder verschiedener Felder – kombiniert werden.

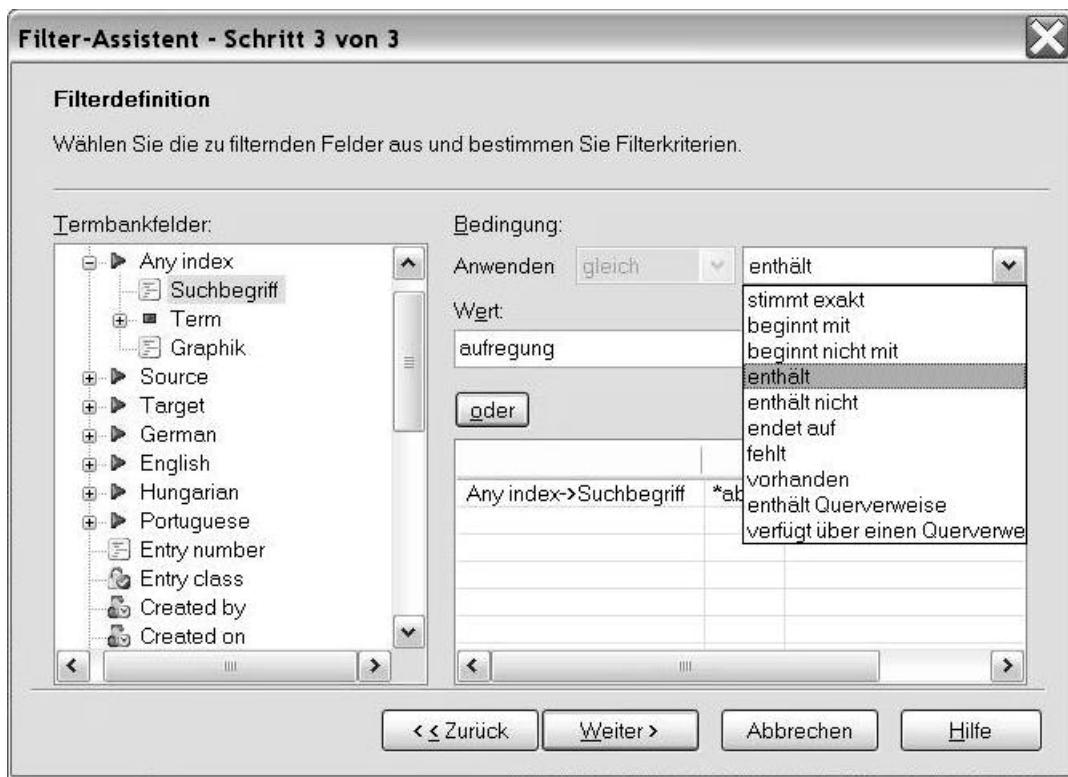


Abb. 15: Setzen eines Filters

Ein Filter wird gespeichert und ist standardmäßig für häufigeren Gebrauch bestimmt. Er ist also nicht von Haus aus als Suchfunktion gedacht, sondern diesbezüglich nur ein Provisorium. Daher wäre es ratsam, einen zum Zweck der Suche angelegten Filter gleich nach Gebrauch wieder zu löschen, da sich sonst eine unüberschaubare Zahl von Filtern in der Datenbank sammelt.

Zur Ansicht eines Eintrags kann wie in 6.2.3 beschrieben ein Layout gewählt werden, dass den jeweiligen Interessensschwerpunkten entspricht (vorausgesetzt, dass entsprechende Layouts zuvor definiert wurden). Innerhalb eines Eintrags selbst können Felder minimiert werden (*Kompaktansicht* bzw. *collapsed structure*), wodurch nur der Feldname angezeigt wird, Feldinhalt und untergeordnete Felder hingegen ausgeblendet werden. In Abb. 16 wurden z.B. im Index *Englisch* die Felder *Status* und *Kontext* sowie die Sprachen Arabisch, Russisch und Serbisch minimiert³⁴.

³⁴ Quelle: <http://mtsrv.trans.univie.ac.at:8090/TermbaseFinderApp/>, dort Asyltermbank_60, eine im Rahmen von Forschungsprojekten am Zentrum für Translationswissenschaft der Universität Wien entstandene Termbank.

Entry level
Entry number:
Created on: 2006-04-11T17:33:42
Entry class:
Fachgebiet: Recht
Unterfachgebiet: Asylrecht

Deutsch

Term: Abweisung des Asylantrages

Quelle: [Merkblatt](#)

Status: freigegeben

Term: Abweisung des Antrages auf internationalen Schutz

Quelle: § 33 [AsylG 2005](#)

Status: freigegeben

Sprachebene: Gesetzliche Benennung

Definition: **Abweisung des Antrages auf internationalen Schutz** ist die negative Entscheidung über den Antrag aufgrund seiner inhaltlichen Prüfung (Prüfung der Fluchtgründe und der Gründe für eine **subsidiäre Schutzgewährung**).

Quelle: [Leikau/Krainz](#)

Englisch

Term: rejection of the application for international protection

Quelle: [Regulation \(EC\) No. 343/2003](#)

Status:

Kontext:

Term:

[Arabisch](#)

Französisch

Term: rejet de la demande de protection internationale

Quelle: [Christine Plunger](#)

Status: neu

Term: rejet de la demande d'asile

Quelle: Règlement (CE) n° 343/2003 [Link](#)

Status: freigegeben

Siehe: [rejeter une demande d'asile sur le fond](#)

Kontext: Les obligations prévues au paragraphe 1, points d) et e), cessent également dès que l'État membre responsable de l'examen de la demande d'asile a pris et effectivement mis en oeuvre, à la suite du retrait ou du rejet de la demande d'asile, les dispositions nécessaires pour que le ressortissant d'un pays tiers se rende dans son pays d'origine ou dans un autre pays où il peut légalement se rendre.

Quelle: Règlement (CE) n° 343/2003 [Link](#)

[Russisch](#)

[Serbisch](#)

Abb. 16: Beispieleintrag mit minimierten Feldern

In den folgenden Abschnitten wird untersucht, wie *MultiTerm* für phraseographische Zwecke genutzt werden kann. Dafür wurden drei verschiedene Datenmodelle erstellt, deren Vor- und Nachteile jeweils besprochen werden sollen.

6.3 Phraseographische Nutzung von *MultiTerm*

Wie in 5.3 beschrieben, ist es hilfreich, einsprachige und übersetzungsbezogene Daten getrennt zu verwalten, um den Bestand einer Sprache leicht mit weiteren Zielsprachen verknüpfen zu können bzw. um einem Phraseologismus mehrere mögliche Übersetzungen mit jeweiliger Erläuterung zuordnen zu können. Eine solche Möglichkeit sieht *MultiTerm* nicht explizit vor.

Daher muss ein Weg gefunden werden, die übersetzungsbezogenen Daten dennoch auf sinnvolle Weise einzugliedern, und zwar möglichst so, dass sie für jede Übersetzungskombination getrennt angegeben werden können und dass ein bestehender einsprachiger Datensatz unkompliziert mit weiteren Übersetzungen verknüpft werden kann, sei dies in derselben oder einer weiteren Zielsprache. Dafür wurden drei verschiedene Eintragsmodelle entworfen, die in 6.3.3 besprochen werden.

6.3.1 Einsprachige Daten

Die Herausforderung beim Entwurf eines geeigneten Datenmodells bestand darin, die übersetzungsbezogenen Daten auf geeignete Weise unterzubringen. Die einsprachigen Daten wurden jedoch von Modell zu Modell nicht verändert – sie liegen einfach auf der Termebene unter dem betreffenden Phraseologismus. Daher werden sie bei der Beschreibung der Modelle zum Teil zusammenfassend dargestellt. Tabelle 7 zeigt eine Übersicht über die verwendeten einsprachigen Daten. Die Felder seien hier nur kurz genannt; der genaue Inhalt entspricht den Ausführungen in 5.4. Die Optionen der mit * markierten Picklist-Felder werden in Tabelle 6 aufgeschlüsselt.

| Stilschicht | | | |
|---|---|---|---|
| formal/ gehoben umgangssprachlich Slang vulgär | Allgemeinsprache Fachsprache | mündlich schriftlich | Neologismus veraltet veraltet |
| dialektal | literarisch journalistischer Stil | pejorativ euphemistisch | Gaunersprache Jugendsprache |
| Regionen DE | Regionen EN | Regionen HU | Regionen PT |
| Alemannisch CH + Ö Alemannisch D + F Bairisch D Bairisch Ö Mitteldeutsch Niederdeutsch/ Friesisch Ostdeutsch | Afrika Australien Indien Irland Kanada Neuseeland USA Süd- und Mittelamerika Südostasien und Pazifik Vereinigtes Königreich | Ungarische Standardsprache Nordwestdialekte (palóc nyelvjárások) Mittelsiebenbürgische und Szekler-Dialekte (mezőségi és székely nyelvjárások) | Angola Brasilien Guinea-Bissau Indien Kap Verde Macau Mosambik Osttimor Portugal São Tomé e Príncipe |
| Fachgebiet (exemplarisch) | syntakt. Klasse | Externe Leerstellen | |
| Allgemeinsprache Biologie Mathematik Medizin Physik Politik Recht Technik Wirtschaft | Satz Satz mit Leerstelle satzwertiges Gebilde Nebensatz Nominalphrase Verbalphrase Prädikativ Adjektivphrase Adverbialphrase Präposition Postposition Interjektion Konjunktion Subjunktion | Subjekt Prädikat Akkusativobjekt Dativobjekt Genetivobjekt Ergänzung Attribut NP VP AdjP AdvP PP | |

Tabelle 6: Einträge der Picklist-Felder

| Feld | | Feldart; Einfach- (1) bzw Mehrfachnennung (x) | Inhalt |
|--------------------------------------|----------------|--|---|
| Term | | | |
| Def | Text | 1 | Definition |
| QUE Def | Text | 1 | Quelle der Definition |
| Syn/ Paraphrase | Text (Verweis) | x | synonyme Mehr- oder Einworteinheiten |
| Variante | Text (Verweis) | x | andere Repräsentanten mit minimaler Abweichung |
| Verwandte Phraseologismen | Text (Verweis) | x | nicht-synonyme Phraseologismen mit verwandter Beziehung zum Eintrag, z.B. Teil- oder Obermengen, Antonyme, Phraseologismen mit ähnlichem Bild |
| Phraseologizität | Boolean | 1 | Angabe, ob Eintrag phraseologisch ist (ja/nein) |
| Idiomatizität | Picklist | 1 | Idiom/ Teil-Idiom/ Kollokation |
| Stilschicht* | Picklist | x | verschiedene Angaben zu Stil und Konnotation |
| Regionen (DE/EN/HU/PT)* | Picklist | x | Regionen, in denen der Phraseologismus geläufig ist |
| Fachgebiet* | Picklist | x | Auswahl aus mehreren Fachgebieten |
| Kontextbeispiel | Text | x | einsprachiges Kontextbeispiel |
| QUE Kontext | Text | 1 | Quelle des Kontextes |
| Typische Textsorten | Text | 1 | ggfs. Textsorten, die Eintrag typischerweise enthalten |
| Bemerkungen zum Gebrauch | Text | 1 | ggfs. Erläuterung des Verwendungskontextes |
| Syntakt. Klasse* | Picklist | 1 | NP/ VP/ Akk.-Objekt/ etc. |
| Kontextanbindung | Picklist | 1 | keine/ logisch/ durch Satzglied |
| Externe Leerstellen (obligatorisch)* | Picklist | x | zur syntaktischen Vollständigkeit zwingend |
| semant. Restriktionen | Text | x | Spezifizierung der Leerstelle, z.B. "menschlich", "Vorgang" |
| Externe Leerstellen (fakultativ) | Picklist | x | syntaktisch nicht zwingend, aber typisch |
| semant. Restriktionen | Text | x | s.o. |
| Transformations-Defekte | Picklist | x | z.B. Konstituentennegation, Expansion durch Attribute, Quantifizierung, Tempusabwandlung |
| Traditionelle Klasse | Picklist | x | Sprichwort/ Redewendung/ Zwillingsformel/ etc. |
| Besonderheiten | Picklist | x | unikale Komp./ untypische Valenz/ syntakt. Irregularität |
| Ursprung | Text | 1 | ursprüngliche Quelle, falls bekannt |
| Graphik | Multimedia | x | |
| Def ZS | Text | 1 | zielsprachliche Definition |
| QUE Def ZS | Text | 1 | Quelle der Definition |
| SYN/ Paraphrase ZS | Text (Verweis) | 1 | Synonym bzw. Paraphrase in der Zielsprache |
| Verwandter Phraseologismus ZS | Text (Verweis) | x | verwandter Phraseologismus in der Zielsprache |

Tabelle 7: Einsprachige Daten mit Eingabemodus und Inhalt

Eine Methode wäre daher, einen **Eintrag als phraseologisch oder nicht-phraseologisch** zu kennzeichnen. *MultiTerm* bietet hierfür die Möglichkeit, ein Feld als Boolesche Variable – neben Text- und Optionsfeld eine weitere Feldart – anzulegen, bei der durch ein einzelnes Ankreuzfeld eine Eigenschaft zugewiesen werden kann oder nicht. Eine andere Möglichkeit wäre, eine freie Wortverbindung von vornherein nur als ziel-sprachliche Paraphrase, nicht aber als Äquivalent einzutragen. Welche Methode geeigneter ist, kann hier aufgrund mangelnder einschlägiger Beispieldaten nicht abschließend geklärt werden.

Die Auswahl der Felder ist recht ausführlich und sollte natürlich durch Erstellen eines geeigneten Layouts bedarfsgerecht eingegrenzt werden (s. 5.5.4/5). Die Felder *Traditionelle Klasse*, *Besonderheiten* und *Ursprung* des Phraseologismus sind vor allem von linguistischem Interesse, während Übersetzer vor allem Angaben zum Gebrauch wie *Kontextbeispiel*, *Stilschicht* oder *Textsorte* benötigen, für Sprachlernende kann zusätzlich syntaktische Information von Interesse sein etc. Entsprechend ließe sich diese Liste auch ausweiten, etwa um die Angabe eines Unternehmens bei firmenspezifischer Terminologie, um etymologische Daten etc.

Andererseits sollten einige Felder im Praxisfall in weitere Unterfelder aufgeschlüsselt werden, was im Rahmen der zu besprechenden Testdatenbanken nicht möglich war. Bei den *Fachgebieten* sollte eine vollständige Fachgebietssystematik wie z.B. die Lenoch-Klassifikation (vgl. Arntz et al. ⁵2004:241) integriert werden. Zu Versuchs-zwecken wurden hier nur einige Beispiefachgebiete aufgelistet. Eine weitere Aufschlüsselung wäre auch bei allen Quellenangaben angezeigt; hier könnten Unterdatenfelder Art des Mediums sowie bibliographische Daten enthalten.

Auch die Felder *Restriktionen* sollten idealerweise nicht als Text-, sondern als Optionsfeld gestaltet sein. Ein Textfeld wurde als provisorische Lösung verwendet, da zum Zeitpunkt des Erstellens dieser Arbeit keine umfassende Typologie möglicher semantischer Restriktionen³⁵ bzw. kein semantisches Netz vorlag. Auch bei den *typischen Textsorten* wäre eine solche Entwicklung denkbar – auf Basis häufig vorkommender Nennungen ließe sich später ein Optionsfeld erstellen.

Allgemein ist davon auszugehen, dass sich im Laufe der Verwendung einer Datenbank aus der Beschaffenheit der Daten Bedarf an weiteren Feldern und Picklist-Optionen ergibt. Daher sollte eine Datenbank grundsätzlich flexibel, sprich: ausbaufähig gestaltet sein. Bei Optionsfeldern ist daher die Aufnahme der Option *Sonstiges* unerlässlich. Strenggenommen sollte auch unter jedem Optionsfeld ein Textfeld zur Verfügung

³⁵ Syntaktische Restriktionen werden durch die Notation der Nennform und der externen Leerstelle dargestellt.

gestellt werden, um dieses Sonstige gegebenenfalls zu spezifizieren. Darauf wurde hier zur Vereinfachung verzichtet.

6.3.2 Übersetzungsbezogene Daten

Tabelle 8 zeigt die übersetzungsbezogenen Datenfelder, die in den Beispieldatenbanken angelegt wurden.

| Feldname | Feldart; Einfach- (1) bzw. Mehrfachnennung (x) | Inhalt |
|----------------------------|---|--|
| ? | | |
| Äquivalenz AS-ZS | Picklist | 1 total - pragmatisch - partiell |
| Kommentar AS | Text | 1 Kommentar zur Äquivalenz in der Ausgangssprache |
| Kommentar ZS | Text | 1 Kommentar zur Äquivalenz in der Zielsprache |
| Übersetzungsbeispiel AS-ZS | Text | x Übersetzungsbeispiel (Ausgangstext) |
| QUE AT | Text | 1 Quelle des Ausgangstextes |
| Zieltext | Text | x Zieltext |
| QUE ZT | Text | 1 Quelle des Zieltextes (z.B. Name der Übersetzerin) |

Tabelle 8: Übersetzungsbezogene Datenfelder

Das Fragezeichen weist darauf hin, dass es vom jeweiligen Datenmodell abhängt, wo in der Datenbank die übersetzungsbezogenen Datenfelder untergebracht werden. „AS“ und „ZS“ stehen hier als Variablen für die Ausgangs- und Zielsprache, die dem Feldnamen jeweils beigefügt werden, damit klar ist, worauf sich die Daten beziehen. Es wird somit nicht ein Feld Äquivalenz AS-ZS angelegt, sondern je eines pro Übersetzungsrichtung. Der Fall der Null-Äquivalenz wurde nicht angeführt, da in diesem Fall erst gar kein Übersetzungsdatensatz angelegt wird.

Das Feld *Übersetzungsbeispiel* enthält nur den Ausgangstext; der Zieltext befindet sich in einem untergeordneten Feld. Logischer wäre, beide Felder gleichrangig unter eine Art Überschrift „Übersetzungsbeispiel“ zu legen; ein „leerer“ Knoten ohne Feldinhalt zur bloßen Strukturierung lässt sich in *MultiTerm* jedoch nicht anlegen.

6.3.3 Mögliche Datenmodelle

6.3.3.1 A: Übersetzungsdaten als Indexfeld

Die erste Idee, übersetzungsbezogene Daten getrennt zu verwalten, bestand darin, neben den Sprachen eigene Indexfelder für die Übersetzungsdaten anzulegen. Für eine deutsch-ungarische Termbank gäbe es also folgende Indexfelder:

| | |
|---------|-------------------------------------|
| DE | Deutsch |
| HU | Ungarisch |
| Ü DE-HU | Übersetzungsdaten Deutsch-Ungarisch |
| Ü HU-DE | Übersetzungsdaten Ungarisch-Deutsch |

Beide Übersetzungsrichtungen zu trennen scheint sinnvoll, da sich hinsichtlich der Äquivalenzbeziehung Asymmetrien ergeben können, wodurch Erläuterung und Übersetzungsbeispiele entsprechend unterschiedlich ausfallen können. Abb. 17 zeigt einen Beispieleintrag dieses Datenmodells.

Abb. 17 (umseitig): Eintragsmodell A - Beispieleintrag

Entry level

Entry number: 62

Hungarian

Term: [vmt] színe-jáva

phraseologisch: ✓

Definition: emberi közösségek, termények, terméknek a legértékesebb része.

QUE Def: Magyar Értelmező Kéziszótár

Idiomatizität: Idiom

Fachgebiet: Allgemeinsprache

Stilschicht: Allgemeinsprache, journalistisch

Kontext: Az óceán kincsei - a tenger gyümölcséinek színe-jáva

QUE http://www.fo.hu/hu/konyv/ismeretterjeszto_1/

Kontext: szakacsoknyvek/az_ocean_kincsei_a_tenger_gyumolcseinek_szine_java#bovebben

Kontext: A magyar sport színe-jáva

QUE Kontext: http://www.mno.hu/portal/682672

syntaktische Klasse: NP

Externe Leerstellen (obligatorisch): Attribut

Restriktionen: birtokos/ Besitzer

German

Term: die erlesensten [NP]

phraseologisch: X

Definition: Eine Auswahl der besten, qualitativ hochwertigsten Elemente einer Menge. Bezieht sich häufig auf Kulinarisches, aber auch auf Düfte.

QUE Def: HK

Syn, Paraphrase: die besten [NP]

Idiomatizität: Kollokation

Fachgebiet: Allgemeinsprache

Stilschicht: Allgemeinsprache, formal/ gehoben, literarisch

syntaktische Klasse: NP mit Leerstelle

Externe Leerstellen (obligatorisch): NP

Restriktionen: nicht menschlich, konkret, häufig essbar oder trinkbar

Ü HU-DE

Term: [vmt] színe-jáva - die erlesensten [NP]

Äquivalenz: partiell

Kommentar DE: "erlesen" häufig im Zusammenhang mit kulinarischen Genüssen (Wein).

Nicht bei Personen.

Kommentar HU: Az "erlesen" szót sokszor a kulináris specialitásokkal kapcsolatban használjuk (pl. bor). nem vonatkozhat személyekre.

Übersetzungsbispiel: Az óceán kincsei - a tenger gyümölcséinek színe-jáva

QUE AT: http://www.libri.hu/konyv/az-ocean-kincsei.html (Buchtitel von Robert Longworth)

Zieltext: Schätze des Ozeans - die erlesensten Meeresfrüchte

QUE ZT: HK

Ü DE-HU

Term: die erlesensten [+NP] - [vmt] színe-jáva

Äquivalenz: partiell

Kommentar DE: "[vmt] színe-jáva" kann auch in anderen Kontexten stehen als "erlesen" - allg. das Beste, die Elite von etw/jdm

Übersetzungsbispiel: Im Restaurant wird Ihnen eine leckere ungarische Küche serviert und erlesenste Weine aus der Region Villány.

QUE http://booking.budapest-hotel-

AT: guide.hu/ungarn/villany/hotels/crocus_gere_bor_hotel_wine_spa/hotel-informationen.de.html

Zieltext: Az étteremben ízletes magyar ételeket és a villányi térség borainak színe-jávát szolgáljuk fel Önnek.

QUE ZT: HK

Übersetzungsdaten werden wie eine weitere Sprache behandelt, enthalten jedoch andere Unterdatenkategorien. Dazu muss bei der Termbankdefinition zweierlei beachtet werden:

Erstens sieht *MultiTerm* nicht vor, einen Index für etwas anderes als eine Sprache zu verwenden. Daher muss man sich damit behelfen, in der Termbankdefinition pro forma irgendeine Sprache zu wählen und diese z.B. in \ddot{U} DE-HU umzubenennen. Der Indexname lässt sich problemlos ändern, das zugehörige unzutreffende Sprachenkürzel bleibt jedoch bei den Übersetzungsdaten stehen, wie in Abb. 18 ersichtlich ist.

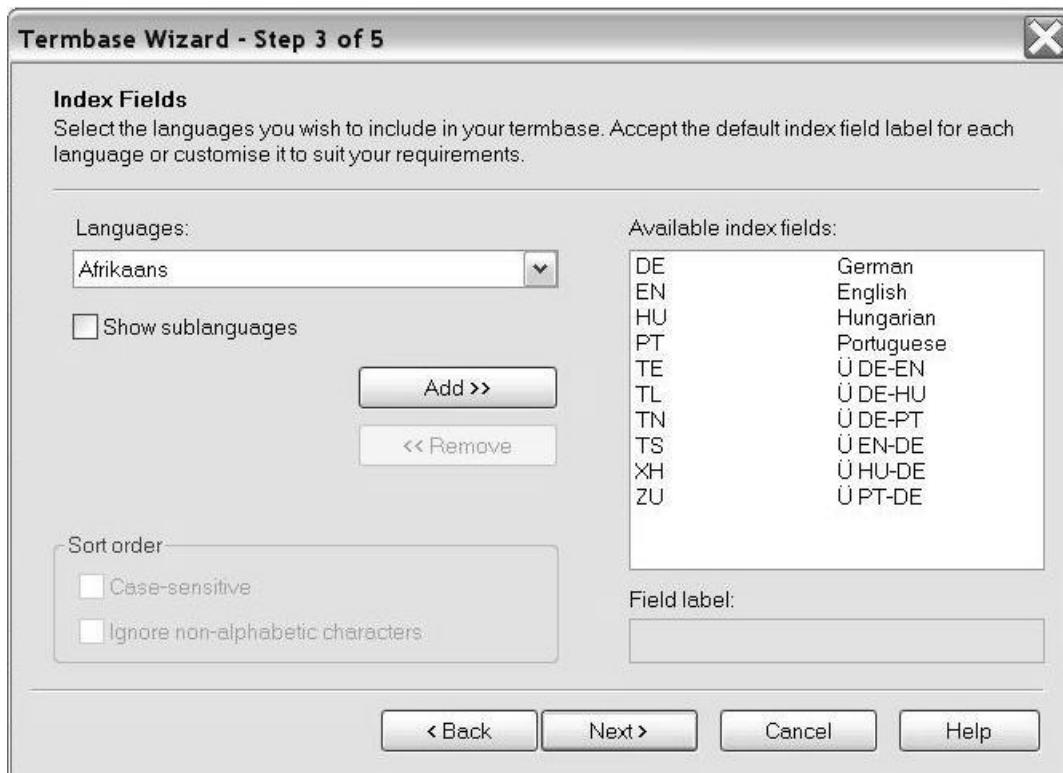


Abb. 18: Umbenennung von Indexfeldern

Die Beispieltermbank, aus der Abb. 17 stammt, wurde für die Sprachen Deutsch, Englisch, Portugiesisch und Ungarisch erstellt, wobei Übersetzungsbeziehungen jeweils nur zum und aus dem Deutschen hergestellt werden. Bereits hier besteht Bedarf an sechs verschiedenen Indexfeldern für Übersetzungsdaten; bei mehr Sprachen und vielfältigeren Verknüpfungen würde diese Zahl exponentiell ansteigen, was die Erweiterungsfähigkeit der Termbank klarerweise einschränkt.

Zweitens muss beim Speichern der Übersetzungsdaten als eigener Index ein besonderes Augenmerk auf die hierarchische Anordnung in der Termbankdefinition gelegt werden (s. 6.2.3). Die hierarchische Struktur in der Termbankdefinition orientiert sich zunächst nur daran, auf welcher Ebene ein Feld angesiedelt werden soll (Eintrags-, Index-, Termebene oder tiefer), nicht aber daran, unter welchem Index, sprich, unter welchem spezifischen Ast der Termbank ein Feld später benötigt wird. Daher

müssen einsprachige und Übersetzungsdaten in der Termbankdefinition nebeneinander gelegt werden, wie Abb. 19 zeigt:

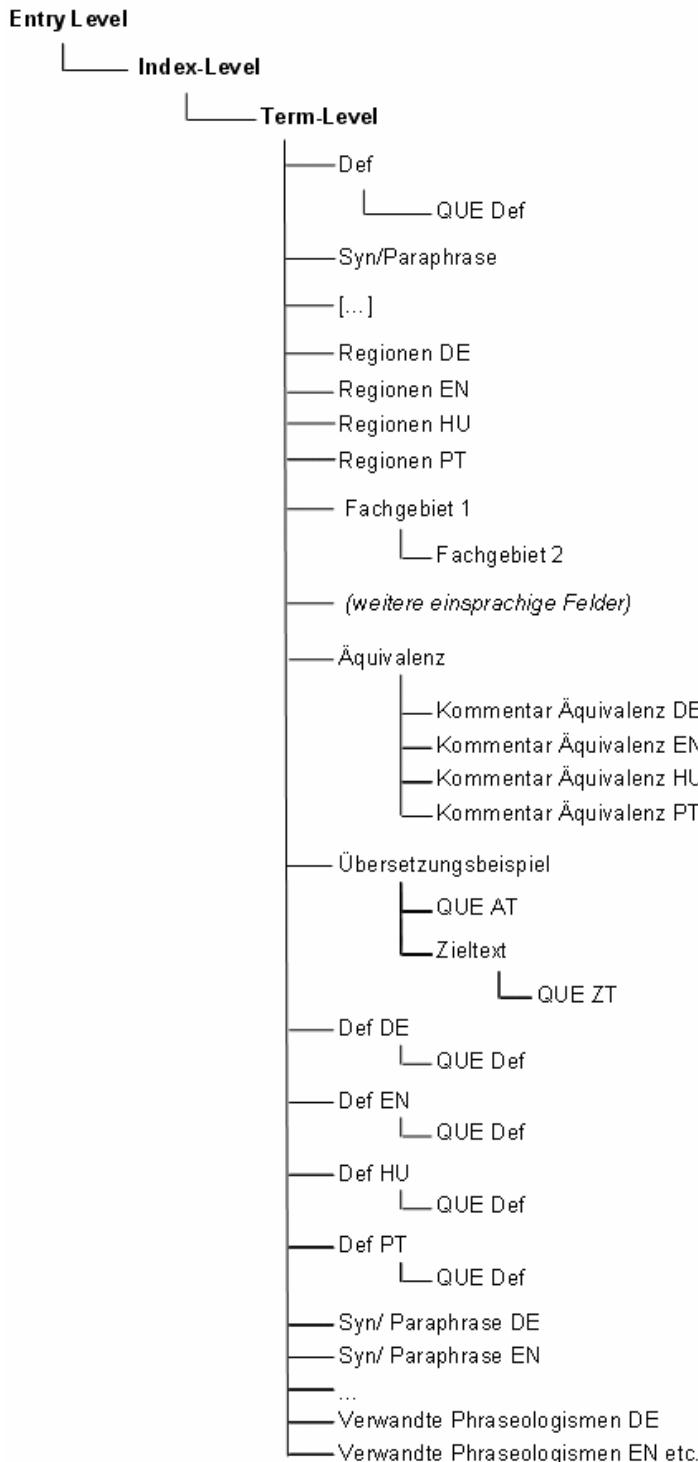


Abb. 19: Nebeneinander von einsprachigen und übersetzungsbezogenen Feldern in der Termbankdefinition

Auch sprachspezifische Felder, die jeweils nur zu einem Index gehören, werden hier nebeneinander gereiht. Im Eingabemodell (s. Abb. 13) werden die Felder dann nur un-

ter dem entsprechenden Index – Sprache oder Übersetzungsdatensatz – zur Verfügung gestellt, wie aus dem Beispieleintrag in Abb. 17 ersichtlich ist.

Die Aufnahme eines nur einsprachigen Datensatzes – für den Fall, dass ein ziel-sprachliches Äquivalent entweder nicht vorhanden oder dem User nicht bekannt ist – ist bei diesem Datenmodell ohne weiteres möglich. In diesem Fall wird der ausgangssprachliche Index bei der Eingabe ausgefüllt, die anderen Indices gelöscht. Die gelöschten Indices können später bei Bedarf wieder hinzugefügt werden.

Für die Umsetzung dieses Datenbankmodells stellt sich die Frage, ob ein Eintrag nur **genau ein zweisprachiges Phraseologismenpaar** enthalten soll wie im obigen Beispiel, oder **mehrere** – *MultiTerm* bietet die Möglichkeit, einem Index mehrere Terme, d.h. Synonyme in einer Sprache, unterzuordnen. Letzteres hätte den Vorteil, dass mehrere Übersetzungsvarianten auf einen Blick sichtbar werden und User sich nicht durch verschiedene Einträge mit demselben ausgangssprachlichen Phraseologismus klicken müssen.

Abb. 20 zeigt einen Versuch, **mehrere Übersetzungsvorschläge** in einen Eintrag dieses Datenmodells aufzunehmen.

Abb. 20 (umseitig): Mehrere Äquivalente in einem Eintrag

Entry level

Entry number: 62

Hungarian

Term: [vmi] színe-java

phraseologisch: ✓

Definition: emberi közösségek, termények, termékek a legértékesebb része.

QUE Def: Magyar Értelező Kéziszótár

Idiomatizität: Idiom

Fachgebiet: Allgemeinsprache

Stilschicht: Allgemeinsprache, journalistisch

Kontext: Az óceán kincsei - a tenger gyümölcséinek színe-java

QUE http://www.fo.hu/hu/konyv/ismeretterjeszto_1/

Kontext: szakacsoknövek/az_ocean_kincsei_a_tenger_gyumolcseinek_szine_java#bovebben

Kontext: A magyar sport színe-java

QUE Kontext: http://www.mno.hu/portal/682672

syntaktische Klasse: NP

Externe Leerstellen (obligatorisch): Attribut

Restriktionen: birtokos/ Besitzer

German

Term: (die) Spitzen- [+ Gen.Attr.]

phraseologisch: X

Definition: Diejenigen Personen einer Gruppe, die in einem bestimmten Bereich Höchstwerte/
Höchstleistungen erreichen, idR -sportler, -kandidaten(Politik), -verdiener

QUE Def: HK

Syn, Paraphrase: Top-

Idiomatizität: Kollokation

Fachgebiet: Allgemeinsprache

Stilschicht: Allgemeinsprache, journalistisch

Kontext: In der "Wahlarena 2010" im WDR Fernsehen stellen sich die Spitzenkandidaten von CDU,
SPD, Grüne, FDP und Linke den Fragen von NRW-Bürgern.

QUE http://www.wdr.de/themen/politik/landtagswahl_2010/kandidatenduell/100428.jhtml
Kontext:

Bemerkungen zum Gebrauch: nicht phraseologisch

syntaktische Klasse: NP mit Leerstelle

Externe Leerstellen (obligatorisch): Substantiv

Restriktionen: menschlich, häufig -sportler, -kandidaten, -politiker

Externe Leerstellen (fakultativ): Attribut

Restriktionen: Genetiv

Term: die erlesensten [NP]

phraseologisch: X

Definition: Eine Auswahl der besten, qualitativ hochwertigsten Elemente einer Menge. Bezieht sich häufig auf Kulinarisches, aber auch auf Düfte.

QUE Def: HK

Syn, Paraphrase: die besten [NP]

Idiomatizität: Kollokation

Fachgebiet: Allgemeinsprache

Stilschicht: Allgemeinsprache, formal/ gehoben, literarisch

syntaktische Klasse: NP mit Leerstelle

Externe Leerstellen (obligatorisch): NP

Restriktionen: nicht menschlich, konkret, häufig essbar oder trinkbar

Ü HU-DE

Term: [vmi] színe-java - (die) Spitzen- [+ Gen.Attr.]

Äquivalenz: partiell

Kommentar der bestimmte Artikel wird bei nachgestelltem Genetivattribut benötigt:
DE: "Ungarns Spitzensportler",
aber "die Spitzenpolitiker des Landes".

Kommentar a határozott névelő a főnevet követő birtokos jelző esetén szükséges: "Ungarns
HU: Spitzensportler",
de "die Spitzenpolitiker des Landes".

Übersetzungsbeispiel: A magyar sport színe-java

QUE AT: <http://www.mno.hu/portal/682672>

Zieltext: Ungarns Spitzensportler

QUE ZT: HK

Term: [vmi] színe-java - die erlesensten [NP]

Äquivalenz: partiell

Kommentar "erlesen" häufig im Zusammenhang mit kulinarischen Genüssen (Wein). Nicht bei

DE: Personen.

Kommentar Az "erlesen" szót sokszor a kulináris specialitásokkal kapcsolatban használjuk (pl.
HU: bor). nem vonatkozhat személyekre.

Übersetzungsbeispiel: Az óceán kincsei - a tenger gyümölcséinek színe-java

QUE AT: <http://www.libri.hu/konyv/az-ocean-kincsei.html> (Buchtitel von Robert Longworth)

Zieltext: Schätze des Ozeans - die erlesensten Meeresfrüchte

QUE ZT: HK

Ü DE-HU

Term: die erlesensten [+NP] - [vmi] színe-java

Äquivalenz: partiell

Kommentar "[vmi] színe-java" kann auch in anderen Kontexten stehen als "erlesen" - allg.
DE: das Beste, die Elite von etw/jdm

Übersetzungsbeispiel: Im Restaurant wird Ihnen eine leckere ungarische Küche serviert und
erlesenste Weine aus der Region Villány.

QUE <http://booking.budapest-hotel->

AT: guide.hu/ungarn/villany/hotels/crocus_gere_bor_hotel_wine_spa/hotel-informationen.de.html

Zieltext: Az étteremben ízletes magyar ételeket és a villányi térség borainak színe-javát
szolgáljuk fel Önnek.

QUE ZT: HK

Term: (die) Spitzen[NP] [Gen.Attr.] - [vmi] színe-java

Äquivalenz: partiell

Kommentar a határozott névelő a főnevet követő birtokos jelző esetén szükséges: "Ungarns
HU: Spitzensportler", de "die Spitzenpolitiker des Landes".

Übersetzungsbeispiel: Im Restaurant wird Ihnen eine leckere ungarische Küche serviert und erlesenste
Weine aus der Region Villány.

QUE <http://booking.budapest-hotel->

AT: guide.hu/ungarn/villany/hotels/crocus_gere_bor_hotel_wine_spa/hotel-informationen.de.html

Zieltext: Az étteremben ízletes magyar ételeket és a villányi térség borainak színe-javát szolgáljuk fel
Önnek.

QUE ZT: HK

Es wurden nur zwei deutsche Äquivalente für den ungarischen Phraseologismus aufgenommen; trotzdem zeigt sich bereits, dass der Eintrag auf diese Weise unübersichtlich lang wird. Bei mehreren Äquivalenten im Zielsprachenindex (hier Deutsch) rücken die Übersetzungsdaten weiter von den einsprachigen Daten weg, was optisch Verwirrung stiftet. Zwar bietet *MultiTerm* die Möglichkeit, Felder in der Ansicht zu minimieren, so dass nur der Feldname sichtbar ist. Jedoch wird hierbei auch der Feldinhalt, also die Äquivalente, ausgeblendet, und jeweils nur noch die Feldbezeichnung *Term* angezeigt. Somit ist dadurch keine bessere Übersicht gewonnen ist. Zur Schnellübersicht über mehrere Übersetzungsmöglichkeiten müssen daher doch wieder alle Termfelder geöffnet werden.

Wie Beispiele weiter unten (Abb. 20) zeigen werden, kann ein Phraseologismus unter Umständen sehr viele zielsprachliche Äquivalente haben. Dies leuchtet bereits ein, wenn man allein die Zahl einsprachiger phraseologischer Repräsentanten eines Begriffs betrachtet. Z.B. für <sterben> gibt es auf Deutsch: *ins Gras beißen, das Zeitliche segnen, den Löffel abgeben, ums Leben kommen, [Jds] letztes Stündlein hat geschlagen* etc.

Sollen in einem Eintrag drei oder mehr Sprachen mit wechselseitigen Übersetzungsbereichungen angezeigt werden, liegt die Übersichtlichkeit des Eintrags wohl jenseits des Zumutbaren.

Ein weiterer Nachteil besteht darin, dass eine klare Entscheidung getroffen werden muss, welche Phraseologismen zu einem Begriff zusammengefasst werden sollen und welche nicht. Da, wie gezeigt wurde, in den meisten Fällen keine totale Äquivalenz vorliegt, dürfte die Entscheidung, welche Äquivalente noch in den Eintrag aufgenommen und welche gesondert verzeichnet werden, nicht immer leicht und teilweise willkürlich sein.³⁶

Weiterhin wäre bei der Aufnahme mehrerer zielsprachlicher Äquivalente zu bedenken, dass für jedes Phraseologismenpaar ein eigener Übersetzungsdatensatz benötigt wird, bei dem jeweils kenntlich gemacht werden muss, auf welche beiden Phraseologismen er sich bezieht. Dies geschieht, indem im Termfeld des Übersetzungsindex der ausgangs- und zielsprachliche Phraseologismus erneut eingetragen werden. Da das Termfeld als Pflichtfeld aber ohnehin ausgefüllt werden muss, stellt dies keinen eigentlichen Nachteil im Vergleich zur Aufnahme nur eines Phraseologismenpaares dar. Ein

³⁶ Ein plausibler Ansatz wäre in einem solchen Fall, alle Phraseologismen mit rein denotativer Begriffsäquivalenz – ungeachtet konnotativer Unterschiede – in einem Eintrag zusammenzufassen, und alle stärker abweichenden Phraseologismen (die nur kontextbezogen bzw. bei entsprechender Aktualisierung zu Äquivalenten werden) in eigenen Einträgen zu verzeichnen, zu denen Cross-References gelegt werden.

Manko ist jedoch die Tatsache, dass innerhalb ein- und desselben Eintrags keine Querverweise („Cross-References“) gelegt werden können, und eine Verlinkung von Phraseologismen zu den jeweiligen Übersetzungsdatensätzen somit nicht möglich ist.

Eine Vereinfachung ergibt sich aus der Aufnahme mehrerer zielsprachlicher Äquivalente nicht. Das Anlegen separater Einträge für weitere zielsprachliche Äquivalente lässt sich in *MultiTerm* sehr einfach umsetzen, indem man einen bestehenden Eintrag dupliziert, so dass im Duplikat nur noch die zielsprachlichen und übersetzungsbezogenen Daten geändert werden müssen.

Alles in allem ist die Aufnahme mehrerer Äquivalente in einen Eintrag somit nicht geeignet, insbesondere aufgrund mangelnder Übersichtlichkeit bei vielen Äquivalenten. Der Ansatz wird daher wieder verworfen.

Mehrere Sprachen

Weiterhin soll kurz skizziert werden, wie sich das Datenmodell A bei der Aufnahme weiterer Sprachen in die Datenbank verhält. Da ohnehin jeder Eintrag nur ein Phraseologismenpaar enthält, ändert sich in der Struktur eines Eintrags nichts. Soll ein vorhandener Phraseologismus in eine weitere Zielsprache übersetzt werden, wird ein bestehender Eintrag dupliziert, die vorhandene Zielsprache aus dem Duplikat gelöscht und eine neue hinzugefügt. Auch hier wird wiederum nur *ein* zielsprachliches Äquivalent eingetragen; für jedes weitere Äquivalent wird das Verfahren wiederholt.

Wichtig ist dabei, für das Duplizieren und Ändern eines Eintrags ein Eingabemodell und Layout zu wählen, in dem alle betreffenden Sprachen und jeweiligen Übersetzungsindices angezeigt werden. Die Sichtbarkeit der ersten Ausgangssprache ist wichtig, damit die ursprüngliche Übersetzung im Duplikat gelöscht werden kann; das Hinzufügen der neuen Zielsprache muss ebenfalls möglich sein.

Nachteile dieses Modells

Insgesamt weist der Ansatz, Übersetzungsdaten als Indexfeld anzulegen, etliche Nachteile auf, unabhängig davon, ob ein Eintrag nur ein Phraseologismenpaar oder mehrere enthält.

Ein wesentliches Manko besteht in der beschriebenen Tatsache, dass zum Anlegen der Übersetzungsindices nur vorhandene Sprachindices umfunktioniert werden können, wodurch die Datenbank nur eingeschränkt erweiterbar ist. Für eine Testphase genügen sicherlich wenige Sprachen; für eine langfristige Etablierung einer phraseologischen Datenbank sollte diese jedoch zumindest theoretisch uneingeschränkt erweiterbar sein.

Ein weiteres Problem ergibt sich daraus, dass jeder Index automatisch ein Termfeld enthält, das Pflichtfeld ist und ausgefüllt werden muss. Im Übersetzungsindex gibt es von Natur aus keinen Term, daher muss ein Inhalt definiert werden. Sinnvollerweise könnten hier beide Phraseologismen – nacheinander in Reihenfolge der Überset-

zungsrichtung und durch einen Spiegelstrich getrennt – eingetragen werden, wie dies im Beispieleintrag in Abb. 17 und Abb. 20 erfolgt ist. Bei der Aufnahme mehrerer Äquivalente in einen Eintrag muss ohnehin die Zugehörigkeit des Übersetzungsdatensatzes zu den jeweiligen Phraseologismen kenntlich gemacht werden. Enthält die Datenbank aber nur ein Phraseologismenpaar pro Eintrag, was sich als die bessere Lösung erwiesen hat, bedeutet die erneute manuelle Eingabe einen zusätzlichen Arbeitsaufwand.

Ein dritter Nachteil besteht darin, dass Cross-References nur zu externen Einträgen (etwa *Verwandten Phraseologismen*) gelegt werden können, nicht aber zu anderen Feldern desselben Eintrags. Bei der Aufnahme mehrerer Äquivalente kann somit ein Übersetzungsdatensatz nicht auf einfache Weise mit dem Ausgangs- und Zielphraseologismus verknüpft werden.

Aus diesen Gründen scheint dieses Datenmodell nicht optimal; es sollen daher weitere Möglichkeiten untersucht werden.

6.3.3.2 B: Übersetzungsdaten auf Eintragsebene

Um das Problem zu vermeiden, dass das Termfeld im übersetzungsbezogenen Index künstlich gefüllt werden muss, wurden in einem zweiten Versuch die Übersetzungsdaten auf der Eintragsebene angesiedelt. Dies scheint insofern logisch, als es sich bei den Übersetzungsdaten um übergeordnete, sprachübergreifende Daten handelt. Diese werden der Grundidee von *MultiTerm* zufolge auf einer höheren Ebene angezeigt.

Abb. 21 zeigt einen Beispieleintrag nach diesem Modell. Auch hier können der Zielsprachliche Index sowie Übersetzungsdaten freigelassen werden, falls nur einsprachige Daten bekannt sind.

Entry level

Äquivalenz HU-DE: partiell

| | |
|-----------|--|
| Kommentar | eltérő kép; a német kifejezés egy kicsit választékosabb, mint a magyar; írt (pl. zsurnalisztikai) |
| HU: | szövegekben is használatos; arra is vonatkozhat, hogy vki/vmi eltűnik a nyilvánosság szeme elől |
| Kommentar | anderes Bild; der ung. Ausdruck wird nur umgangssprachlich verwendet und bezieht sich allgemein auf |
| DE: | plötzliches spurloses Verschwinden; nicht unbedingt auf das Verschwinden aus der Öffentlichkeit wie der deutsche Ausdruck. |

Übersetzungsbsp HU-DE: Diego néha ragyog, néha viszont eltűnik, mint szürke szamár a ködben

QUE AT: <http://www.sport-net.hu/sporthirek/keves-esely-van-diego-vb-s-szereplesere>

Zieltext: Manchmal glänzt Diego, manchmal verschwindet er komplett von der Bildfläche.

QUE ZT: HK

Äquivalenz DE-HU: partiell

| | |
|---------------|--|
| Kommentar DE: | anderes Bild, aber ähnlicher sprachlicher Stil |
| Kommentar HU: | eltérő kép, azonos stilisztikai szint |

Übersetzungsbsp DE-HU: Der Inhaber verschwand einfach von der Bildfläche, hinterließ ein Chaos.

QUE AT: <http://www.drachenlachen.com/index.php/sabine-richter.html>

Zieltext: A tulajdonos egyszerűen eltűnt, mint szürke szamár a ködben,
és káoszt hagyott maga mögött.

QUE ZT: HK

Hungarian

Term: eltűnik, mint szürke szamár a ködben

Definition: nyomatlanul eltűnik, teljesen nyoma vész

QUE Forgács, Tamás (32004): Magyar szólások és közmondások szótára. Mai nyelvünk állandósult
Def: szókapcsolatai példákkal szemléltetve. Budapest: Tinta Könykiadó.

SYN/ Paraphrase: nyomatlanul eltűnik

Variante: eltűnik, mint szamár a ködben

Stilschicht: ugs.

Regionen HU: Standard-Ungarisch

Fachgebiet 1: Allgemeinsprache

Kontextbeispiel: Engem az úgynevezett "emberi tényező" izgat. Igen, az az ember, aki a malörért felelős, de eltünt,
mint szürke szamár
a ködben. Méghozzá akkor tünt el, amikor még nem lehetett pontosan tudni, miként oldódik majd
meg a százharminc
utas sorsa.

QUE Kontext: Magyar Hírlap 2996. JAN. 9., 7; az idézetet közli: Forgács, Tamás (32004):

Magyar szólások és közmondások szótára. Mai nyelvünk állandósult
szókapcsolatai példákkal szemléltetve. Budapest: Tinta Könykiadó.

syntaktische Klasse: VP

Externe Leerstellen (obligatorisch): Subjekt

Kontextanbindung: durch Satzglied

Def DE: spurlos verschwinden, nicht auffindbar sein. (Wörtlich: "verschwinden wie ein grauer Esel im Nebel")

SYN/ Paraphrase DE: spurlos verschwinden

SYN/ Paraphrase DE: nicht auffindbar sein

Verwandte Phraseologismen DE: wie vom Erdboden verschluckt sein

Verwandte Phraseologismen DE: (völlig) von der Bildfläche verschwinden

German

Term: von der Bildfläche verschwinden

Definition: nicht mehr in der Öffentlichkeit zu sehen, spurlos verschwinden

QUE Def: <http://de.thefreedictionary.com/Bildfl%C3%A4che>, HK

SYN/ Paraphrase: wie vom Erdboden verschluckt sein

Variante: vollständig von der Bildfläche verschwinden

Variante: völlig von der Bildfläche verschwinden

Variante: komplett von der Bildfläche verschwinden

Verwandte Phraseologismen: nicht mehr im Rennen sein

Idiomatisität: Teil-Idiom, Kollokation

Stilschicht: ugs., Allgemeinsprache, journalistisch

Fachgebiet 1: Allgemeinsprache

syntaktische Klasse: VP

Externe Leerstellen (obligatorisch): Subjekt

Kontextanbindung: durch Satzglied

Def HU: hirtelen eltűnik a színről (sokszor a közéletről/ a nyilvánosság szeme elől), nyomatlanul eltűnik

QUE Def: HK

SYN/ Paraphrase HU: eltűnik a színről

SYN/ Paraphrase HU: eltűnik a nyilvánosság szeme elől

Abb. 21: Eintragsmodell B – Beispieleintrag

Auch bei diesem Datenmodell stellt sich die Frage, ob ein Eintrag mehrere Äquivalente oder nur ein Phraseologismenpaar enthalten soll.

Bei mehreren müssten unter *Äquivalenz* und *Übersetzungsbeispiel* Felder hängen, die Aufschluss geben, welche Phraseologismen gemeint sind, und zwar jeweils unter beiden Feldern, weil diese ja nicht wie beim Datenmodell A gebündelt sind, sondern einzeln auf der Eintragsebene liegen und allenfalls optisch gegliedert werden können. Dies würde nicht nur zu Unübersichtlichkeit, sondern auch zu einem ungleich höheren Aufwand bei der Dateneingabe führen, was sich bei mehr als zwei Sprachen natürlich noch verstärkt. Daher wird dieser Ansatz wieder verworfen.

Genau wie in Modell A enthält ein Eintrag somit nur ein Phraseologismenpaar in zwei verschiedenen Sprachen. Bezuglich der Mehrsprachigkeit der Datenbank ergibt sich daraus, dass zur Übersetzung eines vorhandenen Phraseologismus in eine neue Sprache ein bestehender Eintrag dupliziert wird, die alte Zielsprache im Duplikat gelöscht und eine neue hinzugefügt wird. Die Eintragsstruktur bleibt gleich.

Mit dem unkomplizierten Duplizieren von Einträgen in den Datenmodellen A und B geht ein kleiner Nachteil einher. Werden die duplizierten Daten später verändert, müssen sie auch in den Duplikaten aktualisiert werden³⁷. Wird dies nicht beachtet, besteht die Gefahr, dass sich widersprüchliche oder unvollständige parallele Einträge zu einem Phraseologismus in der Datenbank sammeln.

Die kontrastive Darstellung eines Phraseologismenpaares ist mit diesem Modell in übersichtlicher Weise möglich und mit weniger Aufwand verbunden als im Modell A.

6.3.3.3 C: Übersetzungsdaten auf Termebene

Schließlich soll noch ein drittes Datenmodell getestet werden, das von der ursprünglichen Idee, Übersetzungsdaten getrennt zu verwalten, am weitesten entfernt ist. In diesem Modell werden alle Übersetzungsdaten auf die Termebene gelegt. Zu diesem Zweck wird ein neues Feld *Übersetzung (Zielsprache)* auf der Termebene angelegt, das ein zielsprachliches Äquivalent enthält und dem die Felder *Äquivalenz* und *Übersetzungsbeispiel* untergeordnet werden.

Ein Eintrag enthält jeweils nur einen ausgangssprachlichen Phraseologismus, unter dem das Feld *Übersetzung ZS* mehrfach vorkommen kann, so dass für jeden ausgangssprachlichen Term mehrere zielsprachliche Äquivalente sichtbar sind. Abb. 22 zeigt einen ungarisch-deutschen Beispieleintrag mit mehreren Übersetzungen.

³⁷ Dies ist relativ unkompliziert möglich, indem alte Duplikate gelöscht und die neue Version erneut dupliziert wird. Die neuen Duplikate können mit früheren Übersetzungen zusammengeführt werden.

Entry level

Entry number: 4

Hungarian

Term: visszatér a régi kerékvágásba

Definition: rendeződik, normalizálódik, megszokott útjára tér

QUE Def: Magyar szólások és közmondások szótára

SYN/ Paraphrase: rendeződik

SYN/ Paraphrase: normalizálódik

SYN/ Paraphrase: megszokott útjára tér

Variante: beáll a régi kerékvágásba

Variante: beáll a rendes kerékvágásba

Variante: visszazökken a régi kerékvágásba

Variante: visszazökken a rendes kerékvágásba

Variante: visszatér a rendes kerékvágásba

Idiomatizität: Teil-Idiom

Stilschicht: Allgemeinsprache, journalistisch, mündlich, ugs.

Regionen HU: Standard-Ungarisch

Kontextbeispiel: Ha a tárgyalásoknak akár csak egy része realizálódna, bármikor visszaállhat az élet a rendes kerékvágásba.

QUE Kontext: Magyar Hírlap 1997. FEBR. 8., 11; az idézetet közli: Magyar szólások és közmondások szótára

Kontextbeispiel: Nekünk az a dolgunk? hogy fenntartsuk a rendet. Ha elmennénk, minden visszatérne a régi kerékvágásba. (MH 1994. DEC. 27.,8)

A Sojourner nevű kicsiny Mars-jármű kedden visszazökkent a rendes kerékvágásba - közöttük a pasadenai rányítóközpontból.

QUE Kontext: Magyar Hírlap 1997. JÚL. 17.,17; az idézetet közli: Magyar szólások és közmondások szótára

syntaktische Klasse: VP

Externe Leerstellen (obligatorisch): Subjekt

Restriktionen: häufig "a dolgok"

Kontextanbindung: durch Satzglied

Def Wörtlich: "ins alte Geleise zurückkehren". Nach einer Unterbrechung oder kurzfristigen Veränderung werden

DE: die alten Zustände wieder hergestellt. Dies kann sowohl im positiven Sinne gemeint sein, dass die alte
Ordnung wieder hergestellt ist, als auch im negativen Sinne, dass ein Änderungsversuch nichts bewirkt hat.

QUE Def: HK

Übersetzung DE: wieder [seinen/ ihren] gewohnten Lauf nehmen

Äquivalenz: **partiell**

Kommentar DE: Anderes Bild. Mit dem dt. Ausdruck ist im positiven Sinne gemeint, dass nach einer Turbulenz die alte Ordnung/ Gewohnheit wieder hergestellt wurde.

Kommentar HU: A német kifejezésnek pozitiv értelme van: Akkor használjuk, amikor egy rövidebb vagy hosszabb közjárat után visszatérünk a dolgok korábban megszokott rendjéhez.

Übersetzungsbeispiel: Alkonyat – Napfogyatkozás (2010), feliratos

Minden egy döntéssel kezdődik; Végre újra együtt Bella Swan (Kristen Stewart) és Edward Cullen (Robert Pattinson). A halandó Bella élete visszatér a régi kerékvágásba, minden olyan, mint mielőtt a vámpír család, Cullenék elhagyták volna Forks városát.

QUE AT: <http://divxmozi.com/alkonyat-napfogyatkozas-2010-feliratos/>

Zieltext: Eclipse – Bis(s) zum Abendrot (2010), OmU

Alles beginnt mit einer Entscheidung; endlich sind Bella Swan (Kristen Stewart) und Edward Cullen (Robert Pattinson) wieder vereint. Das Leben der sterblichen Bella nimmt wieder seinen gewohnten Lauf; alles ist wie früher, bevor die Vampirfamilie Cullen die Stadt Forks verlassen hatte.

QUE ZT: HK

Übersetzung DE: wieder ins alte Fahrwasser geraten

Äquivalenz: **partiell**

Kommentar DE: Mit dem deutschen Ausdruck ist im negativen Sinn gemeint, dass alte unerwünschte Zustände wieder eingekehrt sind, eine gewünschte Veränderung nicht erzielt wurde.

Kommentar HU: A német kifejezés jelentése negativ: Akkor használjuk, amikor egy kívánt változás nem sikerült (pl. a kábitószerről való leszokás), és visszatérünk a régi állapothoz.

Übersetzung DE: der Markt erholt sich

Äquivalenz: **partiell**

Kommentar DE: andere Satzgliedfunktion; diese Übersetzung ist nur im Zusammenhang mit dem Markt möglich.

Kommentar HU: Elterő mondatrészükben használatos. Ez a fordítás csak a piaccal kapcsolatos szövegkörnyezetben lehetséges.

Übersetzungsbeispiel: Visszatér a piac a régi kerékvágásba

QUE AT: <http://www.vg.hu/penzugi/hitel/visszatert-a-piac-a-regi-kerekvagasba-221742>

Zieltext: Der Markt hat sich wieder erholt/ Erholung am Markt

QUE ZT: HK

Übersetzung DE: sich normalisieren

Äquivalenz: **partiell**

Übersetzungsbeispiel: Hamarabb visszatérnek a régi kerékvágásba a Daimlernél

A Daimler a tervezetnél korábban véget vet a munkaidő-csökktéseknek a kereslet fellendülésének köszönhetően.

QUE AT: <http://www.napi.hu/m/default.asp?cCenter=article.asp&nID=444933>

Zieltext: Arbeitszeiten bei Daimler normalisieren sich früher als geplant

Durch einen Anstieg der Nachfrage beendet Daimler die Reduzierung der Arbeitszeiten früher als geplant.

QUE ZT: HK

Übersetzung DE: Alles ist wieder beim Alten

Äquivalenz: **partiell**

Kommentar DE: andere Satzgliedfunktion; auf Ungarisch ist das Substantiv zu aktualisieren.

Kommentar HU: Elterő mondatrész. A német kifejezés egy egész mondat, melynek az alanya "minden": "visszatér minden a régi kerékvágásba".

Abb. 22: Eintragsmodell C – Beispieleintrag mit mehreren ZS-Äquivalenten

Dadurch, dass pro Eintrag nur ein Phraseologismus in einer Sprache gespeichert wird, erübrigen sich Entscheidungen über begriffliche Zusammengehörigkeit und Abgrenzung.

Dieses Datenmodell hat im Gegensatz zu den vorherigen den Vorteil, dass mehrere Äquivalente mit Übersetzungsbeispielen auf einen Blick sichtbar sind, so dass man nicht mehrere Einträge durchforsten muss, bis sich eine passende Übersetzung

findet. Ist bei der Eingabe noch kein Äquivalent bekannt, werden die Übersetzungsfelder vorerst freigelassen.

Die zu den zielsprachlichen Äquivalenten gehörenden einsprachigen Daten werden jeweils in eigenen Einträgen gespeichert. Das Feld *Übersetzung (Zielsprache)* kann eine Cross-Reference zum zielsprachlichen Phraseologismus enthalten, oder aber nur Text, falls ein entsprechender Eintrag noch nicht existiert.

Bei der Aufnahme weiterer Sprachen enthält ein Eintrag folglich Übersetzungen in mehrere Sprachen. Abb. 23 zeigt, wie ein solcher Eintrag aussehen könnte. Die Übersichtlichkeit ist hier in keiner Weise eingeschränkt. Nicht relevante Sprachen können in der Ansicht minimiert oder durch Wahl eines entsprechenden Layouts ausgeblendet werden.

Entry level

Entry number: 53

German

Term: in Kraft treten

Definition: Gültigkeit oder Wirksamkeit erlangen

QUE Def: http://de.wiktionary.org/wiki/in_Kraft_treten

Idiomatizität: Kollokation

Stilschicht: Allgemeinsprache, Fachsprache, formal/ gehoben

Fachgebiet 1: Allgemeinsprache, Politik, Recht

Kontextbeispiel: Die Verordnung EZB/2001/13 ist daher dahin gehend zu berichtigten, dass sie am 1. Januar 2003 in Kraft tritt.

QUE Kontext: <http://eur-lex.europa.eu/Notice.do?mode=dbl&lang=de&ihmlang=de&lng1=de,hu&lng2=bg,cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,mt,nl,pl,pt,ro,sk,sl,sv,&val=273875:cs&page=>

typische Textsorten: Zeitungstexte, Nachrichten

Satzgliedfunktion: VP

Externe Leerstellen (obligatorisch): Subjekt

Restriktionen: idR Gesetz, Regelung o.ä.

Kontextanbindung: durch Satzglied

Übersetzung HU: hatályba lép

Äquivalenz: total

Übersetzungsbeispiel: Die Verordnung EZB/2001/13 ist daher dahin gehend zu berichtigten, dass sie am 1. Januar 2003 in Kraft tritt.

QUE AT: <http://eur-lex.europa.eu/Notice.do?mode=dbl&lang=de&ihmlang=de&lng1=de,hu&lng2=bg,cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,mt,nl,pl,pt,ro,sk,sl,sv,&val=273875:cs&page=>

Zieltext: Ezért helyesíteni kell az EKB/2001/13 rendeletet annak egyértelművé tételeire, hogy a rendelet 2003. január 1-jén lép hatályba.

QUE ZT: <http://eur-lex.europa.eu/Notice.do?mode=dbl&lang=de&ihmlang=de&lng1=de,hu&lng2=bg,cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,mt,nl,pl,pt,ro,sk,sl,sv,&val=273875:cs&page=>

Übersetzung PT: entrar em vigor

Äquivalenz: total

Übersetzungsbeispiel: Zwölf Monate nach ihrer Annahme im Rahmen einer multilateralen Konsultation treten die Änderungen in Kraft, es sei denn, dass ein Drittel der Vertragsparteien Einwände mitgeteilt hat.

QUE AT: <http://eur-lex.europa.eu/Notice.do?mode=dbl&lang=de&ihmlang=de&lng1=de,pt&lng2=bg,cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,mt,nl,pl,pt,ro,sk,sl,sv,&val=276750:cs&page=>

Zieltext: As alterações entrarão em vigor 12 meses após a sua adopção quando de uma consulta multilateral, a menos que um terço das partes tenha apresentado objecções.

QUE ZT: <http://eur-lex.europa.eu/Notice.do?mode=dbl&lang=de&ihmlang=de&lng1=de,pt&lng2=bg,cs,da,de,el,en,es,et,fi,fr,hu,it,lt,lv,mt,nl,pl,pt,ro,sk,sl,sv,&val=276750:cs&page=>

Abb. 23: Modell C - Eintrag mit mehreren Zielsprachen

Ein klarer Vorteil dieses Modells ist seine große Übersichtlichkeit. Eine Abfrage liefert mehrere Übersetzungsmöglichkeiten auf einen Blick, einsprachige Daten zu einem zielsprachlichen Phraseologismus sind mit einem Klick erreichbar. Auch für die Dateneingabe ist dieses Modell vorteilhaft: Häufig stellt sich erst im Übersetzungsprozess ein bestimmter Ausdruck als geeignetes Äquivalent heraus, bzw. es eröffnen sich mehrere Möglichkeiten der Formulierung in der Zielsprache. Phraseologische Einträge als Abstraktum entspringen also dem praktischen Kontext der Übersetzung. Die Möglichkeit, mehrere Ideen unmittelbar in einem Eintrag festzuhalten, lädt dazu ein, immenses Sprachwissen, das durch den kreativen Übersetzungsprozess zutage tritt, systematisch zu dokumentieren.

Ein Nachteil dieses Modells ist, dass es dazu verleitet, einen zielsprachlichen Phraseologismus nur manuell im Übersetzungsfeld einzutragen, ohne dafür einen eigenen Eintrag mit präzisen Informationen anzulegen. Im Prinzip wäre dies für den Anfang kein Problem, da in einem Kollaborationsprojekt genauere Informationen jederzeit nachgetragen werden können. Allerdings gibt es keinen Grund für eine Hinauszögerung der Dateneingabe, da eine Person, die aktiv in die Zielsprache übersetzt, vermutlich auch über nötiges anwendungsbezogenes Wissen über den zielsprachlichen Phraseologismus verfügt.

Zudem gibt es keine einfache Möglichkeit, noch fehlende einsprachige Einträge für bereits verzeichnete Übersetzungen auf einen Blick anzeigen zu lassen. Es besteht daher die Gefahr, dass das Fehlen zielsprachlicher Daten von Personen, die über entsprechendes Wissen verfügen, gar nicht bemerkt wird, und die Bestände zweier Sprachen bzw. Übersetzungsrichtungen sich völlig unabhängig voneinander entwickeln.

MultiTerm erlaubt zwar, Suchfilter danach zu kreieren, ob ein bestimmtes Feld eine Cross-Reference enthält, nicht aber danach, ob in einem bestimmten Feld noch *keine* Cross-Reference besteht, worauf es in diesem Fall ja ankommt. Daher lassen sich noch fehlende Bezüge zu zielsprachlichen Einheiten nur manuell ausfindig machen.

Ein Schwachpunkt liegt aber nicht nur im möglichen Fehlen von Einträgen in der jeweiligen Zielsprache, sondern auch darin, dass bestehende Einträge beim Legen von Cross-References möglicherweise nicht gefunden werden. Anders als die normale Suchfunktion in *MultiTerm*, die auch eine Fuzzy-Search, also eine unscharfe, fehlertolerante Suche beinhaltet, muss beim Legen einer Cross-Reference die genaue Zeichenfolge des Verweisziels eingegeben werden. Die Eingabe darf zwar verkürzt sein, muss aber mit dem Anfang des Zielphraseologismus übereinstimmen. Dies kann im phraseologischen Bereich zum Problem werden, da zur Notation von Nennformen unterschiedliche Möglichkeiten bestehen (s. 5.1.4.2), insbesondere, was die Notation externer Leerstellen, den Umgang mit Artikeln am Anfang (*(eine) heikle Angelegenheit*) sowie die Satzgliedfolge betrifft.

Eine weitere Schwachstelle betrifft die Logik der Eingabereihenfolge. Eine Cross-Reference kann erst gelegt werden, wenn ein entsprechender zielsprachlicher Eintrag existiert. Beim Eingeben einer neuen Übersetzungsbeziehung wäre es also am effizientesten, zuerst einen Eintrag des zielsprachlichen Phraseologismus anzulegen, und erst danach den des ausgangssprachlichen Phraseologismus, in dessen Feld *Übersetzung ZS* eine Cross-Reference gelegt wird. Diese Reihenfolge entspricht aber nicht der Intuition und erfordert Vorausplanung. Intuitiv würde man zunächst den ausgangssprachlichen Phraseologismus anlegen, danach eine oder mehrere Übersetzungsbeziehungen, und am Schluss noch genauere Informationen zu den zielsprachlichen Äquivalenten. Auf diese Weise muss man aber nach dem letzten Schritt noch einmal zum ausgangssprachlichen Eintrag zurückkehren, um dort die Cross-Reference zu legen. Dies ist technisch keinerlei Problem, vermindert aber das Usability-Niveau und somit die Attraktivität des Programms für Nutzer.

Ansprechend ist dieses Datenmodell trotz der beschriebenen Schwachstellen aufgrund seiner großen Übersichtlichkeit bei der Abfrage. Umsetzbar wäre es möglicherweise bei einer kleineren, homogenen Nutzergruppe, z.B. Übersetzer mit fundierten Kenntnissen beider Sprachen, die hauptsächlich an den Übersetzungsdaten interessiert sind. Für die Eingabe wäre das Aufstellen allgemeiner Regeln zur Vorgehensweise sinnvoll, die für jeden zugänglich sind.

6.3.4 Vergleich der drei Datenmodelle

Modell A liegt zwar am nächsten an der ursprünglichen Idee, Übersetzungsdaten gesondert zu speichern, ist aber in der Praxis am wenigsten brauchbar, weil die Datenbank erstens durch die Belegung von Sprachindices mit Übersetzungsdaten nur eingeschränkt erweiterbar ist, und zweitens die Notwendigkeit, das Termfeld im Übersetzungsdatensatz zu füllen, einen erhöhten Aufwand bei der Eingabe bedeutet.

Die Modelle B und C sind gleichermaßen brauchbar und weisen unterschiedliche Vor- und Nachteile auf: Modell B liefert eine umfassende Gegenüberstellung zweier Phraseologismen mit Informationen in beiden Sprachen, dafür sind mehrere Übersetzungsvarianten hier nicht auf einen Blick erkennbar, sondern es müssen zum Vergleich mehrere gleichlautende ausgangssprachliche Einträge einzeln angeklickt werden.

Modell C bietet einen schnellen Überblick über mehrere Übersetzungsvarianten, hat aber den Nachteil, dass zu den jeweiligen Äquivalenten möglicherweise nicht zeitgleich genügend einsprachige Informationen eingetragen werden, oder aber, dass ein einsprachiger Eintrag des Äquivalents zwar schon existiert, beim Versuch, eine Cross-Reference zu legen, aber nicht erkannt wird.

Für die Datenabfrage ist Modell C zweifellos am angenehmsten, da mehrere Übersetzungsmöglichkeiten auf einen Blick verglichen werden können. Für die Eingabe ist C zwar günstig, wenn gleich mehrere zielsprachliche Äquivalente eingegeben werden, das Festhalten einsprachiger Daten des Zielphraseologismus ist hier aber komplizierter als im Modell B, weil dafür ein eigener Eintrag angelegt werden muss.

Welches der Modelle B und C sich besser eignet, müsste in einer Pilotphase getestet werden. Unabhängig vom gewählten Datenmodell weist *MultiTerm* auch generelle Schwächen bei der Aufzeichnung phraseologischer Daten auf. Wo allgemeine Stärken und Schwächen von *MultiTerm* für die Phraseographie liegen, wird in den folgenden Absätzen besprochen.

6.4 Vorteile von *MultiTerm*

Auch wenn die in 5.6 zusammengestellten Idealbedingungen für eine phraseographische Datenbank nicht in vollem Umfang erfüllt sind, weist *MultiTerm* zweifelsohne einige Vorteile auf.

Eine wesentliche Stärke des Programms besteht in der **flexiblen Datenbankmodellierung**, die dem Nutzer bei der Definition der Felder, deren Eigenschaften und der hierarchischen Anordnung völlig freie Hand lässt. Dies ermöglicht die Schaffung einer differenzierten Datenbankstruktur, wie sie in 5.8 gefordert wurde, und die Aufnahme umfassender gebrauchsbezogener Angaben zum Phraseologismus.

Ebenso ist die **Zuordnung mehrerer Äquivalente** zu einem Phraseologismus sowie eine **separate Speicherung übersetzungsbezogener Daten** möglich, wobei die konkrete Umsetzung dieser beiden Faktoren von der Wahl des jeweiligen Datenmodells abhängt. Sowohl Modell B als auch Modell C liefern eine übersichtliche Strukturierung der Information; Schwächen ergeben sich womöglich hinsichtlich der Intuitivität und Effizienz der Eingabe, nicht aber bezüglich der Logik des Aufbaus.

Ein weiterer wichtiger Faktor besteht in der Möglichkeit, sowohl für Anzeige als auch Bearbeitung je nach Kompetenz und Interesse **spezifische Benutzerprofile** zu erstellen. Administratorinnen können unterschiedliche Benutzerrollen definieren, für die jeweils spezifische Rechte für unterschiedliche Bereiche der Datenbank festgelegt werden können. Mögliche Rechte sind Lesezugriff, Schreibzugriff, sowie die Berechtigung, Einträge anzulegen bzw. sie zu löschen. Die Anzeige kann durch das Anlegen unterschiedlicher Layouts auf spezifische Benutzerinteressen zugeschnitten werden.

Auch die Möglichkeit, **Cross-References** zwischen Einträgen zu legen, stellt einen Vorteil von *MultiTerm* für die Phraseographie dar.

6.5 Grenzen von *MultiTerm*

Die Besprechung der drei Eintragsmodelle hat gezeigt, dass *MultiTerm* auch Grenzen hat, auf die man auch bei optimaler Strukturierung der Daten zwangsläufig stößt.

Durch die Möglichkeit der flexiblen Datenbankmodellierung lassen sich phraseologische Inhalte in *MultiTerm* gut organisieren; diesbezüglich gibt es nur kleinere Schwächen.

Ein wesentlicher Nachteil ist jedoch, dass sich *MultiTerm* nur in eingeschränktem Maße für Online-Kollaboration nutzen lässt, wie sie in 5.5.2 vorgeschlagen wurde. Hierbei stößt man zum einen auf systembedingte Hindernisse, die mit den Möglichkeiten des Programms selbst zusammenhängen, und zum anderen auf praktische Hindernisse, die sich aus äußeren Rahmenbedingungen wie Kosten und technischer Umsetzbarkeit ergeben.

Ein wesentliches Manko bezüglich der Funktionalitäten der Software selbst ist, dass sich neben den Eingabefeldern **keine Erläuterungen** einfügen lassen, die, wie in 5.5.5 festgestellt wurde, zur Gewährleistung der Qualität der Daten dringend erforderlich wären. Erstens lässt sich in manchen Fällen allein durch die Feldbezeichnung nicht eindeutig darstellen, welche Information genau gefragt ist, bzw. worin die Abgrenzung zu anderen Feldern besteht. Zweitens kann die Eingabe intuitiver und leichter verständlich werden, wenn der gewünschte Feldinhalt durch ein Beispiel veranschaulicht wird. Drittens können in der Erläuterung formale Standards vorgeschrieben werden, etwa die Notation der Nennform betreffend, die User sonst datenbankextern irgendwo nachschlagen müssen. Insbesondere in folgenden Fällen besteht Erläuterungsbedarf:

- Notation der Nennform;
- was als *Suchbegriff* eingegeben werden soll;
- Unterscheidung zwischen *SYN/Paraphrase*, *Variante* und *verwandte Phraseologismen*; insbesondere müsste hier erklärt werden, dass Paraphrasen derselben syntaktischen Klasse angehören müssen wie der Eintrag, was als Variante und was als eigener Phraseologismus gilt und welche Phraseologismen als verwandt betrachtet werden;
- Bestimmung der syntaktischen Klasse;
- was mit Restriktionen und transformationellen Defekten gemeint ist;
- dass bei *fakultativen externen Leerstellen* nur typische Ergänzungen genannt werden sollen;

Aufgrund der hohen Komplexität der Materie ist es nicht möglich, Feldbezeichnungen so zu wählen, dass sie intuitiv verständlich sind. Eine vereinfachende Zusammenfassung von Feldern hingegen ist im Sinne der Granularität (s. 5.1) nicht ratsam.

In 6.3.4 wurde festgestellt, dass sowohl bei Datenmodell B als auch bei Datenmodell C die Aufstellung von **Regeln** erforderlich ist, um eine konsistente Führung der Datenbank zu gewährleisten. Bei Modell B muss vereinbart werden, dass bei Änderungen eines einsprachigen Eintrags auch eventuell bestehende Duplikate entsprechend geändert werden. Bei Modell C muss daran erinnert werden, beim Eintragen eines zielsprachlichen Äquivalents auch einen entsprechenden einsprachigen Eintrag in der Zielsprache anzulegen.

Solche Regeln müssen datenbankextern für User einsehbar fixiert werden. Die Einhaltung dieser Regeln erfordert einen zusätzlichen Aufwand und verhindert – ebenso wie das Fehlen von Erläuterungen – eine intuitive Dateneingabe. Dadurch erhöht sich einerseits die Fehlerwahrscheinlichkeit, und andererseits sinkt mit steigender Komplexität die Motivation zur Teilnahme, so dass voraussichtlich mehr User mangels Zeit oder technischem Verständnis „abspringen“.

Dies ließe sich durch eine höhere **Automatisierung von Routineprozessen** vermeiden, was bei *MultiTerm* nicht möglich ist. In Modell B sollten Duplikate bei Änderungen automatisch wechselseitig synchronisiert werden, oder aber von vornherein technisch als ein Eintrag gespeichert sein, der unterschiedliche übersetzungsbezogene Verknüpfungen enthält. In Modell C wäre es hilfreich, wenn beim Legen einer Cross-Reference eine unscharfe Suche zum Auffinden des zielsprachlichen Äquivalents möglich wäre. Außerdem könnte eine ideale Benutzeroberfläche den User nach der Eingabe eines Phraseologismus samt Übersetzung direkt zur einsprachigen Eingabemaske des zielsprachlichen Äquivalents weiterleiten, so dass das Anlegen dieses Eintrags nicht vergessen wird.

Diese Faktoren machen *MultiTerm* nur eingeschränkt für Web-Kollaboration brauchbar, die nur einem begrenzten, interessierten Nutzerkreis zugänglich ist. Regeln, Normen und Erläuterungen zur Eingabe müssten datenbankextern verzeichnet werden, auch Support kann erforderlich werden. Dadurch wäre die Durchführung eines solchen Projekts mit *MultiTerm* mit einem erhöhten administrativen Aufwand verbunden.

Ein „Schönheitsfehler“ von *MultiTerm* gemessen an dem in Kapitel 5 entworfenen Ideal einer phraseographischen Datenbank besteht im **Fehlen eines Statistik-Tools** (s. 5.3.3.1) zur User-Befragung über Gebräuchlichkeit, Varianten, areale Verbreitung, Stilschicht etc. Immerhin verfügt *MultiTerm* über Andockstellen für weitere Programmtools, so dass sich hier für die Serverbetreiberin eine Möglichkeit ergeben könnte, ein eigenes Statistik-Tool zu integrieren.

Abgesehen von systembedingten Grenzen des Programms stehen auch **praktische Rahmenbedingungen** einer kollaborativen Nutzung im Wege.

Diese bestehen unter anderem in den **hohen Anschaffungs- und Lizenzkosten für den Serverbetreiber**. Für User ist die Nutzung kostenlos, es sei denn, der Serverbetreiber erhebt Gebühren für die Nutzung. Eine Erhebung von Gebühren hätte jedoch den Nachteil, dass sich die Teilnehmerzahl reduzieren würde, was unter Umständen zu einem weniger umfassenden Bild des phraseologischen Sprachgebrauchs führt.

Ein weiteres Problem kann sich dadurch ergeben, dass durch die Lizenzbestimmungen eine **maximale Anzahl von Benutzern** festgesetzt ist, die gleichzeitig eine Verbindung zu *MultiTerm Server* herstellen und Terminologiedaten bearbeiten können. Am Zentrum für Translationswissenschaft der Universität Wien sind derzeit 25 gleichzeitige Schreibzugriffe lizenziert. Für ein breit angelegtes phraseographisches Projekt kann eine solche Höchstzahl ein einschränkender Faktor sein.

Zudem haben Erfahrungen an der Universität Wien gezeigt, dass bei der kollaborativen Nutzung über die Online-Webschnittstelle *MultiTerm Online* häufig **technische Probleme** auftreten. Derzeit ist die Universität Wien der einzige Serverbetreiber, der *MultiTerm Server* über *MultiTerm Online* verwendet.

Außer den beschriebenen Schwierigkeiten im Hinblick auf eine kollaborative Nutzung von *MultiTerm* wurden weitere Nachteile beobachtet, unter anderem die fehlende Möglichkeit einer expliziten **Beschlagwortung**. Suchbegriffe, die in ein hierfür angelegtes Feld eingetragen wurden, können daher nur über eine Volltextsuche gefunden werden.

Ein weiteres kleineres Problem stellt die Tatsache dar, dass beim Anlegen eines **Index nur Sprachen** zur Auswahl stehen, die bei Bedarf umbenannt werden. Damit gibt es für die Gesamtzahl möglicher Indexfelder für spätere Erweiterungen eine Obergrenze. Unter anderem aus diesem Grund wurde Modell A verworfen, bei dem Übersetzungsdaten in Indexfeldern gespeichert werden. Dieser Punkt ist jedoch nicht gravierend, da andere Lösungen für die Eintragsstruktur gefunden wurden.

7 Zusammenfassung und Ausblick

Die vorliegende Arbeit befasste sich mit der Konzeption eines multilingualen elektronischen phraseographischen Wörterbuchs, wobei zunächst ein konzeptuelles Datenmodell auf theoretischer Basis entworfen wurde, das anschließend mit Hilfe der Software *MultiTerm* praktisch umgesetzt wurde.

Als theoretische Fundierung wurden eingangs Grundbegriffe aus der einsprachigen und kontrastiven Phraseologieforschung dargestellt und reflektiert. Kernpunkt dabei war es, die Vorteile einer weiten Begriffsauffassung aufzuzeigen, die auch Kollokationen sowie stark variable Phraseologismen einschließt (s. 2.2.1, 2.2.3). Dies entspricht dem Usus in der maschinellen Sprachverarbeitung und trägt den sprachproduktionsbezogenen Bedürfnissen des Spracherwerbs Rechnung. Aus kontrastiver Sicht wurden Phraseologismen daher als Mehrworteinheiten definiert, die sich nicht allein durch Anwendung des grammatischen Regelwerks und des lexikalischen Inventars einer Sprache aktiv bilden lassen, sondern zusätzliches Vorwissen erfordern (s. 4.2). Insbesondere sind für die aktive Verwendung von Phraseologismen funktionale Angaben wie Stilschicht, Textsortenbezogenheit, areale Verbreitung etc. von großer Bedeutung. Solche Angaben sind in gedruckten Wörterbüchern häufig unterrepräsentiert (s. 4.4).

Bezüglich der Übersetzung von Phraseologismen wurde einerseits die Bedeutung des Translationsskopos, andererseits die des umgebenden Textes betont (s. 4.1). Dies führt häufig dazu, dass ein nur partielles Äquivalent die beste Übersetzung im jeweiligen Kontext darstellt. Für ein Wörterbuch wurde daher die Möglichkeit gefordert, einem Eintrag mehrere, u.a. partielle Äquivalente zuzuordnen, versehen mit Übersetzungsbispieln und Kommentaren zur Äquivalenz (s. 4.3).

Auf Basis dieser Erkenntnisse wurde ein konzeptueller Entwurf für eine mehrsprachige phraseologische Datenbank erstellt, die als aktives Wörterbuch in mehreren Übersetzungsrichtungen funktionieren und die Prinzipien der Granularität und Elementarität berücksichtigen sollte (s. 5.1). Das Konzept basierte auf zwei zentralen Elementen: Erstens werden einsprachige und übersetzungsbezogene Daten getrennt angelegt, um die einsprachigen Bestände auch mit weiteren Sprachen verknüpfen zu können. Zweitens werden Phraseologismen als Ganzes in einem Eintrag gespeichert, der diverse gebrauchsbezogene Angaben sowie Verknüpfungen mit verwandten Einträgen und Äquivalenten enthält (s. 5.2/3).

Nach der Besprechung einzelner ein- und mehrsprachiger Datenkategorien wurden Möglichkeiten der Contenterstellung und -bearbeitung diskutiert (s. 5.5). Es wurde vorgeschlagen, einen Kernbestand aus Korpora zu extrahieren, der manuell weiter bearbeitet und ergänzt wird. Diesbezüglich wurde für Online-Kollaboration plädiert, die

die Verteilung der Arbeit sowie Aktualität und Repräsentativität der Daten ermöglicht (s. 5.5.2). Dabei können auch User einbezogen werden, und zwar entweder durch interaktive Bewertung von Phraseologismen hinsichtlich Gebräuchlichkeit und Stil, oder in Form von direkter Dateneingabe. Wesentlich für den Erfolg eines kollaborativen Ansatzes ist eine nach Usability-Kriterien gestaltete Benutzeroberfläche, was insbesondere durch ein schrittweises Anzeigen von Information sowie Direkthilfen und Erläuterungen begleitend zur Eingabe erreicht wird (s. 5.5.5).

Basierend auf den Ausführungen in Kapitel 5 wurde abschließend ein Kriterienkatalog für ein elektronisches phraseographisches Wörterbuch erstellt (s. 5.8).

In Kapitel 6 wurde die praktische Umsetzung der entworfenen Datenbankstruktur mit der Terminologieverwaltungssoftware *MultiTerm* dokumentiert. Nach einer allgemeinen Beschreibung der Software wurden drei verschiedene Eingabemodelle entworfen und anhand eines ungarisch-deutschen Beispielkorpus getestet. Die wesentliche Herausforderung hierbei bestand in der Unterbringung der übersetzungsbezogenen Angaben.

Im Modell A erfolgte dies über je ein eigenes Indexfeld für beide Übersetzungsrichtungen, die mit den zu vergleichenden Phraseologismen in einen Eintrag zusammengefasst wurden. Dieses Modell erwies sich aus diversen Gründen als unpraktikabel und wurde wieder verworfen (s. 6.3.3.1).

Die Modelle B und C waren gleichermaßen brauchbar, wobei beide unterschiedliche Vor- und Nachteile aufwiesen. Bei Modell B wurden die Übersetzungsdaten auf die Eintragsebene gelegt, also je einem Phraseologismenpaar übergeordnet. Dadurch ergab sich ein übersichtlicher kontrastiver Vergleich zwischen zwei Phraseologismen. Modell C enthielt nur einen Phraseologismus in einer Sprache, dem auf der Termebene verschiedene zielsprachliche Äquivalente zugeordnet wurden, versehen mit weiteren Unterfeldern zur Äquivalenz. Dies ergab eine übersichtliche Darstellung verschiedener Übersetzungsmöglichkeiten auf einen Blick, die bei B fehlte. C hingegen lieferte keine direkte kontrastive Gegenüberstellung (s. 6.3.3.2, 6.3.3.1).

Beide Modelle enthielten mögliche Fehlerquellen, die durch die Aufstellung expliziter Regeln zur Eingabe verhindert werden müssen, um eine konsistente Führung der Termbank zu ermöglichen (s. 6.4). Dadurch macht die Mitarbeit eine gewisse theoretische Einarbeitung erforderlich, was sich einschränkend auf die Usability und damit auf die Attraktivität der Software auswirkt.

Diese Faktoren führten zu dem Ergebnis, dass *MultiTerm* zwar zur strukturierten Darstellung von phraseologischem Inhalt gut geeignet ist, in puncto Benutzerfreundlichkeit und intuitiver Eingabe jedoch Mängel aufweist, die die Möglichkeit einer kollaborativen phraseographischen Nutzung einschränken. Weitere Hürden bestehen in der fehlenden Möglichkeit der Beschlagwortung, in hohen Anschaffungskosten für den

Serverbetreiber, begrenzte Möglichkeit des gleichzeitigen Serverzugriffs sowie in technischen Problemen, die an der Universität Wien häufig auftraten (s. 6.5)

Dies wirft die Frage auf, ob überhaupt ein bestehendes System geeignet ist, den in Kapitel 5 beschriebenen Desiderata in vollem Umfang zu entsprechen. Vermutlich wäre eine Neuprogrammierung sinnvoller, die technisch durchaus umsetzbar scheint.

Die Umsetzung eines solchen Projekts erfordert die Einbeziehung von Wissen aus unterschiedlichen Disziplinen, insbesondere der linguistischen Phraseologieforschung, der Translationswissenschaft, Lexikographie und Terminologielehre, der Computerlinguistik, v.a. der computergestützten Lexikographie und deren Anwendung in sprachverarbeitenden Systemen wie maschineller Übersetzung, sowie weiteren Bereichen der Informatik, z.B. Datenbankmodellierung und Usability.

Hierbei liefert die Phraseologieforschung die theoretischen Grundlagen, die Informatik dagegen die formale Repräsentation des Wissens und die Möglichkeit dessen praktischer Nutzbarmachung. Die Translationswissenschaft setzt die Standards bezüglich der Übersetzungsqualität und notwendiger kontextbezogener Angaben, die das Wörterbuch zu einem attraktiven translatorischen Werkzeug machen.

Somit bedeutet die multilinguale elektronische Phraseographie einen interdisziplinären Brückenschlag, von dem alle Seiten profitieren können, z.B. durch Vereinheitlichung von Terminologie und lexikalischer Repräsentation. Projekte wie XMELLT, das Möglichkeiten der Standardisierung und Austauschbarkeit polylexikalischer mehrsprachiger Lexika für Sprachtechnologie untersuchte (vgl. Calzolari et al. 2002; Ide 2000), weisen bereits in diese Richtung.

Die bewusste Einbeziehung von Nutzern, die Schaffung einer allgemeinverständlichen Oberfläche stellt einen weiteren wichtigen Baustein in diesem Mosaik dar: Anwenderinnen wie Übersetzerinnen, Sprachlernende und -didaktinnen würden nicht nur von diesem Projekt profitieren, sondern könnten als Mitwirkende einen wesentlichen Beitrag leisten, nämlich die Inhalte selbst. Das häufig nur über „Sprachgefühl“ zugängliche Wissen kann über ein kollaboratives phraseographisches Wörterbuch einem breiten Nutzerkreis zugänglich gemacht werden.

8 Literatur- und Quellenverzeichnis

8.1 Fachliteratur

- ARNTZ, REINER (2001): Fachbezogene Mehrsprachigkeit in Recht und Technik. Hildesheim, Zürich, New York: Georg Olms Verlag.
- ARNTZ, REINER/ PICTH, HERIBERT/ MAYER, FELIX (2004): *Einführung in die Terminologiekarbeit*. Hildesheim, Zürich, New York: Georg Olms Verlag.
- BUDIN, GERHARD (1990): Terminological Analysis of LSP Phraseology. In: Terminology, Science & Research. Vol. 1 (1990), no. 1-2, Wien.
- BUDIN, GERHARD/ GALINSKI, CHRISTIAN (1992): Übersetzungsorientierte Phraseologieverwaltung in Terminologiedatenbanken. In: terminologie et traduction. No. 2/3, 1992, 565-574.
- BUBENHOFER, NOAH/ PTASHNYK, STEFANIYA (2010): Korpora, Datenbanken und das Web: State of the Art computergestützter Forschung in der Phraseologie und Lexikographie. In: Ptashnyk et al (Hg.) (2010), 7-19.
- BURGER, HARALD (2003): Phraseologie. Eine Einführung am Beispiel des Deutschen. Berlin.
- BURGER, HARALD/ BUHOFER, ANNELIES/ SIALM, AMBROS (1982): *Handbuch der Phraseologie*. Berlin, New York: De Gruyter.
- BURGER, HARALD/ DOBROVOL'SKIJ, DIMITRI/ KÜHN, PETER/ NORRICK, NEAL R. (Hg.) (2007): *Phraseologie. Ein internationales Handbuch der zeitgenössischen Forschung/ Phraseology. An International Handbook of Contemporary Research*. Berlin, New York: De Gruyter.
- CALZOLARI, NICOLETTA/ FILLMORE, CHARLES/ GRISHMAN, RALPH/ IDE, NANCY/ LENCI, ALESSANDRO/ MACLEOD, CATHERINE/ ZAMPOLLI, ANTONIO (2002): Towards Best Practice for Multiword Expressions in Computational Lexicons. In: *Proceedings of the Third International Conference on Language Resources and Evaluation* (LREC 2002), S. 1934-40, Las Palmas. URL (10. Okt. 2010) <http://gandalf.aksis.uib.no/lrec2002/pdf/259.pdf>
- CARSTENSEN, KAI-UWE/ EBERT, CHRISTIAN/ EBERT, CORNELIA/ JEKAT, SUSANNE/ KLABUNDE, RALF/ LANGER, HAGEN (Hg.) (2010): *Computerlinguistik und Sprachtechnologie. Eine Einführung*. Heidelberg: Spektrum Akademischer Verlag.
- COWIE, ANTHONY PAUL (1998): Introduction. In: Cowie, Anthony Paul (Hg.) *Phraseology: Theory, Analysis and Application*. Oxford: Oxford University Press, 1-20.
- DRÄGER, MARCEL/ JUSKA-BACHER, BRITTA (2010): Online-Datenerhebung im Dienste der Phraseographie. In: Ptashnyk et. al. (Hg.) (2010), 165-178.
- DOBROVOL'SKIJ, DMITRIJ (1999): Kontrastive Phraseologie in Theorie und Wörterbuch. In: Baur, Ruprecht S./ Chlosta, Christoph/ Piirainen, Elisabeth (Hg.) (1999):

- Wörter in Bildern, Bilder in Wörtern. Beiträge zur Phraseologie aus dem Westfälischen Arbeitskreis.* Essen: Schneider Verlag Hohengehren GmbH.
- DUHME, MICHAEL (1991): *Phraseologie der deutschen Wirtschaftssprache*. Eine empirische Untersuchung zur Verwendung von Phraseologismen in journalistischen Fachtexten. Essen: Die Blaue Eule.
- ETTINGER, STEFAN (1989): Einige Probleme der lexikographischen Darstellung idiomatischer Einheiten (Französisch-Deutsch). In: Gréciano (Hg.) (1989), 95-115.
- FLEISCHER, WOLFGANG (2¹⁹⁹⁷): *Phraseologie der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- GLÄSER, ROSEMARIE (1986): *Phraseologie der englischen Sprache*. Leipzig, Wien: Langenscheidt.
- GÖPFERICH, SUSANNE (2²⁰⁰³): Text, Textsorte, Texttyp. In: Snell-Hornby et al. (Hg.) (2²⁰⁰³), 104-107.
- GRÉCIANO, GERTRUD (Hg.) (1989): *Europhras 88. Phraseologie Contrastive. Actes du Colloque International*. Strasbourg: Klingenthal.
- HEID, ULRICH (2007): Computational linguistic aspects of phraseology II. In: Burger et al. (Hg.) (2007), 1036-1044.
- HUBER, EVELYN (1993): *Übersetzungsorientierte Phraseologieverwaltung mit Terminologieverwaltungssystemen*. Wien, Zentrum für Translationswissenschaft, Diplomarbeit.
- HUTCHINS, JOHN W./ SOMERS, HAROLD L. (1992): *An Introduction to Machine Translation*. London, San Diego: Academic Press.
- IDE, NANCY (2000): XMELLT. Cross-Lingual Multi-word Expression Lexicons for Language Technology. URL (10. Okt. 2010):
<http://www.cs.vassar.edu/~ide/transparencies/xmellt.ppt>
- ISENBERG, HORST (1978): Probleme der Texttypologie. Variation und Determination von Texttypen. In: *Wiss. Zeitschrift der Karl-Marx-Universität Leipzig* 27, 565-579.
- KEMPER, ALFONS / EICKLER, ANDRÉ (2006): *Datenbanksysteme*. München: Oldenbourg Wissenschaftsverlag.
- KJÆR, ANNE LISE (1990): Context-Conditioned Word Combinations in Legal Language. In: Terminology, Science & Research. Vol. 1 (1990), no. 1-2, Wien.
- KOLLER, WERNER (7²⁰⁰⁴): *Einführung in die Übersetzungswissenschaft*. Wiebelsheim: Quelle & Meyer Verlag.
- KOLLER, WERNER (2007): Probleme der Übersetzung von Phrasemen. In: Burger et al. (Hg.) (2007), 605-613.
- LUBENSKY, SOPHIA/ McSHANE, MARJORIE (2007): Bilingual phraseological dictionaries. In: Burger et al. (Hg.) (2007), 919-928.
- LYONS, JOHN (8¹⁹⁹⁵): *Einführung in die moderne Linguistik*. München: C.H. Beck.
- MEL'ČUK, IGOR (2007): Lexical Functions. In: Burger et al. (2007), 119-131.

- MEL'ČUK, IGOR (1998): The Meaning-Text Approach to the Study of Natural Language and Linguistic Functional Models. In: Embleton, S.: LACUS Forum 24, Chapel Hill: LACUS, 3-20 : URL (10. Okt. 2010):
<http://olst.ling.umontreal.ca/melcuk>.
- MOON, ROSAMUND (2007): Phraseology in general monolingual dictionaries. In: Burger et al. (Hg.) (2007), 909-918.
- NIELSEN, JAKOB/ MACK, ROBERT L. (1994): *Usability Inspection Methods*. New York, Chichester, Brisbane, Toronto, Singapur: John Wiley & Sons.
- PELZ, HEIDRUN (³2005): *Linguistik. Eine Einführung*. Hamburg: Hoffmann und Campe.
- PTASHNYK, STEFANIYA/ HALLSTEINSDÓTTIR, ERLA/ BUBENHOFER, NOAH (Hg.) (2010): Korpora, Web und Datenbanken. Computergestützte Methoden in der modernen Phraseologie und Lexikographie/ Corpora, Web and Databases. Computer-Based Methods in Modern Phraseology and Lexicography. Baltmannsweiler: Schneider Verlag Hohengehren.
- SAILER, MANFRED (2007): Corpus linguistic approaches with German corpora. In: Burger et al. (Hg.) (2007), 1060-1071.
- SCHREIBER, MICHAEL (²2003): Übersetzungstypen und Übersetzungsverfahren. In: Snell-Hornby et al. (Hg.) (²2003), 151-154.
- SNELL-HORNBY, MARY/ HÖNIG, HANS G./ KUSSMAUL, PAUL/ SCHMITT, PETER A. (Hg.) (²2003): *Handbuch Translation*. Tübingen: Stauffen.
- STANTCHEVA, DIANA (2002): *Phraseologismen in deutschen Wörterbüchern. Ein Beitrag zur Geschichte der lexikographischen Behandlung von Phraseologismen im allgemeinen einsprachigen Wörterbuch von Adelung bis zur Gegenwart*. Berlin: Verlag Dr. Kovač.
- VEISBERGS, ANDREJS (1996): Idiom Transformation, Idiom Translation and Idiom Dictionaries. In: Zettersten, Arne/ Hjørnager Pedersen, Viggo (Hg.): *Symposium on Lexicography VII. Proceedings of the Seventh Symposium on Lexicography May 5–6, 1994 at the University of Copenhaben*. Tübingen: Max Niemeyer Verlag, 241-246.
- WOTJAK, GERD (1982): Äquivalenz, Entsprechungstypen und Techniken der Übersetzung. In: Jäger, Gert/ Neubert, Albrecht (Hg.) (1982): *Äquivalenz bei der Translation*. Leipzig: Verlag Enzyklopädie.

8.2 Wörterbücher

- PUSZTAI, FERENC (2006): Magyar Értelmező Kéziszótár. Budapest: Akadémiai Kiado.
- FORGÁCS, TAMÁS (³2004): Magyar szólások és közmondások szótára. Mai nyelvünk állandósult szókapcsolatai példákkal szemléltetve. Budapest: Tinta Könyvkiadó.

8.3 Internetquellen

(Stand jeweils 10. Oktober 2010)

Termbanken des Zentrums für Translationswissenschaft, Universität Wien:

<http://mtsrv.trans.univie.ac.at:8090/TermbaseFinderApp/>,

Beispielquellen:

<http://books.google.de/books?id=XTBbrOvXIxMC&pg=PA7>

http://de.wikipedia.org/wiki/Ich_bin_ein_Berliner

http://de.wikipedia.org/wiki/Liste_gefl%C3%BCgelter_Worte/J

http://www.buecherhallen.de/aw/hoebu/community_hoebu/blog/~jfm/%5BMediennews

[%5D_Lena_kam_sang_und_siegte!/_](#)

http://www.enotes.com/topic/Double_entendre

<http://wissen.de/wde/generator/wissen/ressorts/geschichte/index,page=1178282.html>

10 Abbildungsverzeichnis

| | |
|---|-----|
| Abb. 1: Basisklassifikation nach Burger (2003) | 25 |
| Abb. 2: Phraseologische Klassen im Begriffsfeld | 27 |
| Abb. 3: Struktur eines phraseologischen Wörterbuchs..... | 48 |
| Abb. 4: Areale Untergliederung des deutschsprachigen Raums | 54 |
| Abb. 5: Vollständige Eintragsstruktur des phraseologischen Wörterbuchs | 61 |
| Abb. 6: ER-Diagramm | 62 |
| Abb. 7: Feldtypen in der Eingabemaske | 66 |
| Abb. 8: Eintragsstruktur bei <i>MultiTerm</i> | 80 |
| Abb. 9: Klassischer Eintrag in <i>MultiTerm</i> | 81 |
| Abb. 10: Termbankdefinition – Indexfelder | 83 |
| Abb. 11: Termbankdefinition – beschreibende Felder | 84 |
| Abb. 12: Termbankdefinition – Eintragsstruktur..... | 84 |
| Abb. 13: Eintragsmodell – hierarchische Struktur..... | 85 |
| Abb. 14: Benutzeroberfläche von <i>MultiTerm</i> | 87 |
| Abb. 15: Setzen eines Filters | 89 |
| Abb. 16: Beispieleintrag mit minimierten Feldern | 90 |
| Abb. 17: Eintragsmodell A - Beispieleintrag..... | 96 |
| Abb. 18: Umbenennung von Indexfeldern | 98 |
| Abb. 19: Einsprachige und übersetzungsbezogene Felder in der Termbankdefinition | 99 |
| Abb. 20: Mehrere Äquivalente in einem Eintrag | 100 |
| Abb. 21: Eintragsmodell B – Beispieleintrag..... | 106 |
| Abb. 22: Eintragsmodell C – Beispieleintrag mit mehreren ZS-Äquivalenten | 109 |
| Abb. 23: Modell C - Eintrag mit mehreren Zielsprachen | 110 |

11 Tabellenverzeichnis

| | |
|---|----|
| Tabelle 1: Interne und externe Valenz | 21 |
| Tabelle 2: Strukturelle Abweichungen | 37 |
| Tabelle 3: Lang- und Kurzform von Phraseologismen..... | 52 |
| Tabelle 4: Datenkategorien mit Eingabemodus | 68 |
| Tabelle 5: Übersicht über das Korpus mit quantitativer Verteilung relevanter Kriterien | 76 |
| Tabelle 6: Einträge der Picklist-Felder | 92 |
| Tabelle 7: Einsprachige Daten mit Eingabemodus und Inhalt..... | 93 |
| Tabelle 8: Übersetzungsbezogene Datenfelder..... | 95 |

12 Anhang A: Korpus HU-DE

| Phraseologismus HU | Def HU | Kontext HU | Phraseologismus DU | Aquivalenz | Kommentar zur Äquivalenz DE/HU |
|-----------------------------------|---|---|--|-------------|---|
| [vminek] fele se tréta | 'nem lehet vicceini vele, komolyan kell venni a dolgot, ill. komolyra fordul a dolgoz' (MSZKSZ) | "Radon - ennek már a fele se tréfa [...] A WHO új általása szerint a lakásokban megengedhető radon sugárzás az eddigig max 1000 Bq/m ³ helyett csak 100 Bq/m ³ . A vizsgálatok azt mutatják, hogy a radon sokkal veszélyesebb, mint gondolták. Minden 100 Bq/m ³ 16%-al emeli a tüdőrák kockázatát." (http://psycho.blog.hu/2010/03/05/radon_csaik_a_fele_se_tréfa) | Das ist eine ernste Angelegenheit/ das ist ein ernstes Problem | partiell | |
| [vminek] fele se tréta | " | " | Damit ist nicht zu spaßen | pragmatisch | Leicht abweichende Struktur, aber gleiches Bild und gleiches stilistisches Niveau, Enyhén eltérő struktúra, de azonos kép és stílusúval színvonal |
| [vkinék] nyomdokain lépdeni! [vk] | "azt az ultat, pályát követi, amelyen az illété jár" (MSZKSZ) | "Möllermann könnyével láthatóan Sálamon Béla nyomdokain lépezi." (http://www2.mno.hu/portal/134206) | in jds Fußstapfen treten | total | |
| nyomdokainba lép | "Baj v. botrány várható (nhol)" (MEKSZ) | "Valami bűzlik a gázszámíkákat" (http://www.blikk.hu/blikk_k_aktualis/v-alami-buzlik-a-gazszamlakkal-2025864) | Etwas ist faul an etw/ Hier/ Daran ist etwas faul. | partiell | anderes Bild/ eltérő kép |
| [vmi] bűzlik itt/ vimbiben | " | " | etw erhitzt die Gemüter | partiell | gleiche Bedeutung, anderes Bild/ azonos jelentés, eltérő kép |
| [vmi] felkorlácsolja a kedélyeket | [vmi] (HK) | "1956 elején az SZKP XX. Kongresszusa a személyi kultusz nyilvánítása érte a magyar politikai fejlesztési tervet a kedélyeket a kedélyeket a magyar politikai élében." (http://www.okeonga.hu/kek_daru/2006/08/keidaru_2006_08.pdf) | die Stimmung einheizen | partiell | |
| [vmi] felpezsdti a Kedélyeket | jövedvre derít, derültséget okoz (Dr. Dóra Solti) | " | für gute Laune sorgen | partiell | |
| [vmi] felpezsdti a kedélyeket | "emberi közösségek, terménynek, terméknél a legérfehérkesebb fésze." (MSZKSZ) | "A zenészek színe-java" (MSZKSZ) - "Az oceán kincse - a tenger gyümölcsseinek színe-java" (http://www.tohuuhulkonyvismeretterjeszto_iszakarskonyvek/az_ocean_kincsei_a_tenger_gyumoltseinek_szine_java#bovebben) | (die) Spalten- [+ Gen.Attr.] | partiell | der bestimmte Artikel wird bei nachgestelltem Genitivattribut benötigt: "Ungarns Spitzensportler", aber "die Spitzensportler des Landes" / A határozott névvel a fölvet követő birtokos jelző esetén szükséges: "Ungarns Spitzensportler", de "die Spitzensportler des Landes". |

| Phraseologismus HU | Def HU | Kontext HU | Phraseologismus DU die erstensten [NP] | Aquivalenz | Kommentar zur Äquivalenz DE/HU |
|--|--|--|--|--|---|
| [vml] színe-java | " | " | " | partiell | erfassen häufig im Zusammenhang mit kulinarischen Gerüßen (Wien). Nicht bei Personen. Az 'erlesen' szót sozor a kulináris speciálitásokkal kapcsolatban használjuk (pl. bor). Nem vonatkozhat személyekre. |
| [vml] színe-java | " | " | die Crème-de-la-crème + Gen. | partiell | |
| [vml] színe-java | " | " | die Elite + Gen. | partiell | |
| [vml] színe-java | " | " | die Spitze + Gen./ das Beste von... | partiell | |
| a sótét középkor | "A humanisták számára a Nyugatrómai Birodalom bukása és a saját török kozsát elítélt ezer év a sótétség és a kulturális hanyatlás időszaka volt. A kozérők minden szempontból negativ minősítést azonban csak a 18. századtan, a felvilágosodás idején kapott. Ekkor vált általánossá a sótét kozérők megnevezésük." | "A Los Alamos Nemzeti Laboratórium egy vulkanológusa szolgálinak tartja, hogy a sótét kozépkor, a szellemi sótétség korszaka valódi sótétségeivel kezdődött." (http://web.interware.hu/valasz/valashirek/archidarkages.htm) | das finstere Mittelalter | total | |
| annak idején | "1. abban az időben, régebben" (MÉKSZ) | | zu jener Zeit, damals; | partiell | Nur auf Vergangenheit bezogen - entspricht der 1. Bedeutung von 'annak idején' a multipla vonatkozó - megfelel a magyar kifejezés 1. |
| annak idején | 2. "ha majd idősebbű lesz" (MÉKSZ) | | dann (Zukunft) | partiell | Lebenszeit auf die Zukunft bezogen (wenn es sowohl ist/ der Zeitpunkt gekommen ist ein bestimmtes Ereignis eingetreten ist), entspricht der 2. Bedeutung des Phraseologismus a jövőre vonatkozó - megfelel a magyar kifejezés 2. |
| á-tól cettig/ á-tól z-ig | "elejétől végeig, kihétel nélkül; teljesen" | "Az Eco által írt vizsgált renékritévek nem befjezetlen 'felkesztemelék', amelynek az olvasó annyian 'társzerzője', hanem ától cettig nagyon is ravaaszú kitundító ópuszok (1), amelyekben a 'minitaszerző' -az irodalmi szövegter mérnő pontossággal negaszerző stratéga - a textusra rejt el a 'miniakvásiók' szánt utasításait, hogy miként olvassa a művet." | von A bis Z | total | |
| átvitt értelmemben | "szóirokosaltnak jelenésváltozással, kiakadt értelemszintűségi, a szamár szónak "buta" jelentése)" (MÉKSZ) | "Magical. Jelenése: miszikus, csodálatos, izgalmas. Néha helyettesít az ic- és alakot, különösen ha átvitt értelemben akarunk megnyilvánulni valamiről." (http://www.angoltanszek.hu/UTtitut.php?tid=1858) | im übertragenen Sinn, im übertragenen Sinn | total | |
| "Az érték multibeli akaklása nem jelent talékoztatójában található, a 2001/10/07/EK irányelv által előírt figyelemzetiéles (HK) következetes álap értelembenének jóvábeli akaklására nézve." | egy berétezési álap egyszerűsített varianciában. "A multibeli teljesítmény nem tekintethető irányadónak a jóvábeni teljesítményre vonatkozóan. A multibeli teljesítmény nem jelent útmutatást a nyilvános teljesítményhez, és előfordulhat, hogy az nem ismétlődik meg." | "Die Wertentwicklung der Vergangenheit lässt keine verlässlichen Rückschlüsse auf die zukünftige Wertentwicklung eines Kapitalanlagefonds zu." | pragmatisch | "Die Wertentwicklung der Vergangenheit lässt keine verlässlichen Rückschlüsse auf die zukünftige Wertentwicklung eines Kapitalanlagefonds zu." | |
| leszáll az alkonyat, belealkonyodik (HK) | http://epa.oszk.hu/0160/01659/00019/pdf/vadasz_es_vegesyjav_EP_A01659_1958_02_19_07_0.pdf | die Dämmerung bricht herein | total | anderes Verb, aber pragmatisch völlig äquivalent, más igé, de teljes ekvivalencia a pragmatika szintén. | |

| Phraseologismus HU | Def HU | Kontext HU | Phraseologismus DU | Aquivalenz | Kommentar zur Aquivalenz DE/HU |
|--|---|---|--|-------------|--|
| csődejárás jelent | a csődejárás az ados által kezdeményezett megfizetés (HK) | "az addos általában azért jelent csődöt, hogy megegyezze a hitelezők által kezdeményezett felelősséget a kölcsönök leírást folytán)" http://www.perzugysziget.hu/index.php?option=com_content&view=articledetail&id=249&Itemid=85) | Konkurs anmelden | total | |
| eltűnik, mint a szürke szanár a körben | "nyomattalanul eltűnik, teljesen nyoma vesz" (MSZKSZ) | "Engem az 'úgynevezett' embert ténylező igazat [gen, az az ember, aki a ma köördít felelős, de eltűnt, mint szürke szanár a körben. Még ohozzá akkor tűnt el, amikor még nem lehette pontosan tudni, miként oldódik majd meg a százhaminc útsorsa.]" (Magyar Hirzap 2966. JAN. 9., 7; az idézetet közli: MSZKSZ) | spurios verschwinden, (völlig) von der Bildfläche verschwinden, wie vom Erdoden verschluckt sein | partiell | unterschiedliches Bild/ eltérő kép |
| eltűnik, mint a szürke szanár a körben | "dühös lesz" (http://nevarchivum.klte.hu/szeng/legyeb/gyakori.htm) | "Ha valami nem az akarata szerint történik, könnnyen felkajta a vizet, és mérgében nem tudja, mit cselekszik. Amilyira dühös tud lenni, hogy száma húmagazastrolorosztan híml" | (völlig) von der Bildfläche verschwinden, sich aufregen | partiell | nur auf Ung. phraseologisch/ csak magyarul frazeologizmus |
| fizetést teljesít | | "Az akkreditív olyan fizetési mód, amelynek által a nyitó bank arra vállal kötelezettséget, hogy a vevő megbízása alapján az előírt feltételek teljesítése és megfelelő okmányok benyújtása esetén fizesse teljesít az eladónak." (http://www.hanwhabank.hu/huvalatal/nemzetkozi-fizetesek) | eine Zahlung leisten | total | |
| gazdasági fellendülés | "gyorsabb gazdasági fejlődés" (MÉKSZ) | "Túl vagyunk a válságom. Lassú gazdasági fellendülés vélte kezdetét!" http://ec.europa.eu/news/economy/00225_hu.htm#top | wirtschaftlicher Aufschwung | total | |
| hatályba lép | érvényes lesz (HK) | "Az 1994. március 1-jén hatályba lépett 26.5-ös deférészögötő háromszögöt hívhattunk alapján azt a kulföldi, aki nem tudja magát igazolni, vagy illagálisan tanúsíték Magyarországon, vagy tiltott hatalárispésten érikt. "Kötelező tanítóködési helyre" utalnak." (Magyar Hirlap 1994. SZEP 1. 16., 4. az. részleteti közli: MSZKSZ) | in Kraft treten | total | |
| hegyszögg | "0° és 90° közé eső szög (tárgasság értelmezőben)" (http://www.kislexikon.hu/hegyszagg.htm) | "Segítségül bármilyen 26.5-ös deférészögötő háromszögöt hívhattunk vona, mert a szövegi szemközti betegő és az átfogó aránya a hasonlóság miatt állandó. Ez a hanyados a hegyszögg szimuszának nevezzük." (http://weboarchive.googleusercontent.com/search?q=cache:ptStfb03f0a/www.sulinet.hu/tanar/kompetenciateuletek/2._matematika/3._moduljerasok-tanar-tanulo-eszközök/2._tipus/10-evi/10/3._tanari_segedanyagok__prezentacio_konferenciak_0/mod_nagyesszögek_szogfiguravenei_1-definicioek_ppt-hegyszessz%C3%BBg&cd=1&n=hu&ct=elnk&gl=hu) | ein spitzer Winkel | total | nur auf Deutsch phraseologisch/ csak a német nyelvben alkot frazeologial szókapcsolatot. |
| hitelbe | | "A magyaroknak mincs bérzük, de költségei alaknának. [...] Mindezent veszünk hitelbe." (http://mgv.hu/gazdasag/2008/02/19_cetelem_megfizartas_hitel_kotterezes) | auf Kredit | total | |
| nem látna a fatói az erdőt | "elvezzik a részleteiben, ezért nem látna át a dolog lényegét." (MSZKSZ) | "De milózben erőf és másról támog leg mindent fol kell kutatunk, nem lenne szabad elmerülni kicsinyes vitáiban, s föleg látni kellene a fatól az erdőt." (MH 1996. DEC. 28., 7. Az idézetet közli: MSZKSZ) | vor lauter Bäumen den Wald nicht mehr sehen | pragmatisch | Leicht abweichendes Bild: Auf Ung. ist man so auf einen Baum konzentriert, dass man darüber hinaus den Wald nicht mehr sieht. Auf Dt. sieht man aufgrund der Kép: Magyarul annyi koncentráltuk egyetlen fára, hogy man nem láttuk töle az erdész erődt, nemetű pedig a fák sokasága miatt már nem vagyunk képesek az egész kép befogadására. |

| Phraseologismus HU | Def HU | Kontext HU | Phraseologismus DU | Aquivalenz | Kommentar zur Äquivalenz DE/HU |
|--|--|---|-------------------------------------|------------|---|
| nyáron | "Ha sokan intézkednek v. sokan szólnak bele egy dolguba, az csak karos lehet." (MSZKSZ) | "Magyarországon manapság hat ingatlanos szakmai szövefesig működik, s minden arra figyelemzeti, hogy a sok hába közelvész a gyermek. Ily módon ma Magyarországon senhol sincs ily széges nyilvántartás arról, kik a valóságos ingatlanok közvetítői." (HVG 1997. MÁRC. 22., 69; az időzetet röszli: MSZKSZ) | Viele Köche verderben den Brei. | total | Nur nyár und téi werden mit dem Suffix -N gebildet; dagegen: tavasszal , összel/ gleiche Bedeutung, anderes Bild azonos jelentes, eltérő kép |
| Sok bába között elvész a gyerek. | "nagy felhalás vmi apróság miatt vagy vmi érdemtelennel dologoz" (MSZKSZ) | "S hogy a jelenlén még impozánsabbnak hosszan, a rendőrkétfőn fényses napai légiógrázásnak a budapesti Váci utcáról. Nem kizárt, hogy minden csak sok húrnó semmírt, mint angyal a lakossági fejelmelek sem töbök optikai csalódásánál. Vájában ugyanis nem a közönszövetség rossz, hanem a lakosság közönszövetsége - általják sokan." (HVG 1998. FEBR. 21., 9-12) | Viel Lärm um nichts | partiell | "Lärm" steht hier idiomatisch für Aufuhn. Aufregung. Das ungarische "húrnó" hingegen ist eine unikale Komponente mit ebendieser Bedeutung, insofem nicht idiomatisch. A német "Lärm" "Zauber" jelent. Itt átvitt értelmemben használjuk (fehatalás). Teljes ekvivalencia a pragmatika szintjén. |
| Sok húrnó semmíerti Nagy húrnó semmíerti | "szabad emberként, nem letartóztatva." | "A sajátos Jübürokat jelek azt is megmutatják a februárban másfél évnnyi börtönre tiltott – ám még szabadabron leülő – expolgármesternek, merre ellenne elindulna a kecskeméti börtönbe." http://www.balon.hu/bacs-kiskun/kozeletben/nem-mindenki-kudelene-bortonne-a-polgarmester-318463 | auf freiem Fuß | total | nur auf Deutsch phraselogisch csak a német nyelven alkot frazeologial szókörrel. |
| szabadiabon | "szabadtéri, nem letartóztatva." | "Százötven éve hirdették ki a szepüpülelen fogantás dogmáját" (http://magazin.ujember.catolikus.hu/Archivum/2005_01/09.html) | unbefleckte Empfangnis | total | |
| szepüpülelen fogantatás | "A szepüpülelen fogantatás egy katolikus hitteli, amely szerint Szűz Mária már édesanya, Szent Anna mehénben szepüpüleleni, vagyis eredendő (másképpen áterős) bün nélküli fogant. (Tehát nem azonos azza! a hittelivel, hogy Mária Jézust szüzen foganta és szülte.)" http://hu.wikipedia.org/wiki/Szepüp%C5%91l_1_teli_1_fogantat%C3%A1s | "Ha a tárgyalásoknak akár csak egy része realizálódna, bármikor visszaállna az élet a rendes kerékvágásba." (MH 1997. FEBR. 8., 11) "Nekünk az a dogma? hogy termántsuk a rendet. Ha elinnaménk, minden visszatérne a régi kerékvágásba." (MH 1994. DEC. 27., 8) "A Sojourner nevű kicsiny Mari-jármű visszazárkózik a rendes kerékvágásba - kozonkira a pasadéni rányitóközpontjiba.." (MH 1997. JÚL. 17., 17.) Az idezeteiket közli: MSZKSZ | wieder seinen gewohnten Lauf nehmen | | Andere Bild. Mit dem dt. Ausdruck ist im positiven Sinn gemeint, dass nach einer Turbulenz die alte Ordnung/Gewohtheit wieder hergestellt wurde. A német kifejezés jelentse negatív. Akkor használjuk, amikor egy rövididebő vagy hosszabb közjárat után visszatérünk a dolgok korábban megszokott rendjéhez. |
| visszatér a régi kerékvágásba beall/visszazökkön a (rendes/ régi) kerékvágásba | "rendeződik, megszokott újjára ter" (MSZKSZ) | "Ha a tárgyalásoknak akár csak egy része realizálódna, bármikor visszaállna az élet a rendes kerékvágásba." (MH 1997. FEBR. 8., 11) "Nekünk az a dogma? hogy termántsuk a rendet. Ha elinnaménk, minden visszatérne a régi kerékvágásba." (MH 1994. DEC. 27., 8) "A Sojourner nevű kicsiny Mari-jármű visszazárkózik a rendes kerékvágásba - kozonkira a pasadéni rányitóközpontjiba.." (MH 1997. JÚL. 17., 17.) Az idezeteiket közli: MSZKSZ | wieder ins alte Fahnwasser geraten | | Mit dem deutschen Ausdruck ist im negativen Sinn gemeint, dass alte unerwünschte Zustände wieder eingekreist sind, eine gewünschte Veränderung nicht erzielt wurde. A német kifejezés jelentse negatív. Akkor használjuk, amikor egy kívánt váltózás nem sikerült (pl. a kabítószerről) |
| visszatér a régi kerékvágásba beall/visszazökkön a (rendes/ régi) kerékvágásba | " | " | " | | |

Quellen.

- HK die Verfasserin
 Puszai Ferenc (2006): Magyar Értelmező Kéziszótár. Budapest: Akadémiai Kiadó.
 MÉKSZ Magyar Hirap
 MH Fongács, Tamás (2004): Magyar szóások és közmondások szóára. Mai nyelvünk állandósult szóköröcsölati példákkal szemléltetve. Budapest: Tinta
 NSZKSZ

Lebenslauf

Geboren 1979 in Lahr/ Schwarzwald, Deutschland

Ausbildung

- 2007-2010 Übersetzungswissenschaft (Deutsch/Portugiesisch/Ungarisch), Universität Wien (MA)
- 2004-2007 Übersetzen und Dolmetschen (Deutsch/Portugiesisch/Ungarisch), freies Wahlfach Physik, Universität Wien (Bakk.phil.)
- 1999-2004 Germanistik und Romanistik (Portugiesisch), Albert-Ludwigs-Universität Freiburg und Universidade Nova de Lisboa

Fachbezogene Berufstätigkeit

- 2007-2009 Wissenschaftliche Mitarbeiterin am Zentrum für Translationswissenschaft, Universität Wien: Erstellung eines englisch-deutschen Thesaurus über Flüchtlingshilfe in Zusammenarbeit mit dem Österr. Roten Kreuz
- 2006-2009 Tutorin am Zentrum für Translationswissenschaft, Universität Wien: Präsenztutorium zur VO *Medienkompetenz* (professioneller Einsatz von MS Office), mitverantwortlich für eLearning und elektronisches Anmeldesystem
- 2006 Dolmetscheinsatz für den brasilianischen Autor Paulo Lins am Unabhängigen Literaturhaus Niederösterreich, Krems

Mitgliedschaften

- EUROPHRAS Europäische Gesellschaft für Phraseologie
- UNIVERSITAS AUSTRIA Berufsverband für Dolmetschen und Übersetzen

Kontakt: helenakempf@gmx.at