

Martin Šemelík/Věra Marková/Marie Vachková (Prag)

Parallelkorpora in universitärer Lehre und Forschung. Erste Erfahrungen mit dem deutsch-tschechischen InterCorp¹

1 Zur Einführung

Der vorliegende Text setzt sich zum Ziel, das deutsch-tschechische Parallelkorpus InterCorp vorzustellen (s.u. 4), indem einige Aspekte der parallelkorpusbasierten Wortschatz- und Grammatikarbeit im fortgeschrittenen DaF-Unterricht im Rahmen des Germanistikstudiums gestreift werden. Anschließend wird auf konkrete Beispiele zur Anwendung des Parallelkorpus InterCorp im fortgeschrittenen DaF-Unterricht eingegangen.

In diesem Beitrag wird die *direkte* Anwendung der Sprachkorpora² im Sprachunterricht in Bezug auf die Sprachlehre, -erforschung und -erlernung (*das daten-drivene Lernen*, weiter nur DDL)³ fokussiert. Die *indirekte* Anwendung von Sprachkorpora (etwa deren Einfluss auf die Bestimmung der Lerninhalte und -ziele oder deren mögliche Rolle bei der Gestaltung von Wörterbüchern, Grammatiken und Lehrwerken u.Ä.) wird in Kap. 5.3.3 gestreift. Die Impulse hierfür kommen primär von der korpusorientierten Wortschatz- und Grammatikforschung am Institut für germanische Studien⁴ und am Institut für die Übersetzungswissenschaft der Philosophischen Fakultät der Karls-Universität in Prag.⁵ Allgemeine Fragen der Korpuslinguistik können in diesem Beitrag wegen dessen eingeschränkten Umfangs nur in Auswahl behandelt werden (vgl. Kap. 3).

¹ Dieser Text ist ein Output der Arbeiten am Projekt des Großen Akademischen Wörterbuchs Deutsch-Tschechisch (GADWT), das am Institut für germanische Studien der Philosophischen Fakultät der Karls-Universität in Prag unter Leitung von M. Vachková entsteht (GAČR 405/09/1280). Zur Projektbeschreibung vgl. Vachková 2011 oder <http://lexarchiv.ff.cuni.cz/lexikograficka-sekce/> (Stand 17.6.2011).

² Zur Unterscheidung zwischen direkter und indirekter Anwendung der Sprachkorpora im Sprachunterricht siehe Leech (1997) oder z.B. Römer (2009).

³ Engl. „data-driven learning“ (der Begriff in Johns 1986 geprägt – vgl. Römer 2008: 118).

⁴ Vgl. <http://german.ff.cuni.cz/pmwiki/pmwiki.php/Main/HomePage>

⁵ Vgl. <http://utrlf.ff.cuni.cz/UTRLFF-238.html>

2 Tschechische Germanistik aus der DaF-Perspektive: Status quo

Um den Stand der tschechischen Germanistik beschreiben zu können, muss man auch die aktuelle Situation des Deutschen in der Tschechischen Republik allgemein in Betracht ziehen,⁶ wie auch die neue Position des Germanistikstudiums an den tschechischen Hochschulen. Die neuen Studienprogramme stellen eine Herausforderung an die Lehrer dar: Die Linguistik ist wegen der Pragmatisierung des Bakkalaureatstudiums nicht aufzugeben, sondern, wie bereits angedeutet,⁷ auf eine neue Art und Weise zu vermitteln. Die Rolle der elektronischen Sprachkorpora ist heutzutage nicht mehr in Frage zu stellen; das Interesse der Studierenden an philologischer Sprachbetrachtung kann durch bedachtes Heranziehen moderner korpusanalytischer Methoden entfaltet werden, wie z.B. bei Johansson (2009) kritisch erörtert wird:

Corpora are important in basic research, and they have a role to play in the classroom as well. But let's not exaggerate. Corpora are no replacement for natural communication. They cannot replace the teacher. And, finally, they should not be used in language teaching just because we now have this wonderful tool and would like to apply it in language teaching as well. Their use is vindicated to the extent that it agrees with what we know about language and language acquisition, and can be shown to be an effective learning tool.⁸

⁶ Diese kennzeichnet sich durch gegensätzliche Entwicklungstendenzen. Die Anzahl der Deutschlerner an tschechischen Schulen sinkt: Während im Jahre 1996/1997 an den tschechischen Grundschulen 341 586 Schüler Englisch und 374 502 Schüler Deutsch lernten, waren um zehn Jahre später 571 077 Englischlerner und 148 101 Deutschlerner registriert (vgl. Dovalil 2007: 161 und Vachková 2010, s. ihren Bericht unter der Adresse <http://www.ids-mannheim.de/org/gremien/intwrat.html>; Stand: 17. 6. 2011). Das Deutsche bleibt jedoch immer die zweitwichtigste Fremdsprache. Die Tatsache, dass die Tschechische Republik an zwei deutschsprachige Nachbarländer grenzt, scheint kein großer Stimulus zu sein, obwohl seitens der deutschsprachigen Firmen in Tschechien eine konstant starke Nachfrage nach Arbeitnehmern mit guten Deutschkenntnissen besteht, und obwohl sich mit guten Deutschkenntnissen die Tür nicht nur zu Auslandsstipendien, sondern auch zur Beförderung bzw. Gehaltserhöhung öffnen kann. Die Position des Deutschen wird sich im Hinblick auf die engen ökonomischen Kontakte in absehbarer Zukunft wohl nicht dramatisch ändern. Zudem gilt, dass die Erhaltung der *Vielfalt des kulturellen Lebens in Europa* durch vermehrte Kenntnis nationaler und regionaler Sprachen zu den wichtigsten sprachpolitischen Zielen des Europarats gehört.

⁷ Vachková/Marková/Belica (2008).

⁸ Johansson (2009: 42).

3 (Parallel)Korpora im Fremdsprachenunterricht

Die massive Verbreitung der neuen Informationstechnologien, die sich auch die Sprachwissenschaft zunutze gemacht hat, hat u.A. die Entstehung großer elektronischer Sprachkorpora zur empirischen Erforschung der Sprache(n) ermöglicht. Für den DaF-Unterricht drängen sich jedoch Fragen auf, die ebenso die heranwachsenden Germanisten als auch deren Lehrer bewegen und um die Anwendung der rohen Daten im Unterricht kreisen. Im Folgenden soll deswegen mindestens auf ausgewählte Gesichtspunkte der Anwendung von Korpora im Fremdsprachenunterricht eingegangen werden:⁹

- (a) Wie kann man aus einer endlichen Menge von Sprachverwendungsdaten Konsequenzen für ein Sprachsystem ziehen?
- (b) Für welche Zwecke kann/soll das Korpus verwendet werden?
- (c) Welche Lerntypen profitieren am meisten von der Korpusarbeit? Welche lernpsychologischen Gesichtspunkte sind bei der Korpusarbeit noch zu berücksichtigen?
- (d) Hat der DDL-Ansatz Konsequenzen für das Rollenverständnis der Lehrer/Schüler? Wie sind diese zu beurteilen?

Ad (a)

Die Frage, inwieweit man die durch die Analyse von Korpusdaten gewonnenen Erkenntnisse auf den Sprachausschnitt übertragen kann, den das Korpus repräsentiert, entspringt der Tatsache, dass die Grundgesamtheit, über die man etwas aussagen möchte (die Sprache), nicht abgeschlossen ist, und deshalb nicht präzise definiert werden kann.¹⁰ – Hier gilt: Je größer, repräsentativer und ausgeglichener ein Korpus zur Verfügung steht, desto verlässlichere Ergebnisse der jeweiligen Korpusrecherchen können erwartet werden. Zugleich soll aber nicht vergessen werden, dass Korpusbefunde nur positiv interpretiert werden können:

Wenn ein Phänomen in einem Korpus belegt ist, dann existiert es auch. Wenn ein Phänomen in einem Korpus nicht belegt ist, kann man keine Aussage folgern. [...] In diesem Fall kann es zwei Erklärungen geben: Entweder existiert das Phänomen tatsächlich nicht, oder das Korpus ist in dieser Hinsicht unvollständig.¹¹

Soll die Akzeptabilität einer im Korpus nicht belegten „Wortkombination“ beurteilt werden, kann sich die Befragung von Muttersprachlern als hilfreich erweisen, obwohl sie aber nicht überschätzt werden darf, denn die Sprachkenntnisse eines Einzelnen sind vis-a-vis dem Usus notwendigerweise eingeschränkt

⁹ Die Fragen (a)–(d) werden hier getrennt behandelt, doch sie hängen eng miteinander zusammen.

¹⁰ Lemnitzer/Zinsmeister (2009: 50).

¹¹ Perkuhn/Belica (2006: 4).

und stark subjektiv geprägt. Man muss sich dessen bewusst werden, dass Muttersprachler häufig von ihrem Heimatdialekt beeinflusst sind oder die verkrusteten Sprachvorstellungen konservativer Sprachnormierer unkritisch akzeptieren. Der Vorwurf, dass die Korpora manchmal Unsinn enthalten würden, ist hier mit Perkuhn/Belica wie folgt zu kommentieren:

Eine technisch einwandfreie und authentische Abbildung vorausgesetzt, enthalten Korpora nur Fakten über den Sprachgebrauch. Sie erfassen grammatisch korrekten und falschen Umgang mit der Sprache, so wie ihn die abgebildete Sprachgemeinschaft praktiziert. Darüber hinaus findet sich sicher auch kreativer Umgang mit der Sprache, der von Grammatikschreibern nicht antizipiert werden kann. [...] Es steht den korpusinterpretierenden Linguisten natürlich frei, die Daten entsprechend zu kategorisieren. Als „Quatsch“ können die Daten nur in Hinsicht auf eine bestimmte Anwendung und somit als Folge einer Interpretation gewertet werden.¹²

Dabei ist aber bisher ungeklärt geblieben, wie viele Belege auf eine Regularität hindeuten, die man bei einer linguistischen Beschreibung zu berücksichtigen hat.¹³

Ad (b)

Mit McEnergy/Wilson kann die Rolle der Korpora im Sprachunterricht folgendermaßen zusammengefasst werden:

The corpus encourages the student to act as the producer of research, rather than its passive receptacle. [...] In being presented with corpus materials, students learn by a confrontation with data. [...] a student is immediately forced into active participation in the process of learning. [...] The teacher still has a place in the corpus-based classroom, but the shift of emphasis away from the didactic pedagogue towards the fellow researcher [...] is emphasized by inclusion of the corpus in the teaching equation.¹⁴

Andere DDL-Ansätze sehen den Studierenden weniger als Forscher, sondern eher als Reisenden auf einer Entdeckungsfahrt und betonen die Bedeutung des entdeckenden Lernens. Bernardini schreibt:

[...] learners browse corpora much in the same way as they would explore an unknown land. The teacher accompanies them as a guide, not telling them what to do but advising them on how to pursue their own interests, suggesting alternative ways to proceed, other interpretation of the data or possible ways forward. Interestingly, the role of instructor or oracle does not shift from teacher to corpus: the corpus is not expected to provide the right answers, and often does not, but constantly presents new challenges and stimulates new questions, renewing the user's curiosity and

¹² Perkuhn/Belica (2006: 5).

¹³ Lemnitzer/Zinsmeister (2009: 56).

¹⁴ McEnergy/Wilson (1997: 6f.).

offering ample opportunity for researching aspects of language and culture, which may easily become a subject of research projects, reports and discussion.¹⁵

Ein besonderer Reiz der Parallelkorpora besteht darin, dass „bilingual corpus [...] provides real world contrastive evidence of how utterances or samples of text in one language can be rendered in another, according to a given context, and of how words are used appropriately in combination“,¹⁶ womit die Parallelkorpora eine der Schwächen beheben, an der traditionelle Übersetzungswörterbücher leiden.¹⁷ Während die v.a. vom Beginn bis etwa zu den 60er-Jahren des 20. Jh. entwickelten methodischen Verfahren des Fremdsprachenunterrichts (v.a. direkte Methode, audiolinguale und audiovisuelle Methode) das Übersetzen als Methode aus dem Fremdsprachenunterricht verbannt haben, ist das Übersetzen später im Rahmen oft unterschiedlicher Vermittlungsmethoden wieder rehabilitiert worden – nicht als Methode, sondern als eine mögliche Lern- und/oder Übungsform,¹⁸ der im Fremdsprachenunterricht auch heutzutage ein fester Platz gebührt. – Verstärkt gilt dies für ein universitär verankertes philologisches Sprachstudium, das auf reflektierte Sprachbeherrschung abzielen hat.¹⁹ Für welche konkreten Zwecke die Korpora verwendet werden sollen, hängt mit den unter (a) besprochenen Phänomenen zusammen. So ist es beispielsweise nicht zu ermitteln, wie groß der Anteil der Fragesätze an allen Sätzen der deutschen Sprache ist. – Es liegt hier das Problem der Repräsentativität einer Stichprobe im Verhältnis zur Grundgesamtheit vor: Dieses Verhältnis kann nicht in einem Korpus genau widerspiegelt werden.²⁰ Zu berücksichtigen sind selbstverständlich auch jeweilige Spezifika konkreter Sprachkorpora, wie aus den Ausführungen in Kap. 5.2 ersichtlich ist.

Ad (c)

Es ist mehrmals darauf aufmerksam gemacht worden, dass verschiedene Lerner von der Korpusarbeit in unterschiedlichem Maße profitieren. – Ausschlaggebend scheinen hier das Sprachniveau, die PC-Kenntnisse und Lernziele der Fremdsprachenlerner sowie deren Angehörigkeit zum jeweiligen Lerntyp zu sein.²¹ Boulton (2009) führt an, dass die Korpusarbeit vor allem stark visuelle Lerntypen bevorzugen.²² In diesem Zusammenhang lässt sich feststellen, dass

¹⁵ Bernardini (2002: 166).

¹⁶ Peters/Picchi/Biagini (2000: 74).

¹⁷ Vachková (2007: 31).

¹⁸ Näheres dazu s. Königs (2001).

¹⁹ Z.B. Teubert (2004: 172).

²⁰ Lemnitzer/Zinsmeister (2009: 115ff.).

²¹ Römer (2009: 29).

²² Boulton (2009: 5).

the optimal teaching style is a balanced one in which all students are sometimes taught in a manner that matches their learning style preferences, so they are not too uncomfortable to learn effectively, and sometimes in the opposite manner, so they are forced to stretch and grow in directions they might be inclined to avoid if given the option.²³

Der Vorwurf, dass das DDL schnell langweilig und mechanisch werden könne, ist hier u.E. zu relativieren. Die Lerner verspüren Überdross an *jeder* Lern- und Lehrmethode, die im Unterricht einseitig bevorzugt wird. – Dass es zur Zeit etwa 50 verschiedene deutsche Korpora²⁴ sowie zahlreiche diverse Beispiele für die Korporäumsetzung²⁵ gibt, berechtigt uns zur Behauptung, dass es eine relativ lange Zeit in Anspruch nimmt, bis sich der Überdross an entsprechenden Lern- und Lehrmethoden einstellt. Daneben sei hier Folgendes bemerkt: Indem für die Korpusarbeit bzw. für mehr Korpusarbeit im Fremdsprachenunterricht geworben wird, wird damit sicher weder explizit noch implizit zum Ausdruck gebracht, dass diese zum alleinigen Lerngegenstand gemacht werden solle.

Ad (d)

Die Frage, ob der von McEnery/Wilson (vgl. Ad b) festgestellte Wandel im Rollenverständnis des Lehrers insgesamt gesehen wünschenswert ist, ist u.E. nicht a priori positiv zu beantworten. Gemäß unserer Erfahrung wünschen sich manche Fremdsprachenlerner eine relativ stark instruktive Unterweisung, um effektiv lernen zu können. Diesen Lernern wird ein „didactic pedagogue“ wohl bessere Dienste leisten als ein „fellow researcher“. Bernardinis Konzept des „serendipitous corpus browsing“ ist zum Teil von Mukherjee in Frage gestellt worden, und zwar nicht nur, weil es nicht ganz klar ist, wie sich bei solch einer extrem autonomen Lernform Lerninhalte bestimmen ließen, sondern eben auch, weil die Lehrer die Kontrolle über den Lernprozess völlig einbüßen müssten.²⁶

4 Das InterCorp-Projekt

Die deutsch-tschechische Kontrastierung²⁷ konnte in den letzten Jahren zahlreiche und wertvolle Impulse v.a. vom InterCorp-Projekt (Philosophische Fakultät der Karls-Universität Prag) erhalten.²⁸ Das Ziel des Projektes ist, als

²³ Felder/Spurlin (2005: 105, zitiert nach Boulton 2009: 14)

²⁴ Lemnitzer/Zinsmeister (2009: 115ff.).

²⁵ Z.B. Lemnitzer/Zinsmeister (2009), Flowerdew (2009), Boulton (2009).

²⁶ Mukherjee (2006: 14).

²⁷ Zum Stand des deutsch-tschechischen Sprachvergleichs siehe Šimečková (2001) und kürzlich Šemelík (2011).

²⁸ Das InterCorp-Projekt bildet einen Bestandteil des in den Jahren 2005–2011 unter Nr. 0021620823 realisierten Forschungsvorhabens „Das Tschechische Nationalkorpus

Materialbasis für Lexikographie, den Fremdsprachenunterricht, theoretisch orientierte Studien, studentische Arbeiten und Computerapplikationen dienende Parallelkorpora (mit dem Tschechischen stets als einem der Partner) aufzubauen. Ein besonderer Vorteil des InterCorp gegenüber vergleichbarer Korpora besteht darin, dass es multilingual ist und kontrastive Analysen sowohl von zwei als auch von mehreren, sog. kleineren Sprachen ermöglicht:

Due to a widespread knowledge of English and some other languages it is, in a way, a pair of two small languages that must be viewed as wanting in this respect. This is not meant as a political statement, so popular and vague in today's Europe, but a linguist's conviction that more data for a large-scale comparison and more qualified study of all kinds of languages is necessary and that must come from as many languages as possible.²⁹

Besonders intensive Sprachkontakte hat das im mitteleuropäischen Raum verwendete Tschechisch mit dessen Nachbarsprachen gehabt, d.h. mit dem typologisch verwandten Polnisch und Slowakisch sowie mit dem typologisch nicht verwandten Deutsch. Eben hier kann das InterCorp zu neuen korpusbasierten Forschungen anspornen, denn

specifically, the blurring of differences between two closely related languages (especially with Slovak) might be worth investigating in a parallel corpus. On the other hand, the long-standing contact with German, having a rich history, might be made more interesting if one goes more deeply, beyond mere loan-words, namely into semantics, calques or influences on the grammar system.³⁰

Das durch gemeinsame Pflege der Brünner und der Prager Germanisten kontinuierlich ausgebaute deutsch-tschechische InterCorp³¹ (fast 8,9 Millionen deutsche Wortformen) umfasst deutsche und tschechische nach 1945 erschienene literarische und publizistische Texte (Zeitraum 2000–2008) in geschriebener Form, wobei es neben tschechischen oder deutschen Originaltexten ebenfalls solche Texte enthält, die ins Tschechische aus einer dritten Sprache übersetzt worden sind.

This fact, i.e. having non-original texts on both sides in some cases, has to be taken into account in some kinds of analysis while in other this may not be so important.

und Korpora weiterer Sprachen“ (Leitung F. Čermák und V. Schmiedtová), vgl. mehr u. <http://ucnk.ff.cuni.cz/intercorp/> (Stand: 17. 6. 2011).

²⁹ Čermák (2009).

³⁰ Ebd.

³¹ T. Káňa und H. Peloušková (Pädagogische Fakultät der Masaryk-Universität Brunn), V. Dovalil und Š. Zbytovský (Philosophische Fakultät der Karls-Universität Prag). Auf linguistischen Treffen (*Deutsch und Tschechisch im Vergleich* – Brunn 2008, *Konferenz InterCorp 2009* – Prag) werden dann Ergebnisse der korpusbasierten Forschung des Tschechischen im Vergleich zum Deutschen präsentiert und in Sammelbänden veröffentlicht (Káňa/Peloušková 2009, Čermák/Klégr/Corness 2010).

Technique evaluating relevance of this kind of indirect translations from a third language, in comparison to direct equivalents, has to be found yet.³²

Daraus ergibt sich, dass die Methodologie zur Beurteilung der Äquivalenz auf allen Sprachebenen zu erarbeiten wäre. (In der Korpuslinguistik gilt die Authentizität der Forschungsbasis als eine besondere Stärke.)³³ Es trifft dabei aber auch zu, dass

[t]ranslation effects are not necessarily negative, because translation is an important way of renewing the target language. The degree of source vs. target orientation may vary depending upon a number of factors, such as text category or the degree of importance of the source text.³⁴

Für welche Zwecke das InterCorp verwendet werden kann, wird anschaulich in Čermák/Corness/Klégr (2010) und Čermák/Koček (2010) vorgestellt. Im Folgenden werden mögliche Anwendungen des InterCorp im fortgeschrittenen DaF-Unterricht im Rahmen des Germanistikstudiums besprochen.

5 Zur Anwendung des deutsch-tschechischen InterCorp im DaF-Unterricht

Erfahrungsgemäß ist es sinnvoll, bei einer ersten Annäherung an ein Korpus in drei Schritten vorzugehen: (a) vorbereitende Lektüre, (b) Korpusarbeit – Funktionalitäten des Korpus, (c) Korpusarbeit – Fremdsprachenlernen und linguistische Forschung.

5.1 „*Wer liest, der weiß was*“ – vorbereitende Lektüre

Der praktischen Korpusarbeit sollte u.E. eine theoretische Einführung in die Korpuslinguistik vorausgehen. Zu diesem Zweck eignen sich z.B. ausgewählte Kapitel aus Lemnitzer/Zinsmeister (2009). Die Parallelkorpora werden in Arbeiten wie Čermák (2009) oder Vachková (2010b) u.a.m. thematisiert. Solche Texte können als Referate vergeben werden.

³² Čermák (2009).

³³ Vgl. ebenda.

³⁴ Johansson (2007: 62).

5.2 „Was ist das und wie funktioniert das?“ – Funktionalitäten des InterCorp

Folgende, in mancherlei Hinsicht sicher noch ergänzbare Übungen sind als allererster Einstieg in die InterCorp-Arbeit gedacht.³⁵

1. Lesen Sie Informationen über das InterCorp-Projekt auf dessen Webseiten und beantworten Sie diese Fragen: Was für ein Korpus ist das InterCorp? – Was ist das Ziel des Projekts? – Wie viele Sprachen sind im InterCorp erfasst? – Nennen Sie die drei größten Subkorpora des InterCorp. – Welche Rolle spielt im InterCorp das Tschechische? – Welche Textsorten überwiegen im InterCorp? Welche Konsequenzen kann dies haben?³⁶
2. Sehen Sie sich das InterCorp-Interface an und finden Sie heraus, wie die Diakritik und flektierte Formen von der InterCorp-Software behandelt werden.
3. Bei welchen der folgenden Lexeme ist es am wahrscheinlichsten, dass sie im InterCorp nicht enthalten sind? Warum?³⁷
Hand, dero, mit, operieren, Gstätten, dissolut, extrinsisch, voraussetzen, nehmen
4. Wie können im InterCorp „Wortkombinationen“ gesucht werden?
5. Bei welchen der folgenden „Wortkombinationen“ ist es am wahrscheinlichsten, dass sie im InterCorp aufgefunden werden können? Warum?³⁸
 - (a) kalten Krieg – im kalten Krieg – dass er im kalten Krieg
 - (b) blauen Augen – in ihren blauen Augen – in ihren schönen blauen Augen
 - (c) damit er – damit er sich – damit er sich dort

5.3 „Wozu taugt das?“ – Fremdsprachenlernen und linguistische Forschung

5.3.1 Deutsch-tschechische Kontrastierung im lexikalischen Bereich – das InterCorp als Quelle authentischen Sprachmaterials

Da fehlende Lexikkenntnisse bzw. lexikalische Fehler im Vergleich zu Aussprache- oder Grammatikfehlern die Verständigung zwischen L1-Sprechern und L2-

³⁵ Ausführlichere Einführungsübungen in die Arbeit mit englischen Sprachkorpora sind z.B. in Frankenberg-Garcia (2010) zu finden.

³⁶ Es liegt auf der Hand, dass mit den Textsorten quantitative Fragestellungen zusammenhängen (vgl. auch Frage 3 unten). – Dass *kalt* in 81% der Belege die feste Verbindung *Kalter Krieg* ergibt (vgl. Kap. 5.3.2), ist keine reine Zufallsverteilung, sondern eine korpusbasierte Verteilung, denn das InterCorp enthält auch relativ viele publizistische Texte, in denen der Kalte Krieg häufiger thematisiert wird als z.B. im mündlichen Sprachgebrauch des Alltags.

³⁷ Durch diese Übung sollen die Studierenden daran gewöhnt werden, dass das InterCorp wegen der Textsortenzugehörigkeit und zeitlichen Charakteristika der Texte, aus denen das Korpus zusammengesetzt ist, bestimmte Wortschatzbereiche nicht bzw. nur lückenhaft erfassen vermag. Im InterCorp nicht zu finden sind: *dero* (veraltet), *Gstätten* (regional), *dissolut* (veraltet), *extrinsisch* (Fachsprache).

³⁸ Das Ziel dieser Aufgabe ist es, den Lernern beizubringen, dass je kürzer die gesuchte „Wortverbindung“ ist, desto höher sind deren Frequenzzahlen.

Lernern am stärksten beeinträchtigen,³⁹ kommt den Fragen des Wortschatzes im DaF-Unterricht eine besondere Rolle zu. Manche lexikalischen Fehler sind dabei zurückführbar auf (a) *interlinguale Unterschiede in der Versprachlichung eines bestimmten Bedeutungsspektrums*⁴⁰ und (b) unterschiedliche sprachhistorische Entwicklungen der jeweiligen Sprachen, die u.A. aus der Entstehung der sog. *faux amis*⁴¹ resultieren können. Für diese Unterschiede können Studierende sensibilisiert werden, indem ihnen entsprechende Texte und deren Übersetzungen gegenübergestellt präsentiert werden – das InterCorp bietet hierfür eine ergiebige Materialquelle. In Frage kommen z.B. auch *Lückentexte* als Ausgangspunkt der philologischen Beschäftigung mit diesem Thema – vgl. folgende Übung:

1. *Übersetzen Sie die fett gedruckten Ausdrücke ins Deutsche und setzen Sie sie in den entsprechenden deutschen Lückentext ein.*
 2. *Benennen Sie die sprachlichen Phänomene, die in Übung (A), bzw. (B) thematisiert werden. Was wissen Sie darüber? Nennen Sie weitere Beispiele für diese Erscheinungen.*
- A. (a) *Šel **schválně** pomalu, snažil se nemyslet na sako, **sáhl** po kytáře, **prsty** (1) se dotkly strun, struny **zakvílely** a dost.* Absichtlich ging er langsam, bemühte sich, nicht an das Jackett zu denken, griff nach der Gitarre, die ... (1) berührten die Saiten, die Saiten jaulten auf, und Schluß. (b) *Každá **zkušenost**, od **pobybu prstu na noze** (2) po **stavy blubokého zamýšlení**, se **ukládá** a **organizuje** skrze elektrochemické signály uvnitř neuronů a mezi nimi.* Jede Erfahrung, von der Bewegung ein... (2) bis hin zu tiefen Meditationsstadien, wird in dem elektrochemischen Signalsystem in und zwischen den Nerven gespeichert und von ihm organisiert. (c) *Ruce **měla hranaté**, s **krátkými tlustými prsty** (3), **noby dlouhé** a **hezké**, **chodidla drobná**, s **vysokou klenbou**, na **jedné noze jí chyběl malíček** (4).* Die Hände waren viereckig mit kurzen, dicken ... (3), die Beine lang und wohlproportioniert, die Füße klein mit hohem Spann, und ihr fehlte ein... klein... (4).
- B. (d) *Usmál se: „Vy to **ještě** **nevíte**, **ale** **jste vítěz** **mýho tajnýho konkursu** (5)!“* Er lächelte: „Sie wissen es noch gar nicht, aber Sie sind der Sieger mein... heimlich... ... (5)!“ (e) *„**Pozemečky** **zůstanou pěkně v bezpečí** **pro případ**, **že bychom zkrachovali** (zbankrotovali) (6)“.*
 „Grundstücke bleiben eine ausgezeichnete Sicherung für den Fall, daß wir in ... (6) gehen müßten. (Quelle: InterCorp)

³⁹ Köster (2001: 887).

⁴⁰ Vgl. z.B. folgende Asymmetrie: tsch. *prst* und dt. *Finger/Zeh(e)* – das Deutsche differenziert im Gegensatz zum Tschechischen danach, ob es um Hand oder Fuß geht.

⁴¹ Vgl. z.B. das formal ähnliche, doch in der Bedeutungsstruktur Unterschiede aufweisende Paar *konkurz/Konkurs*, bei dem Interferenzfehler drohen: tsch. *konkurz*: 1. Einstellung der Zahlungen einer Firma/gerichtliches Verfahren gegen eine Firma 2. Stellenausschreibung, dt. *Konkurs* nur in der Bedeutung 1 des tsch. Lexems.

5.3.2 Deutsche substantivierte Infinitive im Fokus. Das InterCorp als Assistenz eines Übersetzungswörterbuchs

Wie Wörterbuch- und Korpusarbeit mit Vermittlung bzw. Wiederholung der Wortbildung verbunden werden kann, soll hier kurz am Beispiel der Infinitivkonversion (*beulen* > *das Heulen*) aufgezeigt werden. Im Hinblick auf geringe Wortbildungsrestriktionen ist die Infinitivkonversion ein universales Mittel zur Substantivierung des Verbs. Das klassische Printwörterbuch als Medium stößt im Falle der substantivierten Infinitive jedoch notwendigerweise auf seine fest umrissenen Grenzen, wenn es aus Platzgründen keine (relativ) erschöpfende Inventarisierung von sämtlichen potenziellen Konversionsprodukten ermöglicht. Laut Fleischer/Barz gilt darüber hinaus, dass „[d]ie Infinitivkonversion weniger ein Mittel zur Bereicherung des Wortschatzes [ist] (obwohl auch diese Seite nicht fehlt) als vielmehr ein syntaktisch relevantes Nominalisierungsverfahren“,⁴² sodass die Inventarisierung dieser Konversionsprodukte beim oberflächlichen Blick wohl auch wenig Sinn macht.⁴³ Aus der kontrastiven Sicht sind aber manche der substantivierten Infinitive in Bezug auf deren in der Verwendung gängige Variationen für Übersetzungszwecke sehr wertvoll, vgl. z.B. bei *Heulen*. Aus dem InterCorp kann man minimal zu 11 Pendants gelangen (*vytí* 6x, *kvílení* 3x, *houkání* 3x, *pláč* 2x, *skučení* 2x, *ston* 1x, *ječivý křik* 1x, *brek* 1x, *stěny* 1x, *řev* 1x, *zaječení* 1x, frei übersetzt 3x, nicht übersetzt 1x). Falls man mechanisch nach den Äquivalentangaben zum Verb *heulen* in einem Übersetzungswörterbuch, etwa im SIEB, vorgehen würde, würde man durch Substantivierungen *vytí*, *houkání*, *hubláni*, *vřestění*, *brek(ot)*, *vřískání*, *řev* nur zu 7 Pendants gelangen. Von insgesamt 14 verschiedenen Übersetzungsäquivalenten können nur noch 4 im SIEB belegt werden, die aus Verben ableitbar sind und zugleich im InterCorp vorkommen. Das InterCorp zeigt sich hier hinsichtlich des Äquivalentsangebots dem SIEB überlegen. – Die immer noch anzutreffende, doch naive Vorstellung, dass das Wörterbuch ausschließlich Systemäquivalenz zu erfassen habe, gilt in der modernen Lexikographie als überholt. Versteht sich ein Wörterbuch als ein korpusbasiertes Übersetzungstool, dann hat es notwendigerweise mehr als nur Systemäquivalente anzubieten, zumal da stets konkrete Texte (Parole-daten), nicht aber Sprachsysteme (Langedaten) übersetzt werden. Stark kontextgebundene oder individuell geprägte Äquivalente bzw. nicht ganz richtige Entsprechungsalternativen können/sollen in die jeweiligen Wörterbuchartikel nicht aufgenommen werden, obwohl auch sie von Interesse wären, und zwar aus kontrastiver Perspektive als Belege für Interferenzen oder Übersetzungsfehler. Nicht einmal alle aufgefundenen, wenn auch richtigen Entsprechungsvarianten der Beispielsätze oder Beispielverbindungen können im Wörterbuchartikel un-

⁴² Fleischer/Barz (2007: 211).

⁴³ Auf lexikalisierte Fälle wie *das Leben* oder *das Versprechen* etc. trifft dies selbstverständlich nicht zu.

tergebracht werden, u.A. aus Platzgründen. An dieser Stelle muss aber ausdrücklich betont werden, dass die Entscheidungen hinsichtlich der (noch) aufzunehmenden Übersetzungsäquivalente zum Teil stark subjektiv sind, weil es keine handfesten Kriterien gibt, mit deren Hilfe sich eindeutig entscheiden ließe, wo genau die Grenze zwischen den Übersetzungsäquivalenten der jeweiligen Wortschatzeinheiten und den Äquivalenten ihrer nahen Synonyme zu ziehen ist. Die wiederholt bewiesene amorphe Beschaffenheit der Bedeutung vor dem Hintergrund der starken kontextuellen Abhängigkeit müsste für die Lexikographen ein Stimulus sein, die Freischaltung von Parallelkorpora für Benutzer elektronischer Übersetzungswörterbücher anzustreben, wie es Vachková für das GAWDT vorschlägt.⁴⁴

Einzelne Schritte bei der Arbeit mit dem InterCorp und anderen Informationsquellen können folgendermaßen aussehen: Der Lehrende (ggf. in Zusammenarbeit mit den Studierenden) wählt einen Text(abschnitt) mit einem substantivierten Infinitiv (z.B. *Heulen*) aus, der von den Studenten zu übersetzen ist. Danach sind folgende Fragen zu beantworten:

- (a) *Wie kann man Heulen in diesem Text übersetzen?*
- (b) *Ist Heulen in einem Übersetzungswörterbuch wie SIEB inventarisiert? Warum nicht?*
- (c) *Wie kann man sich in einer solchen Situation helfen?*
- (d) *Zu welchen Übersetzungen kann man gelangen, wenn man mechanisch nach den Äquivalenten zum Verb heulen vorgeht?*
- (e) *Welche Äquivalente zu Heulen bietet das InterCorp an?*
- (f) *Wie erklären Sie sich, dass das InterCorp mehr bzw. andere Übersetzungsäquivalente anbietet als man anhand des Wörterbuchartikels zu heulen in SIEB mechanisch eruieren kann?* (g) *Gibt es konkurrierende Wortbildungssynonyme zu substantivierten Infinitiven des Typs Heulen? Worin bestehen die Unterschiede zwischen diesen Wortbildungssynonymen?*⁴⁵

Diese Vorgehensweise lässt die philologische Sprachbeschäftigung als einen integrierten Bestandteil einer durchaus praktischen Tätigkeit (die des Übersetzens) erscheinen, womit die Motivation zum Studium gesteigert werden kann. Lernen in breiteren Zusammenhängen (hier: Wortbildung, Lexikographie, Übersetzen, Korpuslinguistik) kommt der Forderung eines ganzheitlich ausgerichteten Unterrichts entgegen.

Ein nächstes Beispiel einer kontrastiven Analyse der deutschen Adjektive *schlecht*, *schlimm* und *böse* und der tschechischen Gegenstücke *zlý*, *špatný* bietet

⁴⁴ Vachková (2011: 140).

⁴⁵ Antizipierte bzw. mögliche Antworten: (a) *vytí*, ... (b) nein, nicht lexikalisiert (c) die Übersetzungsäquivalente von *heulen* (d. h. tschechische Verben) substantivieren (d–e): siehe oben (f) die amorphe Beschaffenheit der Bedeutung vor dem Hintergrund der starken kontextuellen Abhängigkeit, die in einem Printwörterbuch nur lückenhaft dargestellt werden kann (g) ja, *Heulen* – *Gebeul(e)/Heulerei* [+lästig +unangenehm].

Anregung für eine mögliche Perspektivenänderung in der Richtung L2–L1.⁴⁶ Andere auf der InterCorp-Basis durchgeführte kontrastive Analysen sind bei Marková/Vachková⁴⁷ am Material der nahen Synonyme *studený/chladný* vs. *kalt/kühl* oder *schön/hübsch* vs. *pěkný/bezký/krásný* zu finden. Zu dem ersteren Wortpaar (*kühl* 364x im InterCorp, *kalt* 1012x) können hier als Beispiel folgende Teilresultate angeführt werden (ohne Frequenzangaben zu einzelnen Übersetzungsäquivalenten).⁴⁸ Beide Wörter können im Deutschen sowohl als synonymische (indem sich ihre Bedeutungen ergänzen), als auch graduonymische Paare (indem sie eine bestimmte Gradation ausdrücken oder im Kontrast stehen usw.) funktionieren. Nun kann man sich fragen, inwieweit eine solche Bedeutungsstruktur im Tschechischen bei dem Wortpaar *studený – chladný* vorhanden ist. Die Suche nach entsprechenden Textäquivalenten im Parallelkorpus dient dann dem eigentlichen Zweck der Untersuchung: Eine genauere, verwendungsbasierte semantische Beschreibung der beiden nahen Synonyme und deren Distribution wird in beiden Sprachen angestrebt. So wird z.B. *kühl* nicht nur als *chladný* (und verwandte Wörter wie *chladno* usw.) übersetzt, sondern es bietet sich eine ganze Skala von Bedeutungsvariationen auch aufgrund der Betrachtung von Übersetzungsäquivalenten (*studený, mrazivý, svěží, jasný, strážlivý, ledový, racionální, odměřený* usw.), die in verschiedene Kollokationsstrukturen eingebunden sind und die man in einem Übersetzungswörterbuch kaum finden kann. *Kalt* ist durch die usuelle Wortverbindung *kalter Krieg* (*studená válka*) belastet (in dieser Verbindung kommt es in 81 von 100 Fällen vor, was die Lerner einerseits überraschen kann, andererseits können sie aufgrund dessen die natürliche Distribution dieses Wortes im Sprachgebrauch sehen, was sonst kaum möglich wäre). Beeinflusst von dieser festen Verbindung auch in anderen Konstruktionen, wird *kalt* auch z.B. als *postkomunistický* („postkommunistisch“) oder *poválečný* („Nachkriegs-“) übertragen. Es liegt dabei auf der Hand, dass gerade solche Übersetzungsäquivalente wie *racionální* u.Ä. im Falle des Adjektivs *kühl* zu einer feineren Bedeutungsbeschreibung des deutschen Wortes beitragen können (Kühle als Metapher der Rationalität, Nüchternheit bzw. Sachlichkeit). In der „eigentlichen“ Bedeutung kommen dann die Äquivalente *studený* und *chladný* vor (samt deren Variationen). Dabei kommt eine breite Skala von Übersetzungsvariationen zum Vorschein, die für einen zukünftigen Linguisten oder Übersetzer Anregung und Anlass zur tieferen semantischen Analyse in einem darstellen.

⁴⁶ Vachková (2010b).

⁴⁷ Marková/Vachková (i.V.).

⁴⁸ Zur korpuslinguistischen Analyse ausgewählter semantisch naher Wortpaare im Deutschen (u. a. *kalt/kühl*), wo die semantische Strukturierung, Kollokabilität und Produktivität wie auch Frequenz der ausgewählten Wortpaare untersucht werden, s. Marková (2010).

5.3.3 Wortarten in deutsch-tschechischer Sicht – das InterCorp als Assistenz einer künftigen kontrastiven Grammatik

Den tschechischen Germanistikstudenten steht zur Zeit keine umfassende und aktuelle, kontrastiv ausgerichtete Grammatik des Deutschen zur Verfügung – die korpus-basierte, pragmatisch-funktional orientierte vergleichende Grammatik von Štícha ist trotz ihres beeindruckenden Umfangs nur ausgewählten, vorwiegend syntaktischen Fragestellungen gewidmet, während das zweifellos verdienstvolle Werk von Povejšil heutzutage bereits veraltet ist. Hier würde sich z.B. die Möglichkeit bieten, die grammatikorientierte Arbeit mit dem InterCorp in den *Projektunterricht* einzubinden, in dessen Rahmen die Studierenden in Arbeitsgruppen einzelne Teilbereiche der kontrastiven Grammatik des Sprachenpaars Deutsch-Tschechisch erarbeiten würden. Die Aufgabe könnte z.B. wie folgt lauten:

Stellen Sie sich vor, Sie sind vom Academia-Verlag beauftragt worden, an der „Kurzgefassten deutsch-tschechischer kontrastiven Grammatik“ mitzuarbeiten und das Kapitel „Zahlwörter/ Číslovky“ zu verfassen. Die Redaktion arbeitet mit Štícha 2003 und Povejšil 1987 (und neuere Aufl.), die sie als erste Orientierungshilfe verwendet. Als Materialbasis für die Forschung dient jedoch vorzugsweise das Parallelkorpus InterCorp. Versetzen Sie sich in die Rolle eines Grammatikverfassers! Legen Sie Ihre Arbeit in Form eines Fließtextes bis ... (wann) vor. Die Präsentation der Ergebnisse Ihrer Forschung einschl. Diskussion erfolgt im Seminar am... (wann).

Im Rahmen dieser Präsentation könnte der Lehrende z.B. Fragen stellen wie:

- (a) *Wie beurteilen Sie das Informationsangebot bestehender kontrastiver Grammatiken in Bezug auf die Zahlwörter?*
- (b) *Enthalten die Grammatiken genug authentische Beispiele?*
- (c) *Gibt es Übereinstimmungen zwischen den Angaben in Grammatiken und dem Sprachgebrauch, wie ihn das InterCorp erfasst?*
- (d) *Gibt es Übersetzungsmöglichkeiten, die Sie weder in den Grammatiken noch im InterCorp gefunden haben?*
- (e) *Woran könnte dies liegen? Sollen diese Übersetzungsmöglichkeiten in so einer Grammatik, an der Sie gearbeitet haben, berücksichtigt werden? Warum (nicht)?*
- (f) *Gibt es im InterCorp Übersetzungen, die Ihrer Meinung nach nicht korrekt sind?*
- (g) *Welche Ihrer Feststellungen halten Sie für besonders wertvoll? Warum?*

Indem bestehendes Wissen bzw. dessen Darstellung aufgrund Korpusrecherchen in Frage gestellt werden kann, wobei die Korpusdaten einer gemeinsamen Diskussion unterzogen werden, können Studierende im Geiste von Johns' „Every student a Sherlock Holmes!“⁴⁹ zum forschenden Browsen im Parallelkorpus und

⁴⁹ Johns (2002: 108, zit. nach Römer 2006: 124).

kritischen Umgang mit gängigen Grammatiken sowie mit dem Korpus selbst angeleitet werden. Die Lernform des Projektunterrichts stellt darüber hinaus den Raum für kooperatives Verhalten bereit, das unserer Meinung nach nicht einmal im tertiären Schulwesen vernachlässigt werden sollte. Indem sich die Kursteilnehmer in die Rolle der Grammatikschreiber versetzen, erweitern sie nicht nur ihre Sprachkenntnisse, sondern erwerben auch Kompetenzen, die sie eventuell als künftige Lehrer oder Forscher verwerten werden (Auswahl treffender Beispiele, Darbietung des Lehrstoffes u.Ä.) – dadurch kann im Rahmen dieses Rollenspiels *sui generis* auch die indirekte Anwendung der Sprachkorpora im Fremdsprachenunterricht thematisiert werden.

6 Ausblick und Desiderate

In der Linguistik allgemein wie auch im DaF-Unterricht sind es zurzeit in erster Linie elektronische Korpora, die neue Perspektiven eröffnen. Dass in Lehre und Studium Korpora benutzt werden sollen, ist mittlerweile zwar evident geworden, doch der Rückstand des Faches *Korpuslinguistik und DaF* bleibt im Vergleich mit *English as a Foreign Language* weiterhin eklatant. Deutsche Einführungen in die Korpuslinguistik gibt es erst seit 2006, eine spezifische Einführung für das Fach DaF fehlt immer noch. Während die Korpuslinguistik in Tschechien wie in deutschsprachigen Ländern zumindest im universitären Bereich Fuß gefasst zu haben scheint, gibt es unseres Wissens keine Evidenz, dass Korpora den Weg in die Arbeitszimmer der tschechischen Deutschlehrer oder sogar in die Klassenzimmer der tschechischen Gymnasien gefunden haben. Eine Art „Missionsarbeit“, d.h. eine stärkere Werbung für die Korpusarbeit auch im außeruniversitären Bereich, erscheint uns daher angebracht, wenn gar notwendig.

Die Arbeit mit lexikalischen Merkmalskarten im DaF-Unterricht anhand der Kookurrenzdatenbank CCDB⁵⁰ erfolgt auf einer einsprachigen Basis; die kontrastiven Gesichtspunkte können, wie gezeigt, ebenfalls herangezogen werden. – Wie ein deutsch-tschechisches Parallelkorpus im fortgeschrittenen Fremdsprachenunterricht angewendet werden kann, ist hier auf der Grundlage unserer ersten Erfahrungen nur angedeutet worden. Auch im aktuellen Umfang ist das deutsch-tschechische InterCorp eine äußerst nützliche und wertvolle Materialbasis für die Sprach(en)forschung. Trotzdem würde man sich wünschen, dass das Korpus auch in der Zukunft kontinuierlich ausgebaut werden würde, denn das bekannte „Weniger ist mehr“ gilt hier nicht – je größer das Korpus ist, desto verlässlicher und generalisierbarer sind die Ergebnisse der Korpusanfragen.

⁵⁰ <http://corpora.ids-mannheim.de/ccdb/>, s. Vachková/Marková/Belica (2008). Das Thema der lexikalischen Merkmalskarten wird in der vorbereiteten Monographie von Schmidt aus der didaktischen Perspektive näher behandelt.

Fasst man das Sprachenpaar Deutsch-Tschechisch unter der Perspektive einer didaktisch orientierten, korpus-basierten kontrastiven Sprachforschung ins Auge, fällt sofort auf, dass kein *Lernerkorpus*, das Texte von tschechischen Deutschlernern und deutschen Tschechischlernern zusammenfassen würde, zur Verfügung steht.⁵¹ Das ist ein Manko, das möglichst bald behoben werden müsste. Ergebnisse gezielter Korpusanalysen würden dabei nicht nur in der Fremdsprachendidaktik, sondern auch in der deutsch-tschechischen Lexikographie umgesetzt werden – um so eher, als ein Übersetzungswörterbuch zum Teil ebenfalls die Rolle eines Lernerwörterbuchs übernehmen muss und somit didaktische Ziele zu verfolgen hat.

Literatur

(a) Forschungsliteratur

- Bernardini, Silvia (2002): Exploring new directions for discovery learning. In: Kettemann, Bernhard/Marko, Georg (Hrsg.): *Teaching and Learning by Doing Corpus Analysis. Proceedings of the Fourth International Conference on Teaching and Language Corpora*, Graz 19–24 July, 2000. Amsterdam: Rodopi, 165–182. <http://books.google.cz/books?id=8YA6wJ6F-MwC&printsec=frontcover&dq=%22Teaching+and+Learning+by+Doing+Corpus+Analysis> (Stand: 17.6.2011)
- Boulton, Alex (2009): Corpora for all? Learning styles and data-driven learning. In: *Proceedings of the 5th Corpus Linguistics Conference*, 1–20. <http://ucrel.lancs.ac.uk/publications/cl2009/> (Stand: 17.6.2011)
- Čermák, František (2009): Parallel Corpora: The Case of InterCorp, a multilingual corpus. <http://ucrel.lancs.ac.uk/publications/cl2009/> (Stand: 17.6.2011)
- Čermák, František/Koček, Jan (Hrsg.) (2010): *Mnohojazyčný korpus InterCorp: Možnosti studia*. Praha: NLN.
- Čermák, František/Klégr, Aleš/Corness, Patrick (Hrsg.) (2010): *InterCorp: Exploring a Multilingual Corpus*. Praha: NLN.
- Dovalil, Vít (2007): Sprachenpolitik in der Tschechischen Republik (unter besonderer Berücksichtigung der Beziehungen zur EU und zum Europarat). In: Blanke, Detlev et al. (Hrsg.): *Sprachenpolitik und Sprachkultur, Sprache. System und Tätigkeit*, Bd. 57, Frankfurt a. M.: Peter Lang.
- Fleischer, Wolfgang/Barz, Irmhild (2007): *Wortbildung der deutschen Gegenwartssprache*. 3. Aufl. Tübingen: Niemeyer.
- Frankenberg-Garcia, Ana (2010, im Druck): Raising teacher's awareness of corpora. In: *Language Teaching*. Cambridge University Press. <http://anafrankenberg.synthasite.com/publications.php> (Stand: 17.6.2011)
- Ger (2001): Der Gemeinsame europäische Referenzrahmen in seinem politischen und bildungspolitischen Kontext. In: *Europarat/Rat für kulturelle Zusammenarbeit: Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*.

⁵¹ Zum deutsch-tschechischen Fehlerkorpus s. Kratochvílová/Rinas (2003).

- Council for Cultural Co-operation. Educational Committee. Modern Languages Division, Strassbourg, übers. von Quetz, J. et al., hrsg. von Goethe-Institut Inter Nationes, der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK), der Schweizerischen Konferenz der Kantonalen Erziehungsdirektoren (EDK) und dem österreichischen Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK), Berlin/München/Wien/Zürich/New York: Langenscheidt, 14–20.
- Johns, Tim F. (2002): Data-driven learning: The perpetual challenge. In: Kettemann, Bernhard/Marko, Georg (Hrsg.): Teaching and Learning by Doing Corpus Analysis. Proceedings of the Fourth International Conference on Teaching and Language Corpora, Graz 19–24 July, 2000. Amsterdam: Rodopi, 107–117. <http://books.google.cz/books?id=8YA6wJ6F-MwC&printsec=frontcover&dq=%22Teaching+and+Learning+by+Doing+Corpus+Analysis> (Stand: 17.6.2011)
- Johns, Tim F. (1986): Microconcord: A Language-learner's Research Tool. In: System 14(2), 151–162.
- Káňa, Tomáš/Peloušková, Hana (Hrsg.) (2009): Deutsch und Tschechisch im Vergleich. Korpusbasierte linguistische Studien. Brno: Masarykova univerzita.
- Königs, Frank G. (2001): Übersetzen. In: Helbig, Gerhard et al. (Hrsg.): Deutsch als Fremdsprache. Ein internationales Handbuch. Berlin/New York: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 19.2), 955–962.
- Köster, Lutz (2001): Wortschatzvermittlung. In: Helbig, Gerhard et al. (Hrsg.): Deutsch als Fremdsprache. Ein internationales Handbuch. Berlin/New York: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 19.2), 887–893.
- Kratochvílová, Iva/Rinas, Karsten (2003): Das deutsch-tschechische Fehlerkorpus. Beitrag im Rahmen der internationalen Konferenz „Korpuslinguistik deutsch – synchron, diachron, kontrastiv“ Würzburg In: Schwitalla, Johannes/Wegstein, Werner (Hrsg.): Korpuslinguistik deutsch: synchron, diachron, kontrastiv. Tübingen: Niemeyer, 277–281.
- Leech, Geoffrey (1997): Teaching and Language Corpora: A Convergence. In: Wichmann, Anne/Fligelstone, Steven/McEnery, Tony/Knowles, Gerry (Hrsg.): Teaching and Language Corpora. London: Longman, 1–23.
- Marková, Věra (2010): Zur Ermittlung der lexikalischen Beziehungen zwischen semantisch nahen Adjektiven. Institut für germanische Studien, Karls-Universität in Prag (Dissertation).
- Marková, Věra/Vachková, Marie (2011, in Vorbereitung): Neues in der Erforschung lexikalischer Beziehungen. In: Peloušková, Hana/Káňa, Tomáš (Hrsg.): Deutsch und Tschechisch im Vergleich II. Korpusbasierte linguistische Studien. Brno: Masarykova univerzita, Pedagogická fakulta.
- McEnery, Tony/Wilson, Andrew (1997): Teaching and Language Corpora. In: ReCALL, Vol. 9, No. 1, 5–14. <http://www.eurocall-languages.org/recall/pdf/rvol9no1.pdf> (Stand: 17.6.2011)
- Mukherjee, Joybrato (2006): Corpus linguistics and language pedagogy: the state of the art – and beyond. In: Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods. Eds. Sabine Braun, Kurt Kohn & Joybrato Mukherjee. Frankfurt a.M.: Peter Lang, 5–24. <http://www.uni-giessen.de/anglistik/ling/Staff/mukherjee/> (Stand: 17.6.2011)

- Peloušková, Hana/Káňa, Tomáš (Hrsg.) (2009): Deutsch und Tschechisch im Vergleich. Korpusbasierte linguistische Studien. Brno: Masarykova univerzita, Pedagogická fakulta.
- Perkuhn, Rainer/Belica, Cyril (2006): Korpuslinguistik – das unbekannte Wesen oder Mythen über Korpora und Korpuslinguistik. In: Sprachreport 1/2006, 2–8.
- Peters, Carol/Picchi, Eugenio/Biagini, Lisa (2000): Parallel and Comparable Bilingual Corpora in Language Teaching and Learning. In: Botley, Simon/McEnery, Tony/Wilson, Andrew (Hrsg.): Multilingual Corpora in Teaching and Research, Amsterdam/Atlanta: Rodopi, 73–75.
- Procházková, Bára/Komárek, Michal (2009): Český babel. <http://respekt.ihned.cz/c1-38494540-cesky-babel> (Stand: 17.6.2011)
- Römer, Ute (2006): Pedagogical Applications of Corpora: Some Reflections on the Current Scope and a Wish List for Future Developments. In: Zeitschrift für Anglistik und Amerikanistik 54(2), 121–134. (Special Issue: “The Scope and Limits of Corpus Linguistics – Empiricism in the Description and Analysis of English”, ed. Volker Gast). <http://www.utoeroemer.de/publikationen.htm> (Stand: 17.6.2011)
- Römer, Ute (2008): 7. Corpora and language teaching. In: Lüdeling, Anke/Kytö, Merja (Hrsg.): Corpus Linguistics. An International Handbook. Berlin: Mouton de Gruyter, 112–130. <http://www.utoeroemer.de/publikationen.htm> (Stand: 17.6.2011)
- Römer, Ute (2009): Corpus research and practice: What help do teachers need and what can we offer? In: Aijmer, Karin (ed.). Corpora and Language Teaching. Amsterdam: John Benjamins. 83–98. <http://www.utoeroemer.de/publikationen.htm> (Stand: 17. 6. 2011)
- Šemelík, Martin (2011): Zum aktuellen Stand des deutsch-tschechischen Sprachvergleichs. In: Die Stellung der Germanistik in der Tschechischen Republik nach der Bologna-Reform. Brno: Tribun, 87–84. <http://lexarchiv.ff.cuni.cz/lexikograficka-sekce/biblio/download.html> (Draftversion, Stand: 17.6.2011)
- Šemelík, Martin/Vachková, Marie (2011): Zirkumfixbildungen auf Ge-(e) aus kontrastiver (deutsch-tschechischer) und korpuslinguistischer Sicht. In: Lejsková, Alena/Valdrová, Jana (Hrsg.): Die Grammatik, Semantik und Pragmatik des Wortes. Ihre Erforschung und Vermittlung. Augsburg: Wißner-Verlag, 124–139.
- Šimečková, Alena (2001): Kontrastive Analysen Deutsch-Tschechisch/Slowakisch: eine Übersicht. In: Helbig, Gerhard et al. (Hrsg.): Deutsch als Fremdsprache. Ein internationales Handbuch. Berlin/New York: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 19.1), 394–403.
- Teubert, Wolfgang (2004): Units of Meaning, Parallel Corpora, and Their Implications for Language Teaching. In: Connor, Ulla/Upton, Thomas A. (Hrsg.): Applied Corpus Linguistics: A Multidimensional Perspective (Language and Computers, no. 52), Amsterdam: Rodopi, 171–189. <http://books.google.cz/books?id=xiGIoDMOuYQC&printsec=frontcover&dq=Applied+Corpus+Linguistics:+A+Multidimensional+Perspective> (Stand: 17.6.2011)
- Vachková, Marie (2007): Kapitoly k německo-české metalexikografii I. Praha: FF UK.
- Vachková, Marie (2010): Tschechische Republik. In: Germanistik und Deutschunterricht in neun Ländern. <http://www.ids-mannheim.de/org/gremien/intwrat.html> (Stand: 14. 12.2010)
- Vachková, Marie (2010b): Parallel Corpora and Compilation of a Bilingual Dictionary. In: Čermák, František/Klégr, Aleš/Corness, Patrick (Hrsg.): InterCorp: Exploring a Multilingual Corpus. NLN. Praha, 235–252.

- Vachková, Marie (2011): Das große akademische Wörterbuch Deutsch-Tschechisch. Ein erster Werkstattbericht. Frankfurt a.M.: Peter Lang.
- Vachková, Marie/Marková, Věra/Belica, Cyril (2008): Korpusbasierte Wortschatzarbeit im Rahmen des fortgeschrittenen Germanistikunterrichts. In: Zielsprache Deutsch. Zeitschrift für Unterrichtsmethodik und angewandte Sprachwissenschaft 3, 20–35.

(b) Wörterbücher und Grammatiken

- Siebenschein, Hugo et al. (2006): Německo-český slovník (Wörterbuch deutsch-tschechisch), Voznice: Nakladatelství Leda (SIEB).
- Povejšil, Jaromír (1987): Mluvnice současné němčiny (Grammatik des Gegenwartsdeutschen), Praha: Academia.
- Štícha, František (2003): Česko-německá srovnávací gramatika (Tschechisch-deutsche vergleichende Grammatik), Praha: Argo.

(c) Sprachkorpora

- Belica, Cyril: Kookkurrenzdatenbank CCDB. Eine korpuslinguistische Denk- und Experimentierplattform für die Erforschung und theoretische Begründung von systemisch-strukturellen Eigenschaften von Kohäsionsrelationen zwischen den Konstituenten des Sprachgebrauchs. © 2001–2007 Institut für Deutsche Sprache, Mannheim.
- Das deutsch-tschechische InterCorp: InterCorp, Parallelkorpora der Philosophischen Fakultät der Karls-Universität Prag, <http://ucnk.ff.cuni.cz/intercorp/>