

HRISZTALINA HRISZTOVA-GOTTHARDT AND  
ZOLTÁN GOTTHARDT

## SPRICHWÖRTER IM BULGARISCHEN NATIONALEN KORPUS

**Abstract:** Der vorliegende Beitrag präsentiert die Methoden und Ergebnisse einer korpusbasierten Untersuchung zur Vorkommenshäufigkeit bulgarischer Sprichwörter in der geschriebenen Gegenwartssprache. Im Rahmen dieser Untersuchung wurden mit Hilfe eines Suchalgorithmus Texte des Bulgarischen nationalen Korpus nach Sprichwörtern durchsucht. Dabei wurden nicht nur jene Sprichwörter ermittelt, die am häufigsten im Korpus vorkommen, sondern auch ihre Varianten und Verfremdungen. Es konnte festgehalten werden, dass die Ergebnisse dieser Untersuchung im Wesentlichen mit den Resultaten zweier früherer korpusbasierter Untersuchungen korrespondieren und daher erste Rückschlüsse auf die tatsächliche parömiologische Situation in Bulgarien zulassen.

**Keywords:** bulgarische Sprichwörter, Bulgarisches nationales Korpus, geschriebene Gegenwartssprache, Vorkommenshäufigkeit, PERL

### *1. Einleitung*

Bis zum jetzigen Zeitpunkt wurden keine empirischen Untersuchungen durchgeführt, die es zum Zweck hatten, die aktuelle parömiologische Situation in Bulgarien zu ermitteln. Das heißt, es liegen keinerlei Daten über den aktuellen Gebrauch bulgarischer Sprichwörter in der geschriebenen und der gesprochenen Sprache vor. Das zeichnet sich besonders deutlich in den modernen parömiographischen Werken ab, die ausschließlich auf früheren Sammlungen aufbauen und dementsprechend oft zahlreiche veraltete und nicht mehr gebräuchliche Sprichwörter aufzeichnen. Aus diesem Grund werden diese Werke von einzelnen Parömiologen als „Datenfriedhöfe“ bezeichnet (vgl. Baur / Chlosta 1996: 92). Das trifft bedauerlicherweise auch für die bulgarischen Sammlungen zu. Sogar die in dem neuesten bulgarischen Sprichwörterbuch aufgelisteten Texte stammen laut Quellenverzeichnis alle aus

Bänden, deren Erscheinungsjahr über mehrere Jahrzehnte zurückliegt (s. Stojkova 2007: 19).

In diesem Sinne erweist es sich als dringend nötig, erste Schritte in Richtung Ermittlung der aktuellen parämiologischen Situation in Bulgarien einzuleiten. Es können dabei zum einen empirische Untersuchungen zur Bekanntheit von Sprichwörtern durchgeführt werden. Darunter versteht man jene Felduntersuchungen, die mittels des Teilpräsentation-Verfahrens auf die Ermittlung der bekanntesten Sprichwörter einer Sprachgemeinschaft abzielen (s. Permjakov 1988, Grzybek 1991, Baur / Chlosta 1996 und Tóthné Litovkina 1996). Zwar liefern solche Untersuchungen Informationen darüber, welche Texte gerade die bekanntesten sind und in welcher Form sie bekannt sind, sie sagen aber eher wenig über ihre aktive Verwendung und ihre Vorkommenshäufigkeit in der Sprache aus. Zur Beantwortung der letzteren Frage wird ein anderes Herangehen angewendet, nämlich die Korpusanalyse. Dabei werden umfangreiche Korpora der geschriebenen Sprache auf die Gebrauchsfrequenz von Sprichwörtern geprüft, indem nicht nur ihr reales Vorkommen, sondern auch ihre konkreten Kontexteinbettungen, Verwendungssituationen und ihre Abwandlungs- und Transformationsfähigkeit untersucht werden (können) (vgl. Čermák 2003; Ďurčo 2006).

An diesen korpuslinguistischen Ansatz anknüpfend, führte Hrisztova-Gotthardt zwischen 2008 und 2009 eine Pilotuntersuchung durch, in deren Rahmen sie erste Erkenntnisse über die Verwendung und Vorkommenshäufigkeit bulgarischer Sprichwörter in zeitgenössischen Zeitungstexten gewinnen konnte (s. Hrisztova-Gotthardt 2010). Das Korpus der Untersuchung bildeten Artikel aus der bulgarischen Tageszeitung *Стандарт* [Standart], die zwischen dem 4. Januar 2000 und 3. August 2008 in der elektronischen Ausgabe von *Стандарт* erschienen waren. Die Texte wurden mit Hilfe eines für die Zwecke der Untersuchung entwickelten Suchalgorithmus nach 2.301 Sprichwörtern<sup>1</sup> durchsucht. Die Ergebnisse der Untersuchung wurden in einer Tabelle zusammengefasst, in der alle 225 im Korpus vorkommenden Texte absteigend nach Zahl der Treffer aufgelistet wurden. Es wurden dabei nicht nur die exakten Treffer verzeichnet, sondern auch zahlreiche qualitative und quantitative Varianten sowie sehr viele Verfremdungen (Antisprichwörter).

2011 wurde eine weitere Untersuchung in die Wege geleitet. Dieses Mal diente das WWW als Korpus. Unter Zuhilfenahme eines speziellen Programms und der Suchmaschinen *Google* und *Bing* ermittelten Hrisztova-Gotthardt und Gotthardt erste Daten über den Gebrauch bulgarischer Sprichwörter in diversen Internet-texten. Nach der Bearbeitung der Trefferlisten stellte sich heraus, dass die Ergebnisse dieser Untersuchung große Ähnlichkeiten mit den Resultaten der ersten Studie aufzeigen: die Texte mit den meisten Treffern im *Стандарт*-Korpus kamen auch in den WWW-Texten am häufigsten vor (vgl. Hrisztova-Gotthardt/Gotthardt 2011).

In diesem Zusammenhang stellt sich logischerweise die Frage, ob eine gezielte Suche nach Sprichwörtern im Bulgarischen nationalen Korpus ähnliche Ergebnisse erbringen wird. Dementsprechend setzt sich der vorliegende Beitrag zum Ziel, diese Frage zu beantworten. Im Rahmen dieser Arbeit werden die Methoden und Ergebnisse einer weiteren Untersuchung dargestellt, bei der die Texte des Bulgarischen nationalen Korpus nach Sprichwörtern durchsucht wurden. Dabei steht nicht nur die Bestimmung der am häufigsten gebrauchten Texte im Mittelpunkt der Untersuchung, sondern auch die Ermittlung ihrer möglichen Varianten und Verfremdungen (Antisprichwörter).

## **2. Das Korpus der Untersuchung**

Wie bereits weiter oben erwähnt, wurden im Rahmen dieser Untersuchung die Texte des Bulgarischen nationalen Korpus<sup>2</sup> näher unter die Lupe genommen. Das Korpus wurde zwischen 2001 und 2009 von Mitarbeitern des Instituts für bulgarische Sprache an der Bulgarischen Akademie der Wissenschaften aufgebaut und wird seitdem ständig um neue Texte erweitert. Mit seinen ca. 320.000.000 Korpuswörtern (Token) gilt das Korpus als relativ groß (s. Scherer 2006: 7 und Lemnitzer / Zinsmeister 2010: 110). Da die gespeicherten Texte alle nach 1945 entstanden sind und möglichst viele Genres und Textsorten abdecken, steht das Korpus repräsentativ für die moderne geschriebene bulgarische Sprache.

Das Korpus kann frei für wissenschaftliche Zwecke benutzt werden, es wird lediglich eine Registrierung benötigt. Da alle Texte tokenisiert und lemmatisiert<sup>3</sup> wurden, kann im Korpus nicht nur nach wörtlichen Entsprechungen, sondern auch nach sämtlichen Flexionsformen eines Wortes gesucht werden. Die lexikalisch-

semantische Annotation der Lemmata ermöglicht des Weiteren die Suche nach Synonymen, Hyponymen und Hyperonymen eines bestimmten Wortes<sup>4</sup>.

Es besteht außerdem die Möglichkeit, nach Phrasen zu suchen, deren Komponenten je nach einem bestimmten Kriterium variieren dürfen. Zum Beispiel kann man als Suchbedingung angeben, dass an Stelle einer bestimmten Komponente einer Phrase ein anderes Wort einer beliebigen Wortart stehen darf. Es kann auch nach Phrasen gesucht werden, zwischen deren „festen“ Komponenten eine bestimmte Anzahl von ebenfalls „beliebigen“ Wörtern vorkommen darf. Natürlich stehen dem Benutzer auch weitere Such- und Restriktionsmöglichkeiten zur Verfügung, die jedoch an dieser Stelle nicht alle behandelt werden können<sup>5</sup>.

### ***3. Die Methode der Untersuchung***

Die im Rahmen der ersten Untersuchung zur Vorkommenshäufigkeit bulgarischer Sprichwörter in der geschriebenen Gegenwartssprache zusammengestellte Liste mit 2.301 tokenisierten Sprichwörtern wurde auch bei diesem Projekt eingesetzt. Die Suche nach den in der Liste verzeichneten Texten verlief in folgenden Schritten:

- Mit Hilfe eines zum Zwecke der früheren Untersuchungen erstellten PERL-Programms wurde das Korpus nach exakten Treffern für die 2.301 Texte durchsucht. In dieser ersten Phase der Untersuchung startete das Programm eine Suchanfrage an die Korpus-Suchmaschine, indem es die 2.301 tokenisierten Sprichwörter als einzelne Suchbegriffe eingab. Es wurde nach Texten gesucht, die zu 100% mit den vorgegebenen Suchbegriffen übereinstimmen: weder die Flexionsformen der einzelnen Konstituenten noch ihre Reihenfolge durften auch die kleinste Abweichung aufweisen. Die Anzahl der Treffer für die einzelnen Suchbegriffe (Sprichwörter) wurde vom Programm gespeichert. Auf diese Weise wurde die Zahl der exakten Treffer ermittelt.
- In der zweiten Phase der Untersuchung wurde eine weitere, verfeinerte Suche gestartet. Im Rahmen dieser Suche wurde das Korpus nach Varianten und Verfremdungen jener 10 Sprichwörter durchsucht, die in der ersten Phase die meisten exakten Treffer erzielt hatten (s. dazu 4. Ergebnisse der Untersuchung).

- Zum einen wurde nach Texten gesucht, in denen die Konstituenten des traditionellen Sprichwortes in einer anderen Flexionsform vorkamen, z.B. im Plural statt im Singular. Außerdem durften in den Treffern zwischen den einzelnen Komponenten (maximal 5) weitere Wörter vorhanden sein.
- Zum anderen wurden jene Texte ermittelt, in denen an Stelle jeder einzelnen Konstituente des traditionellen Sprichwortes ein anderes Wort stand. Dieses Wort sollte allerdings derselben Wortart angehören wie die ursprüngliche Konstituente. Z.B. durfte die Stelle eines Nomens nur durch ein anderes Nomen eingenommen werden.<sup>6</sup> Auch in diesem Fall durften zwischen den einzelnen Konstituenten maximal 5 weitere Wörter vorhanden sein.

Mit Hilfe des weiter oben dargestellten Suchalgorithmus wurden zahlreiche Varianten und Verfremdungen der traditionellen Sprichwörter ermittelt, beziehungsweise auch Texte, in denen man auf bestimmte Sprichwörter angedeutet hat, ohne die vollständigen Texte wörtlich zu zitieren. Alle Texte, die das Programm als Treffer aufgelistet hatte, mussten manuell durchsucht werden, um die tatsächlichen Varianten, Verfremdungen und Andeutungen bzw. Hinweise auf Sprichwörter von den irrelevanten Hits zu trennen.

Im Endergebnis wurde eine neue Liste zusammengestellt, die alle gefundenen Varianten, Verfremdungen und die eventuellen Andeutungen bzw. Hinweise auf Sprichwörter verzeichnete.

- In der dritten Phase der Untersuchung wurden mit Hilfe des PERL-Programms die neu ermittelten Varianten und Verfremdungen in die Korpus-Suchmaschine als Suchbegriffe eingegeben. Dieses Mal wurde lediglich nach exakten Treffern gesucht. Auf diese Weise wurde die genaue Anzahl der im Korpus vorkommenden Varianten und Verfremdungen der 10 häufigsten Sprichwörter festgestellt (dazu s. 4. Ergebnisse der Untersuchung).

Im Laufe der Untersuchung sind wir auf gewisse Schwierigkeiten gestoßen, was die Suche im Korpus anbetrifft. So zum Beispiel war es nicht möglich, nach einer Phrase über die Satzgrenzen hinweg zu suchen. Das erwies sich als Hindernis besonders dann,

wenn wir auch jene Treffer berücksichtigen wollten, in denen nicht das vollständige Sprichwort zitiert wird, sondern darauf lediglich angedeutet wird. So eine Andeutung bzw. so ein Hinweis können nämlich auch über die Satzgrenze hinaus gehen und werden dementsprechend von der Korpus-Suchmaschine außer Acht gelassen.

Zum anderen haben der Bindestrich und der Gedankenstrich den Status von selbstständigen Wörtern (Token) im Korpus. Bei einer wörtlichen Suche kann das die Suchergebnisse negativ beeinflussen. Die Suchmaschine schließt Phrasen mit einem Binde- bzw. Gedankenstrich als mögliche Treffer automatisch aus, zumal sie ein „Extrawort“ enthalten und dementsprechend mit dem Suchstring nicht hundertprozentig übereinstimmen. Um diese Unzulänglichkeit zu umgehen, haben wir in unserem Suchstring an Stelle des Binde- bzw. des Gedankenstrichs ein beliebiges Wort zugelassen, z.B.:

всяко зло – за добро → всяко зло <beliebig> за добро

Somit waren die exakten Treffer vollzählig, allerdings sollten die kompletten Suchergebnisse durchgelesen werden, um die wörtlichen Entsprechungen von Varianten, Verfremdungen etc. zu trennen.

#### ***4. Ergebnisse***

Im Folgenden werden die Ergebnisse der Untersuchung kurz dargestellt und erläutert.

Die Suche nach wörtlichen Entsprechungen für die bulgarischen Sprichwörter hat ergeben, dass von 2.301 Sprichwörtern 657 im Korpus vorkommen. Die 10 Sprichwörter mit den meisten exakten Treffern sind wie folgt:

	<b>Sprichwort</b>	<b>Exakte Treffer</b>	<b>Deutsches Äquivalent/ deutsche Übersetzung</b>
1.	<i>Всяко зло за добро.</i>	179	Nach jedem Übel kommt auch etwas Gutes. (Übers.)
2.	<i>Нищо ново под слънцето.</i>	127	<i>Es gibt nichts Neues unter der Sonne.</i>
3.	<i>Целта оправдава средствата.</i>	108	<i>Der Zweck heiligt die Mittel.</i>
4.	<i>По-добре късно, отколкото никога.</i>	108	<i>Besser spät als nie.</i>
5.	<i>Съединението прави силата.</i>	106	<i>Einigkeit macht stark.</i>
6.	<i>Всяко чудо за три дни.</i>	77	Jedes Wunder dauert (höchstens) drei Tage lang. (Übers.)
7.	<i>Око за око зъб за зъб.</i>	89	<i>Auge um Auge, Zahn um Zahn.</i>
8.	<i>Кръвта вода не става.</i>	95	<i>Blut ist dicker als Wasser.</i>
9.	<i>Времето не чака.</i>	62	Die Zeit wartet auf niemanden. (Übers.)
10.	<i>Времето е пари.</i>	54	<i>Zeit ist Geld.</i>

Diese Liste zeigt große Ähnlichkeiten mit den Ergebnissen der Untersuchungen, bei denen Artikel aus der Tageszeitung *Стандарт* bzw. das WWW nach Sprichwörtern durchsucht wurden. Man kann deutliche Übereinstimmungen erkennen, was die Top-10-Liste der gebräuchlichsten bulgarischen Sprichwörter anbetrifft (s. dazu Hrisztova-Gotthardt 2010 und Hrisztova-Gotthardt/ Gotthardt 2011).

Die verfeinerte Suche, die in der zweiten Phase der Untersuchung ausgeführt wurde, hat gezeigt, dass die traditionellen (kodifizierten) Sprichwörter nicht nur wortwörtlich im Korpus enthalten sind, sondern auch in Form von qualitativen und quantitativen Varianten und Verfremdungen. Des Weiteren wird an mehreren Stellen auf das ursprüngliche Sprichwort hingewiesen, indem nur Teile davon zitiert werden. Diese Art von Treffern ist ebenfalls enorm wichtig für die Ermittlung der bekanntesten und gebräuchlichsten Texte. Die Tatsache, dass ein Sprichwort mehrere Varianten hat, oft modifiziert wird, beziehungsweise dass man gelegentlich auf seine Botschaft verweist, spricht bereits dafür, dass es mit

großer Wahrscheinlichkeit über einen hohen Bekanntheitsgrad verfügt. Aus Platzgründen werden an dieser Stelle nicht alle Varianten und Verfremdungen der Top-10-Treffer angeführt, es wird nur die Anzahl der Treffer je nach Kategorie angegeben:

		Exakte Treffer	Varianten	Verfrem- dungen	Hinweise	Gesamt- treffer
1.	<i>Всяко зло за добро.</i>	179	29	0	2	210
2.	<i>Нищо ново под слънцето.</i>	127	46	10	4	187
3.	<i>Целта оправдава средствата.</i>	108	11	27	18	164
4.	<i>По-добре късно, отколкото никога.</i>	108	1	15	6	130
5.	<i>Съединението прави силата.</i>	106	9	12	3	130
6.	<i>Всяко чудо за три дни.</i>	77	33	4	2	116
7.	<i>Око за око зъб за зъб.</i>	89	12	5	0	106
8.	<i>Кръвта вода не става.</i>	95	2	2	2	101
9.	<i>Времето не чака.</i>	62	6	0	11	79
10.	<i>Времето е пари.</i>	54	2	7	15	78

Wie aus der Tabelle zu entnehmen ist, übertrifft die Anzahl der exakten Treffer deutlich die Anzahl der Varianten. Das gilt nicht nur für die so genannten Top 10, sondern auch für alle 657 Texte, die im Korpus vorkommen. Das heißt, die Sprichwörterformen, die wir im Rahmen der Untersuchung als Grundformen der jeweiligen Sprichwörter betrachtet haben, weisen tatsächlich eine höhere Vorkommenshäufigkeit im Vergleich zu ihren Varianten auf und wurden mit Recht in den Sammlungen als Nennformen (Lemmata) kodifiziert.

Immerhin sind wir bei der Auswertung der Ergebnisse auf einen Text aufmerksam geworden, der zuerst als Variante von *Времето не чака* [Die Zeit wartet auf niemanden] kategorisiert wurde. Unsere Recherche hat ergeben, dass *Времето и приливът не чакат никого* [Die Zeit und die Flut warten auf niemanden] eigentlich die bulgarische Entsprechung für *Time and tide wait for no man* ist. Zwar wurde das Sprichwort bis dahin in keiner einzigen Sammlung kodifiziert, hat sich aber neunmal im Korpus auf-



gefunden. Es kann davon ausgegangen werden, dass sich das Proverb langsam auch in der bulgarischen Sprache einbürgert und demnach nicht als Variante, sondern als vollwertiges Sprichwort bezeichnet werden kann.

### **5. Zusammenfassung und Ausblick**

In dieser Arbeit wurden die Methoden und Ergebnisse der bereits dritten korpusbasierten Untersuchung zur Vorkommenshäufigkeit von Sprichwörtern in der modernen geschriebenen bulgarischen Sprache präsentiert. In Bezug auf die Frage nach den gebräuchlichsten Sprichwörtern haben alle drei Untersuchungen überraschend ähnliche Resultate hervorgebracht. Es hat sich dabei um drei – ihrer Größe und ihrer Zusammensetzung nach – ganz unterschiedliche Korpora gehandelt. Diese Tatsache lässt uns darauf schließen, dass die Übereinstimmungen in den Ergebnissen keineswegs zufällig sind und dass die Sprichwörter mit den meisten Treffern in den drei Korpora tatsächlich zu den bekanntesten und gebräuchlichsten Proverbien gehören und sehr hohe Chancen als Kandidaten für einen Platz im bulgarischen Sprichwortminimum haben. Dazu gehören unter anderem Texte wie: *Всяко зло за добро* [Nach jedem Übel kommt auch etwas Gutes], *По-добре късно, отколкото никога* [Besser spät als nie], *Съединението прави силата* [Einigkeit macht stark], *Нищо ново под слънцето* [(Es gibt) nichts Neues unter der Sonne], *Кръвта вода не става* [Blut ist dicker als Wasser] etc.

In naher Zukunft sollen mit Hilfe des im Rahmen dieser Arbeit dargestellten Suchverfahrens weitere korpusbasierte Untersuchungen durchgeführt werden. Das Ziel dieser Untersuchungen wird es sein, nicht nur die bekanntesten und gebräuchlichsten bulgarischen Sprichwörter zu ermitteln, sondern auch Informationen über ihre Varianten, Bedeutungen, pragmatischen Funktionen etc. Diese Daten sollen später systematisiert und in Form eines Wörterbuchs der modernen bulgarischen Sprichwörter erscheinen.

### **Anmerkungen**

<sup>1</sup> Die 2.301 Texte entstammten zwei phraseologischen Sammlungen, die in der Anfangsphase der Untersuchung immer noch als die „neuesten“ galten: Grigorov / Katarov 1986 und Vlahov 1996.

<sup>2</sup> Mehr zum Bulgarischen nationalen Korpus unter: [http://www.ibl.bas.bg/en/BGNC\\_en.htm](http://www.ibl.bas.bg/en/BGNC_en.htm).

<sup>3</sup> Der Terminus *Tokenisierung* wird in der Computer- und Korpuslinguistik verwendet und bezeichnet die Segmentierung eines Textes in kleinste Einheiten der Wortebene, auch *Token* genannt. Dabei werden bei der Zerlegung nicht nur Wörter im gängigen Sinne, sondern auch Zahlen, Satzzeichen, Klammern, Anführungsstriche und andere Symbole als Token identifiziert, die (meist) von Leerzeichen begrenzt sind (vgl. Carstensen 2004: 408f.; Lemnitzer / Zinsmeister 2010: 64f.). Unter *Lemmaisierung* wird eine Annotation auf der Wortebene verstanden, bei der die Token morphologisch analysiert und auf ihre Grundform zurückgeführt werden, d.h. Worteinheiten, die sich nur in ihren Flexionsmerkmalen unterscheiden, werden unter dem Begriff *Lemma* zusammengefasst (s. Haß 2005: 75ff.; Scherer 2006: 33).

<sup>4</sup> Die Idee zur lexikalisch-semantischen Organisation von Wörtern wurde zum ersten Mal in der lexikalischen Datenbank *WordNet* an der Princeton University realisiert (<http://wordnet.princeton.edu/perl/webwn>). Nach dem Beispiel des Englischen wurde der Wortschatz weiterer Sprachen semantisch strukturiert, darunter auch der Wortschatz des Bulgarischen ([http://dcl.bas.bg/BulNet/wordnet\\_en.html](http://dcl.bas.bg/BulNet/wordnet_en.html)).

<sup>5</sup> Mehr zur Suche im Korpus: [http://www.ibl.bas.bg/en/BGNC\\_search\\_en.htm](http://www.ibl.bas.bg/en/BGNC_search_en.htm).

<sup>6</sup> Angesichts dieser Einschränkung musste hier ein gewisser Grad an Informationsverlust in Kauf genommen werden. Dennoch haben wir uns für diese Einschränkung entschieden, und zwar aus folgendem Grund: Eine Testsuchanfrage, bei der eine Sprichwort-Konstituente durch ein beliebiges Wort ersetzt werden durfte, hat nämlich gezeigt, dass eine uneingeschränkte Suche eine Vielzahl von irrelevanten Treffern erbringt. Diese Treffer sollen alle manuell durchsucht werden, was wiederum einen enormen Zeitaufwand bedeutet. Im Vergleich dazu ist der befürchtete Informationsverlust bei einer Einschränkung der Suchbedingungen eher unwesentlich.

### **Literatur**

#### **Primärliteratur**

[Grigorov/Kazarov] Григоров, Милко / Кацаров, Костадин (1986): *Български пословици и поговорки*. Трето преработено издание. София: Издателство Наука и изкуство.

[Stojkova] Стойкова, Стефана (2007): *Български пословици и поговорки*. София: Издателство Колибри.

[Vlahov] Влахов, Сергей (1996): *Съпоставителен речник на пословици*. София: Издателство Ето.

**Sekundärliteratur**

- Baur, Rupprecht S. / Chlosta, Christoph (1996): Sprichwörter: ein Problem für Fremdsprachenlehrer wie -lerner? In: *Deutsch als Fremdsprache. Zeitschrift zur Theorie und Praxis des Deutschunterrichts für Ausländer*. 33/2, 91–102.
- Carstensen, Kai-Uwe (Hrsg.) (2004): *Computerlinguistik und Sprachtechnologie: eine Einführung*. 2., überarbeitete und erweiterte Auflage. München: Elsevier.
- Čermák, František (2003): Paremiological Minimum of Czech: The Corpus Evidence. In: Burger, Harald / Häcki Buhofer, Annelies / Gréciano, Gertrud (Hrsg.): *Flut von Texten – Vielfalt der Kulturen*. Baltmannsweiler: Schneider Verlag Hohengehren. 15–31.
- Đurčo, Peter (2006): Methoden der Sprichwortanalysen oder auf dem Weg zum Sprichwörter-Optimum. In: Burger, Harald / Häcki Buhofer, Annelies (Hrsg.): *Phraseology in Motion I*. Baltmannsweiler: Schneider Verlag Hohengehren. 3–20.
- Grzybek, Peter (1991): Sinkendes Kulturgut? Eine empirische Pilotstudie zur Bekanntheit deutscher Sprichwörter. In: *Wirkendes Wort. Deutsche Sprache und Literatur in Forschung und Lehre* 41/2, 239–264.
- Haß, Ulrike (Hrsg.) (2005): *Grundfragen der elektronischen Lexikographie: *elekxiko* - das Online-Informationssystem zum deutschen Wortschatz*. Berlin / New York: Walter De Gruyter.
- Hrisztova-Gotthardt, Hrisztalina (2010): Methoden und Ergebnisse einer korpusbasierten Untersuchung zur Vorkommenshäufigkeit bulgarischer Sprichwörter in zeitgenössischen Zeitungstexten. In: Ptashnyk, Stefaniya / Hallsteinsdóttir, Erla / Bubenhofer, Noah (Hrsg.): *Korpora, Web und Datenbanken. Computergestützte Methoden in der modernen Phraseologie und Lexikographie*. Baltmannsweiler: Schneider Verlag Hohengehren. 55–65.
- Hrisztova-Gotthardt, Hrisztalina / Gotthardt, Zoltán (2011): „Който търси, намира“ – Searching for Bulgarian Proverbs on the Web. In: *Jezikoslovlje* 12/1. (im Druck)
- Lemmitzer, Lothar / Zinsmeister, Heike (2010): *Korpuslinguistik: eine Einführung*. 2., durchgesehene und aktualisierte Auflage. Tübingen: Narr Verlag.
- [Пермяков] Пермяков, Григорий Л. (1988): *Основы структурной паремииологии*. Москва: Главная редакция восточной литературы.
- Scherer, Carmen (2006): *Korpuslinguistik*. Heidelberg: Universitätsverlag Winter.
- Tóthné Litovkina, Anna (1996): Parömiológiai felmérés Magyarországon. (Milyen formában és változatban élnek a legismertebb közmondások, és mi határozza meg az ismeretüket?). In: *Magyar nyelv* XCII/4, 439–458.

**Internetquellen**

- Bulgarisches nationales Korpus*: [http://www.ibl.bas.bg/en/BGNC\\_en.htm](http://www.ibl.bas.bg/en/BGNC_en.htm). Gesehen am 27. November 2011.
- Bulgarisches WordNet*: [http://dcl.bas.bg/BulNet/wordnet\\_en.html](http://dcl.bas.bg/BulNet/wordnet_en.html). Gesehen am 27. November 2011.
- Englisches WordNet*: <http://wordnet.princeton.edu/wordnet/>. Gesehen am 27. November 2011.

Hrisztalina Hrisztova Gotthardt  
Olga utca 1.X.32  
7632 Pécs  
Hungary  
E-mail: xpucu@freemail.hu

Zoltán Gotthardt  
Olga utca 1.X.32  
7632 Pécs  
Hungary  
E-mail: gotty@freemail.hu