

Andreas Meger

Das polnisch-deutsche / deutsch-polnische Parallelkorpus PolGerCorp

Struktur und Suchmöglichkeiten – ein Werkstattbericht

The Polish-German / German-Polish Parallel Corpus (POLGERCORP): Structure and query possibilities – Abstract

The POLGERCORP, developed 2018–2020 under the auspices of the Universities of Mainz and Warsaw (research project “The Development of the Polish Aspect System in the Last 250 Years against the Background of Neighbouring Languages” (Beethoven II DFG/NCN)), includes about 10 million tokens in texts from 1750 to 2020 in both translation directions and from various genres (fictional prose, non-fictional texts, press, legal texts). The texts are automatically tagged, lemmatized and sentence-aligned. The article describes the structure and the practical work on the corpus and focuses on a new interface. This new query interface, which is available in English, aims to be “for all” in the sense that it includes a graphical query builder and allows the user to directly input sophisticated CQP queries. Search facilities offer queries by lemma and grammatical tag, as well as the filtering of results on the basis of metadata, including, for example, a choice the selection of the source language and different genres.

1 Korpuslinguistische Ausgangsposition

Parallelkorpora¹ – also Sammlungen von genuinen Texten und deren Übersetzungen² in eine oder mehrere andere Sprache(n) – führen in der Korpuslandschaft als “jüngere und kleine Geschwister” der großen einsprachigen elektronischen (Referenz-)Korpora nach wie vor ein Nischendasein. Zwar hat die technische Entwicklung, insbesondere die Digitalisierung, die in den letzten 15–20 Jahren das Wachstum der elektronischen Textkorpora enorm begünstigt hat, auch zur Entwicklung der Parallelkorpora beigetragen, doch kommen sie weder bei Textmenge noch öffentlicher und wissenschaftlicher Wahrnehmung an die großen einsprachigen Textkorpora heran, deren Umfang inzwischen

¹ Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 380115654.

² In diesem Beitrag werden die Begriffe Übersetzung, Zieltext und zielsprachlicher Text synonym verwendet; entsprechend synonym werden auch die Begriffe Original, Ausgangstext und ausgangssprachlicher Text behandelt.

das weit über Tausendfache der Parallelkorpora umfasst.³ Der wichtigste Grund hierfür liegt klar auf der Hand: Während den einsprachigen Textkorpora – sehen wir einfachheitshalber von urheberrechtlichen Einschränkungen ab – die gesamte Fülle der in einer Sprache produzierten Texte zur Verfügung steht, können Parallelkorpora nur auf Texte zurückgreifen, für die in dem gewählten Sprachenpaar auch eine entsprechende Übersetzung existiert. Ein weiterer, vor allem den “Korpus-Laien” verborgener Grund für die geringere Anzahl und Größe der Parallelkorpora ist der mit dem Satzalignment verbundene Arbeitsaufwand, der bei der Kompilation eines Parallelkorpus hinzukommt (vgl. Meger 2016: 48–49).

Trotz der unterschiedlichen Ausgangsposition der beiden Korpusstypen ist die Wahrnehmung der Parallelkorpora in den letzten Jahren augenscheinlich gestiegen: Während beispielsweise in dem 2014 zu Korpora der Slavinen erschienenen Sammelband (Hebal-Jeziarska 2014) lediglich zwei Parallelkorpora beschrieben werden (Łaziński 2014), erscheint bereits 2016 ein Sammelband (Gruszczyńska/Leńko-Szymańska 2016), der ausschließlich Parallelkorpora mit Polnisch gewidmet ist. Dieser Sammelband veranschaulicht, dass mit neun Parallelkorpora und den dazugehörigen Projekten eine beachtliche Anzahl von Parallelkorpora mit Polnisch vorhanden ist, doch zeigt er auch, dass die Korpora eher klein, verstreut und durch ihren oft projekt- bzw. institutionsinternen und dadurch “geschlossenen” Charakter weitgehend unbekannt sind. Dabei können Parallelkorpora vor allem in den Bereichen des Sprachvergleichs, der Translationsforschung und -didaktik ein wichtiges Potential als Forschungs- und Unterrichtswerkzeug darstellen, was – neben den beiden erwähnten Sammelbänden, deren Fokus primär auf den Slavinen liegt – auch eine Reihe einschlägiger Publikationen deutlich macht (dazu vgl. Fantinuoli/Zanettin 2015).⁴

2 Das polnisch-deutsche / deutsch-polnische Parallelkorpus POLGERCORP

2.1 Genese und Zweck

Für das Sprachenpaar Polnisch/Deutsch existierte – abgesehen von polnisch-deutschen Anteilen in den multilingualen Korpora PARASOL und INTERCORP – bisher kein Parallelkorpus. Aus diesem Grund entstand die Idee, das im Rahmen des deutsch-polnischen

³ Das wird bei Betrachtung folgender, auf den deutschen und slavischen Bereich beschränkter Zahlen deutlich: Das deutsche Referenzkorpus umfasst insgesamt 46,9 Mrd. Textwörter (DEREKO, Stand 20.06.2020), das gesamte polnische Nationalkorpus 1,8 Mrd. (NKJP, Stand 20.06.2020). Demgegenüber stehen die Parallelkorpora mit Polnisch: das – größte – multilinguale INTERCORP mit insgesamt 311 Mio. Textwörtern (darunter 85 Mio. polnische Textwörter, INTERCORP, Stand 20.06.2020), das polnisch-russische, russisch-polnische Parallelkorpus mit etwa 30 Mio. Textwörtern (PoRoKR, Stand 20. Juni 2020) und das multilinguale slavische Korpus PARASOL mit insgesamt ca. 27 Mio. Textwörtern (PARASOL, Stand 20.06.2020), wobei der Anteil des Polnischen an diesem Korpus unklar ist (Meger 2016: 49).

⁴ Einen vergleichsweise vollständigen Überblick über die z. Zt. vorhandenen Parallelkorpora bietet auch die Internetseite des europäischen Forschungsverbundes CLARIN (2020) (Stand 10.08.2020).

Forschungsprojekts *Die Entwicklung des polnischen Aspektsystems in den letzten 250 Jahren auf dem Hintergrund der Nachbarsprachen*⁵ zu schaffende Parallelkorpus so zu gestalten, dass es nach Projektende der wissenschaftlichen und nichtwissenschaftlichen Öffentlichkeit kostenlos als Nachschlagewerk sowie Forschungs- und Unterrichtswerkzeug auf den Internetseiten der am Projekt beteiligten Universitäten Mainz und Warschau zur Verfügung steht.⁶

Vor diesem Hintergrund hat das POLGERCORP⁷ einen projektinternen und einen projektexternen Zweck: (a) Projektintern soll es beim Aufzeigen grammatikalisch-semanticischer Mittel helfen, die das Deutsche als “aspektlose” Sprache bei denselben funktionalen Oppositionen auf grammatischer Ebene und auf Diskursebene verwendet; (b) projektextern soll es hingegen als Forschungstool für sprach- und translationswissenschaftliche Untersuchungen und als Hilfsmittel für Wissenschaftler, Studierende und Übersetzer dienen und so für alle Interessierten zugänglich und nutzbar sein.⁸

Da sich das POLGERCORP auch an ein breites Publikum (“Korpus-Laien”) richtet, gehört zu seinen grundlegenden Annahmen, dass eine wesentliche Voraussetzung für seine Nutzung ein intuitiv bedienbares Abfrage-Interface ist, welches eine Suche ohne besondere Kenntnisse einer speziellen, korpuseigenen Abfragesyntax ermöglicht (vgl. dazu Abschnitte 2.5 und 2.6).

2.2 Struktur

Nicht zuletzt aufgrund des oben beschriebenen projektinternen Zwecks ist das POLGERCORP ein diachrones Korpus, das genuine polnische und deutsche Texte und deren Übersetzungen in die jeweils andere Sprache der letzten 270 Jahre – also ab etwa 1750 bis heute – umfasst. Es ist auf eine Gesamtmenge von etwa 10 Mio. Textwörtern (Tokens) geplant,⁹ wobei für jede Übersetzungsrichtung (Polnisch-Deutsch und Deutsch-Polnisch) jeweils 5 Mio. Textwörter veranschlagt wurden. Bei der Zusammen-

⁵ Dieses bilaterale Forschungsprojekt wird von der DFG und ihrem polnischen Pendant NCN im Rahmen von “Beethoven II” gefördert (Projektnummer 380115654) und hat eine Laufzeit von 2018 bis 2020/21. Die Arbeiten an dem hier beschriebenen Parallelkorpus werden auf polnischer Seite von Marek Łaziński (Universität Warschau) und auf deutscher Seite vom Autor dieses Beitrags geleitet.

⁶ In Vorbereitung auf das Projekt wurde 2014/15 von der Universität Mainz ein polnisch-deutsches, deutsch-polnisches Probekorpus geschaffen, das ca. 1 Mio. Textwörter enthielt und dessen grundlegende Annahmen (v. a. intuitive Bedienbarkeit und übersichtliches Abfrage-Interface) in die Gestaltung des POLGERCORP eingeflossen sind (vgl. Meger 2016, Meger/Woźniak/von Waldenfels 2016).

⁷ Das hier beschriebene Parallelkorpus wird (nur) in der englischsprachigen Version als “Polish-German / German-Polish Parallel Corpus” (POLGERCORP) im Internet bereitgestellt.

⁸ Dass dies bereits vor der offiziellen Bereitstellung des POLGERCORP für die Öffentlichkeit der Fall ist, zeigen zwei Bachelorarbeiten, die derzeit (Juni 2020) unter der Leitung des Verfassers am Arbeitsbereich Polnisch des FTSK entstehen: Eine der Arbeiten befasst sich mit theoretischen und praktischen Aspekten von Parallelkorpora mit Polnisch (u. a. auch des POLGERCORP); die andere nutzt hingegen das POLGERCORP als Materialbasis für eine translationswissenschaftliche Untersuchung.

⁹ Zum gegenwärtigen Zeitpunkt (Juni 2020) enthält das POLGERCORP 9,6 Mio. Textwörter (vgl. Tab. 1). Da die Arbeiten am Korpus derzeit noch nicht abgeschlossen sind, wird die Gesamtzahl der Textwörter voraussichtlich höher ausfallen.

setzung des Korpus gehörten (a) chronologische Ausgewogenheit und (b) Diversität hinsichtlich Textsorten und Autoren zu den Grundannahmen.

(a) chronologische Ausgewogenheit

Um die in das Parallelkorpus aufzunehmenden Texte möglichst ausgewogen auf den recht langen Zeitraum von etwa 270 Jahren zu verteilen, wurde dieser in 50-Jahres-Abschnitte eingeteilt. (Eine Ausnahme bildet der letzte Zeitraum, der ca. 70 Jahre umfasst.)¹⁰ Für jeden Zeitabschnitt und jede Übersetzungsrichtung wurden 1 Mio. Textwörter veranschlagt. Ausnahmen sollten der erste und letzte Zeitabschnitt bilden: Beim ersten Zeitabschnitt (1750–1799) mit 0,75 Mio. geplanten Textwörtern pro Übersetzungsrichtung erfolgte dies aufgrund der zu erwarteten geringeren Menge an zur Verfügung stehenden übersetzten Texten; beim letzten Zeitabschnitt (1950–2020) waren die veranschlagten 1,25 Mio. Textwörter je Übersetzungsrichtung der Länge des Zeitabschnitts geschuldet. (Auch innerhalb der 50-Jahres-Zeiträume sollten die Texte möglichst so ausgewählt werden, dass sie sich gleichmäßig auf den jeweiligen Zeitraum verteilen.)

Im Zuge der Korpusarbeiten wurden bei der Gesamttextwortzahl der jeweiligen Zeitabschnitte Abweichungen von 5 bis 10 % zu der geplanten Textwortzahl toleriert (vgl. Tab. 1), da eine Kürzung der im Laufe der Zeit bereits ins Korpus eingespeisten, z. T. sehr aufwändig bearbeiteten Texte nicht sinnvoll erschien. In solchen Fällen wurde jedoch darauf geachtet, dass beide Übersetzungsrichtungen eine annähernd gleiche Textwortzahl aufwiesen. Da die Korpusarbeiten am letzten Zeitabschnitt z. Zt. noch andauern, haben die diesen Zeitraum betreffenden Zahlen aus Tab. 1 nur eine eingeschränkte Aussagekraft. (So müssen vor allem noch belletristische Texte aufgenommen und die vorhandenen Textsorten “ausbalanciert” werden.)

¹⁰ Überlegungen, die Zeiträume an prägnante Zeitabschnitte der polnischen und deutschen Geschichte anzulehnen und in der zu berücksichtigenden Textmenge entsprechend zu gewichten, wurden aufgrund der ungleichen Länge der Zeiträume und der häufig bestehenden Unsicherheit über den Entstehungszeitpunkt eines Textes verworfen. Außerdem hätte eine solche Vorgehensweise aufgrund der zeitlichen Heterogenität der Zeiträume die im Abschnitt 2.4 beschriebene Problematik der Textbeschaffung verschärft. Die formale Lösung der 50-Jahres-Zeiträume wird durch die Möglichkeit zur individuellen Einstellung der zeitlichen Filter bei der Recherche gemildert (vgl. dazu Abschnitte 2.6, Punkt (e) Suchfilter).

		Polnisch>Deutsch		Deutsch>Polnisch		Gesamt				
		Textwörter (Tokens) und prozentualer Anteil*	Positionen	Textwörter (Tokens) und prozentualer Anteil*	Positionen	Textwörter (Tokens) und prozentualer Anteil*	Positionen			
1750–1799	literarische Texte	327.400	~44%	**6	450.400	59%	9	777.800	~51%	15
	Sach-, Presstexte	265.000	~36%	4	147.300	19%	4	412.300	~27%	8
	Rechtstexte	153.900	~21%	5	169.000	22%	7	322.900	~21%	12
	Gesamt 1750–1799	746.300		15	766.700		20	1.513.000		35
1800–1849	literarische Texte	637.300	60%	17	613.200	60%	13	1.250.500	60%	30
	Sachtexte	222.400	21%	6	202.800	20%	6	425.200	20%	12
	Rechtstexte	204.900	19%	17	206.800	20%	16	411.700	20%	33
	Gesamt 1800–1849	1.064.600		40	1.022.800		35	2.087.400		75
1850–1899	literarische Texte	601.200	57%	17	610.400	58%	10	1.211.600	57%	27
	Sachtexte	261.000	24%	5	217.000	21%	5	478.000	23%	10
	Rechtstexte	205.700	19%	7	221.500	21%	10	427.200	20%	17
	Gesamt 1850–1899	1.067.900		29	1.048.900		25	2.116.800		54
1900–1949	literarische Texte	655.100	61%	10	649.300	60%	10	1.304.400	60%	20
	Sachtexte	208.600	19%	5	228.600	21%	5	437.200	20%	10
	Rechtstexte	218.500	20%	12	209.000	19%	12	427.500	20%	24
	Gesamt 1900–1949	1.082.200		27	1.086.900		27	2.169.100		54
1950–2020	<i>Sachtexte</i> ***	162.200	~15%	10	38.400	6%	10	200.600	12%	20
	<i>Presstexte</i>	584.900	~54%	113	456.000	72%	150	1.040.900	60%	263
	<i>Rechtstexte</i>	345.300	~32%	10	140.000	22%	4	485.300	28%	14
	Gesamt 1950–2020	1.092.400		133	634.400		164	1.726.800		297
Gesamt-korpus 1750–2020	<i>literarische Texte</i>	2.221.000	44%	50	2.323.300	51%	42	4.544.300	47%	92
	<i>Sach-, Presstexte</i>	1.704.100	34%	143	1.290.100	28%	180	2.994.200	31%	323
	<i>Rechtstexte</i>	1.128.300	22%	51	946.300	21%	49	2.074.600	22%	100
	Gesamt 1750–2020	5.053.400		244	4.559.700		271	9.613.100		515

Tab. 1: Struktur des POLGERCORP (Stand 20. Juni 2020)

* Die Textwort- und Prozentzahlen wurden arithmetisch auf- bzw. abgerundet, wodurch es bei den mit einer Tilde (~) markierten Prozentzahlen zu Unschärfen kommt und ihre Summe weniger/mehr als 100 % ergibt.

** Bei der Anzahl der Positionen handelt es sich um die Anzahl der Originaltexte.

*** Alle kursiv markierten Angaben sind vorläufig, da die Arbeiten am Korpus noch nicht abgeschlossen sind.

(b) Diversität hinsichtlich der Textsorten und Autoren

Damit das Parallelkorpus möglichst unterschiedliche Textsorten enthält, wurde festgelegt, dass neben den in der Regel wesentlich längeren fiktionalen (literarischen) Texten, von denen es häufiger Übersetzungen gibt und die damit leichter zu beschaffen sind, auch ein etwa gleich großer Anteil an nicht-fiktionalen Texten und deren Übersetzungen ins Korpus aufgenommen wird. So machen z. Zt. (Juni 2020) fiktionale Texte etwa 47 % und nicht-fiktionale ca. 53 % des Korpus aus. Letztere umfassen Sach-, Presse- und Rechtstexte. Ein besonderes Augenmerk wurde dabei auf Rechtstexte gelegt, die ein gesondertes Subkorpus der Rechtssprache bilden (vgl. Abschnitt 2.3). Neben der beschriebenen Diversifizierung des Korpus hinsichtlich der Textsorten sollten möglichst viele unterschiedliche Autoren und Texte aufgenommen werden, was zur Folge hatte, dass (Ausgangs-)Texte eine Höchstlänge von 30.000 Textwörtern haben und ggf. gekürzt wurden. Lediglich beim Fehlen von Texten in einem bestimmten Zeitraum oder bei einer bestimmten Textsorte wurden vereinzelt längere Textabschnitte aufgenommen. Diese Vorgehensweise führte dazu, dass das POLGERCORP bisher insgesamt 515 Einzelpositionen (Romane, Sachbücher, Internettexte, Zeitschriften- und Zeitungsartikel sowie Gesetzes- und Verordnungstexte) aufweist. Berücksichtigt man dabei die Tatsache, dass zu jedem Text auch eine Übersetzung gehört, kann von insgesamt 1030 Einzelpositionen ausgegangen werden, die das POLGERCORP momentan enthält.

2.3 Subkorpus der Rechtssprache

Bereits in der Planungsphase des POLGERCORP war die Schaffung eines polnisch-deutschen, deutsch-polnischen Subkorpus der Rechtssprache ins Auge gefasst worden. Dieses Subkorpus – veranschlagt auf 2 Mio. Textwörter (ca. 20 % der Korpus-Texte, mit jeweils 1 Mio. Textwörter pro Übersetzungsrichtung) – macht gegenwärtig (Juni 2020) mit 2,07 Mio. Textwörtern etwa 22 % des Korpus aus und setzt sich aus 100 Positionen (zzgl. Übersetzungen) zusammen. Bei der Textauswahl wurde zunächst geplant, Texte aus unterschiedlichen Bereichen der Rechtssprache – also nicht nur Gesetze und Verordnungen, sondern auch privatrechtliche Urkunden und Dokumente – aufzunehmen (vgl. Meger 2016: 51–52). Erkenntnisse aus den Vor- und Rechercharbeiten führten jedoch zu der Entscheidung, dass in das POLGERCORP fast ausschließlich Gesetze, Verordnungen und Vorschriften aufgenommen werden sollen.¹¹ Diese stammen aus den Bereichen des Straf-, Zivil-, Handels- und Verwaltungsrechts. Territorial beschränkte sich die Auswahl primär auf Polen und Deutschland mit ihren historischen Teil- und

¹¹ Folgende Gründe sprachen für die Fokussierung auf diese wenigen juristischen Textsorten: Zunächst verkörpern Gesetze und Verordnungen eine abstraktere und “reiner” Form der Rechtssprache als beispielsweise privatrechtliche Verträge oder andere juristische “Gebrauchstexte”, die “individueller” Natur sind. (In Letztere fließen ohnehin häufig hohe Anteile aus Gesetzen und Verordnungen ein, vgl. z. B. Gesellschaftsverträge). Ein gewichtiges Argument für die getroffene Entscheidung war auch die praktische Tatsache, dass Gesetze und Verordnungen und ihre Übersetzungen publiziert und so besser zugänglich sind (Meger 2018: 397).

Vorgängerstaaten (Preußen, Herzogtum Warschau, Königreich Polen). In Ermangelung anderer Rechtstexte wurden für die 2. Hälfte des 19. Jh. auch Gesetze und Verordnungen aus dem zweisprachigen “Landes-Gesetz- und Verordnungsblatt für das Königreich Galizien und Lodomerien” aufgenommen (insgesamt 6,5 % aller Rechtstexte im Korpus). Hier bestand vielfach das Problem der eindeutigen Identifizierung der Ausgangssprache. Nicht aufgenommen wurden die reichlich vorhandenen und leicht zugänglichen EU-Rechtstexte, da diese keine Übersetzungen aus dem Polnischen bzw. aus dem Deutschen in die jeweils andere Sprache sind.

2.4 Textbeschaffung

Die bereits gemachte Feststellung, wonach das Vorhandensein entsprechender Texte und ihrer Übersetzungen entscheidende Auswirkungen auf die Zusammensetzung und Ausgewogenheit eines Parallelkorpus und somit auf seine Brauchbarkeit als Forschungsinstrument hat, wurde auch bei der Textbeschaffung für das POLGERCORP deutlich. Im ersten Schritt mussten Texte, die für einen bestimmten Zeitabschnitt infrage kamen, bibliografiert werden, was trotz der Digitalisierung von Quellen und Bibliothekskatalogen mit einem nicht unerheblichen Zeit- und Arbeitsaufwand verbunden war. Die Recherche nach polnischen und deutschen Texten, die in den letzten 270 Jahren ins Deutsche bzw. Polnische übersetzt wurden, spiegelte nicht nur die Zeiten der gegenseitigen positiven und negativen Wahrnehmung von Polen und Deutschen wider, sondern förderte auch Autoren, Werke und Übersetzungen zutage, die inzwischen vergessen sind.¹²

Die bibliografierten älteren, nicht urheberrechtlich geschützten Positionen wurden teilweise als digitalisierte Werke (Bücher, Zeitschriften und Zeitungen) aus Quellen wie Wikiźródła/Wikisource, Gutenbergprojekt, Deutsches Textarchiv, Polona, Federacja Bibliotek, Google Books besorgt. Ein großer Teil der älteren wie neueren, digital nicht zugänglichen Werke musste aus Bibliotheken bestellt und digitalisiert (vor allem OCR-erkannt) werden. Als technische Probleme erwiesen sich vor allem die unterschiedlichen Formate und starken Qualitätsunterschiede der vorhandenen Digitalisate. Inhaltlich problematisch waren freie, stark verkürzende Übersetzungen, die in Wirklichkeit Nachdichtungen waren, was häufig erst bei genauer Analyse der Texte festgestellt werden konnte; außerdem erwies sich Literatur in Versform (Lyrik, Dramen) als ungeeignet für ein satzaligniertes Parallelkorpus.¹³ Bei neueren, nach etwa 1930/40 entstandenen

¹² Dies gilt vor allem für frühere Texte aus dem Bereich der Trivial-/Unterhaltungsliteratur und des Rechts. Während Erstere heute kaum rezipiert werden, da sie unserem Zeitgeist nicht mehr entsprechen, sind Letztere inzwischen vergessen, da sie ihre Rechtskraft verloren haben und als Rechtsnorm nicht mehr angewandt werden. Welche Fülle von inzwischen “vergessenen”, aber korpusrelevanten Texten und ihren Übersetzungen es geben kann, verdeutlicht exemplarisch die detaillierte Untersuchung zur Geschichte der polnisch-deutschen Rechtsübersetzung von Meger (2018), die allein für die Übersetzungsrichtung Polnisch-Deutsch – und ohne dabei den Anspruch auf Vollständigkeit zu erheben – für den Zeitraum von 1750 bis heute etwa 86 Gesetzesübersetzungen auflistet.

¹³ Dies schränkte z. B. die Auswahl vor allem bei literarischen Texten aus der 2. Hälfte des 18. Jh. und den ersten Jahrzehnten des 19. Jh. so stark ein, dass die geplante Anzahl von 450.000 Textwörtern bei polnisch-deutschen literarischen Texten nicht erreicht werden konnte und der Anteil der anderen

Texten und Übersetzungen erwies sich das sowohl für Originale wie Übersetzungen geltende Urheberrecht als starkes Hemmnis bei der Textauswahl und -beschaffung. Es verursachte arbeitsaufwändige Recherchen nach den (heutigen) Urheberrechtsinhabern und machte Vereinbarungen zur Nutzung der Texte im POLGERCORP notwendig.

In seltenen Fällen, in denen mehrere Übersetzungen eines Textes vorlagen, was meist bei literarischen und juristischen Werken der Fall war, wurde nur eine Übersetzung ins Parallelkorpus aufgenommen. Bei vorhandenen mehreren Übersetzungen wurde – nicht zuletzt wegen der Urheberrechtsproblematik – auf ältere, möglichst nicht rechtgeschützte Werke zurückgegriffen. Dadurch wurden seltenere, häufig nicht mehr bekannte Übersetzungen aufgenommen, die zudem auch näher am Entstehungs-/Publikationszeitpunkt des Originals lagen und so auch die Zielsprache in einem dem Original näheren zeitlichen Stadium abbildeten.¹⁴ Auch bei den Originalen war das Ziel, möglichst die älteste zugängliche Version zu berücksichtigen, was sich aber nicht immer realisieren ließ. Alle Texte wurden in ihrer Vorlageform übernommen, das heißt, dass sowohl polnische wie deutsche Texte mit unterschiedlichen Orthografien im Parallelkorpus vorhanden sind.¹⁵

Als Übersetzer sollten nach Möglichkeit zielsprachliche Muttersprachler fungieren, was sich bei der Übersetzungsrichtung Deutsch-Polnisch ausnahmslos verwirklichen ließ; bei der Übersetzungsrichtung Polnisch-Deutsch ergab sich vereinzelt aus der Vita der Übersetzer, dass es polnische Muttersprachler waren, die das Deutsche mit fast muttersprachlicher Kompetenz beherrschten.¹⁶

2.5 Technische Aspekte: Annotation, Alignment, Abfrage-Interface

Das POLGERCORP übernimmt viele technische Lösungen seines „Vorläufers“, des bereits oben erwähnten, 2014/15 an der Universität Mainz entwickelten Probekorpus. So erfolgte die morphosyntaktische Annotation bei polnischen Texten auch im POLGERCORP automatisch mit Hilfe des Tagsets, das für das polnische Nationalkorpus entwickelt wurde (NKJP Tagset 2020); bei den deutschen Texten wurde das Tagset der Universitäten Stuttgart und Tübingen STTS verwendet (Tree Tagger 2020). Um eine Nutzung des POLGERCORP für den oben genannten projektinternen Zweck zu ermöglichen, wurde ein besonderes Augenmerk auf die Annotation des Verbalaspekts in den polnischen Texten gelegt. Hier erfolgte eine gesonderte halbautomatische Annotation der Aspektpartner, die es ermöglicht, Suchabfragen nur unter aspektpaarigen Verben durchzuführen (dazu vgl. Łaziński 2020: 268–271).

Textsorten (Sach- und Presstexte) in diesem Zeitraum erhöht werden musste (vgl. dazu oben Abschnitt 2.2 und Tab. 1).

¹⁴ Grundsätzlich ist die Aufnahme von mehreren Übersetzungen eines Textes (z. B. im Rahmen einer künftigen Erweiterung des Korpus) möglich.

¹⁵ In die aufgenommenen Texte wurde so wenig wie möglich eingegriffen; lediglich offensichtliche Druckfehler wurden, dort, wo sie aufgefallen sind, korrigiert.

¹⁶ Bei einer recht beträchtlichen Anzahl von Übersetzungen – vor allem bei Rechtstexten – ließen sich die Übersetzer nicht ermitteln.

Das Alignment erfolgte ausschließlich auf Satzebene. Auf ein Alignment auf Wortebene (Wordalignment) – wie es das Mainzer Probekorpus aufwies – wurde verzichtet, da sich dieses nicht nur als zeit- und arbeitsaufwändig, sondern auch als zu fehlerbehaftet erwies (Meger 2016: 63).

Das Abfrage-Interface des POLGERCORP wurde von Informatikern des Instituts für Polnische Sprache der Polnischen Akademie der Wissenschaften (IJP PAN) entwickelt. Es basiert – wie auch das Interface des Mainz-Germersheimer Probekorpus – auf dem als Open Source angebotenen Softwarepaket PARAVOZ, welches 2006–2014 von Meyer, Waldenfels und Zeman für das multilinguale Parallelkorpus PARASOL entwickelt wurde (Meger/Woźniak/von Waldenfels 2016: 100).

2.6 Wichtigste Funktionen des Abfrage-Interface

Ein zentraler Punkt des in Abschnitt 2.1 beschriebenen projektexternen Zwecks des POLGERCORP (Ausrichtung an ein breites Benutzerpublikum, “Korpus-Laien”) ist ein bedienfreundliches Abfrage-Interface. Dazu gehören vor allem Aspekte wie eine intuitive Bedienbarkeit mit Hilfe von Eingabefeldern, Check-Boxen, Autovervollständigungs- und Drop-down-Menüs, die die Formulierung einer Suchabfrage ohne besondere Kenntnisse einer speziellen, korpuseigenen Abfragesyntax ermöglichen. Gegenwärtig (Juni 2020) liegt das Abfrage-Interface in einer Betaversion vor: POLGERCORP (2020). Die endgültige Version des POLGERCORP (Abfrage-Interface und Korpusdateien) wird zum Projektende auf den Servern der Universitäten Mainz und Warschau der Öffentlichkeit zur Verfügung stehen.

In den nachfolgenden Abschnitten werden die wichtigsten Funktionen des Abfrage-Interface dargestellt.

(a) Aufbau des Abfrage-Interface

Das englischsprachig gestaltete Abfrage-Interface (vgl. unten Abb. 1) besteht aus zwei Teilfenstern und enthält Abfragemöglichkeiten für Polnisch und Deutsch (*Query for Polish*, *Query for German*). Auf der linken Seite ist die erste Sprache (*Primary Language*) positioniert, die für die Darstellung (u. a. Sortierung) der Ergebnisse maßgeblich ist. Sie kann über das “Fähnchen”-Symbol (oben rechts) gewechselt werden. Die erste Sprache entspricht nicht grundsätzlich der Sprache des Originals. Sollen auf der linken Seite nur originalsprachliche Texte berücksichtigt werden, kann dies unter *Filters* über die Funktion *Source language* eingestellt werden. Die Möglichkeit, bei der Suche die Sprache des Originals explizit bestimmen und entsprechende Ergebnisse herausfiltern zu können, schien vor allem aus translationswissenschaftlicher Sicht wichtig.

Jede Sprache enthält eine einfache Suche (*Basic Search*) und eine Corpus-Query-Processor-Suche (*CQP Search*).

The screenshot displays the search interface for the Polish-German / German-Polish Parallel Corpus (POLGERCORP). The interface is divided into two main sections: 'Query for Polish' and 'Query for German'. At the top, there are navigation tabs for 'Search', 'Frequency', 'Collocations', and 'N-grams'. The 'Query for Polish' section includes a 'Basic search' area with fields for 'Token:', 'Lexeme:', and 'Gram. tag:'. The 'Token:' field contains 'asp. value', 'Lexeme:' contains 'asp. determiner', and 'Gram. tag:' contains 'Superlemma'. There are also checkboxes for 'begins with', 'ends with', and 'case sensitive'. Below this is a 'Search only in:' section with buttons for 'Fiction', 'Non-fiction', 'Press texts', and 'Law texts'. A 'Filters' section is also present. The 'Query for German' section includes an 'exclude (find everything except query match)' checkbox, and checkboxes for 'begins with', 'ends with', and 'case sensitive'. At the bottom, there is a 'CQP Search' section with a search input field. The interface also features a 'Primary language' dropdown and a settings gear icon.

Abb. 1: Abfrage-Interface des POLGERCORP

(b) Einfache Suche (*Basic Search*)

Die einfache Suche kann über drei Eingabefenster mit Drop-down-Menü erfolgen: Es kann nach einer bestimmten Wortform, einem Textwort (*Token*), nach allen Formen eines Lexems (*Lexeme*) und nach grammatikalischen Tags aus den Tagsets der jeweiligen Sprache (*Gram. tag*) gesucht werden. Beim Polnischen ist die einfache Suche mit drei Eingabefeldern (*asp. value*, *asp. determiner*, *Superlemma*) zur Suche nach Aspektpartnern ausgestattet, was dem projektinternen Zweck des POLGERCORP geschuldet ist. Zusätzlich kann bei beiden Sprachen eine Suche nach Teilen von Wörtern (u. a. Präfixen, Suffixen, Kompositagliedern, Wortteilen) (Checkbox: *begins/ends with*) und mit Hilfe der Groß-/Kleinschreibung (Checkbox: *case sensitive*) durchgeführt bzw. verfeinert werden.

(c) CQP-Suche (*CQP Search*)

Die CQP-Suche ist mit der einfachen Suche gekoppelt. Alle Eingaben, die über die Eingabefenster der einfachen Suche und durch die Filtersetzung erfolgen, werden in einen CQP-Befehl "übersetzt" (vgl. Abb. 2, wo bei einer Mehrwortsuche die über die Eingabefenster eingegebenen Suchparameter automatisch als CQP-Suchanfrage erscheinen). Diese CQP-Suchanfrage kann zusätzlich manuell modifiziert werden. Gleichzeitig kann hier von Benutzern, die die CQP-Abfragesyntax beherrschen, auch eine direkte, teils sehr komplexe Suchanfrage – ohne Nutzung der Eingabefenster – formuliert werden. Diese Verzahnung der einfachen mit der CQP-Suche kann eventuell zu einem Lerneffekt führen und bei "Korpus-Laien" das Interesse wecken, die CQP-Abfragesyntax zu verstehen und ggf. zu erlernen, um so mit der Zeit auch selbst komplexere Suchanfragen zu formulieren, als es die vorgegebenen Eingabefenster erlauben.

(d) Mehrwortsuche

Bei der einfachen Suche kann auch eine Suchanfrage für zwei oder mehr Wörter gleichzeitig formuliert werden. Für die Mehrwortsuche wird mit Hilfe des Plus-Symbols (+) eine zusätzliche einfache Suche geöffnet, die über alle Eingabemöglichkeiten wie die erste Suche verfügt (vgl. Abb. 2). Hier kann die Suche für ein zweites oder ein weiteres Wort formuliert werden. Außerdem kann der bei der Suche zu berücksichtigende Abstand zwischen den gesuchten Wörtern eingestellt werden. Letzteres erleichtert die Suche nach Phraseologismen und im Polnischen häufig vorkommenden Mehrwortbenennungen.

The screenshot shows the search interface for the POLGERCORP corpus. At the top, there is a green header with the text "Query for German" and a flag icon. Below this, the "Basic search" section contains two search boxes. The first search box has the token "Kopf" entered, and the second has "stellen". Both boxes have a "Gram. tag:" field to their right. Below the search boxes, there are checkboxes for "begins with", "ends with", and "case sensitive". A "Tokens in between:" field is set to 0 and 3. Below the search boxes, there are buttons for "Fiction", "Non-fiction", "Press texts", and "Law texts". At the bottom, there is a "Filters" section with a plus sign icon, and a "CQP Search" section with a search box containing the query: "[lemma="Kopf"] [{0,3} [lemma="stellen"]]".

Abb. 2: Mehrwortsuche im POLGERCORP

(e) Suchfilter (*Filters*)

Das Abfrage-Interface verfügt über eine Reihe von Filtern, mit deren Hilfe eine Suche verfeinert/eingeschränkt werden kann. Standardmäßig werden Textsortenfilter angezeigt, die die Auswahl einer bestimmten Textsorte (*Fiction*, *Non-fiction*, *Press texts*, *Law texts*) ermöglichen. Zusätzlich kann über das Plus-Symbol bei *Filters* (+) eine Reihe von weiteren Filtern geöffnet und über Drop-down-Menüs gesetzt werden (u. a. *Author*, *Translator*, *Title*, *Work*) (vgl. Abb. 3). Diese Filter ermöglichen es, die Suche auf bestimmte Autoren, Übersetzer, Titel/Werke usw. einzuschränken. Als besonders wichtig wären hier zwei Filter zu nennen: die bereits beschriebene Funktion, die die Suche auf eine bestimmte Originalsprache einschränken lässt (*Source language*), und die Möglichkeit, in bestimmten, individuell gewählten Zeiträumen zu suchen (*Creation year (from) / (to)*).

Search only in:		
Fiction	Non-fiction	Press texts
	Law texts	

Filters		
Author	Title (Polish)	Title (German)
Translator	Work (Polish)	Work (German)
Type	Creation year (from)	Creation year (to)
Translation year	Publication year (Polish)	Publication year (German)
Source language		

Abb. 3: Suchfilter im PolGerCorp

(f) Darstellung der Suchergebnisse und Export-Funktion

Standardmäßig werden die Suchergebnisse chronologisch sortiert (vom jüngsten bis zum ältesten Textbeleg) und ohne Metadaten angezeigt (vgl. Abb. 4). Die Sortierung der Ergebnisse kann über das oben rechts positionierte Feld *Actions* individuell nach allen Parametern, die in den Metadaten enthalten sind (z. B. Autor, Titel, Originalsprache usw.), neu vorgenommen werden. Die Anzeige der Metadaten kann oben über das ebenfalls im Feld *Actions* positionierte Symbol “i” global (*Show metadata (global)*) oder individuell für jeden Textbeleg (“i”-Symbol neben jedem Textbeleg) aktiviert werden.

Die Ergebnisse einer Suche können im tsv-Format exportiert und u. a. in Excel (Microsoft Office) bzw. Calc (Apache OpenOffice) importiert und dort individuell weiter bearbeitet werden.

Results for query: [lemma="Kopf"] [0,3] [lemma="stellen"]

Number of results: 8
 Primary language: German

Show metadata (global) Actions
ⓘ A ↕ ↻ ⬇ ⬇ ⬇ ↔

German	Polish
<p>Wenn man alle die anwesenden „ Säulen “ des Orleanismus, die Herren Thiers, Remusat, Duvergier, Broglie, Piskatory, Lasteyrie, Kousin, Quinette, Beaumont und noch ein ganzes Schock anderer sammt und sonders auf den Kopf gestellt hätte, es würde aus keiner dieser Säulen ein rettendes Wort, geschweige eine rettende That herausgefallen sein.</p> <p>Title: 1848. Ein weltgeschichtliches Drama. Bd. 1 Publication year: 1875 Author: Scherr, Johannes Creation year: 1875 Type: Fiction Original language: DE</p>	<p>Gdyby wszystkie zebrane tu kolumny orleanizmu: pan Thiers, Remusat, Duvergier, Broglie, Piscatory, Lasteyrie, Cousin, Quinette i jeszcze kopę innych postawiono na głowach, to żaden nie wydałby słowa zbawczego, nie mówiąc już o zbawczym czynie. ⓘ <input type="checkbox"/></p> <p>Title: Rok 1848. Dramat dziejowy. T. 1 Publication year: 1903 Translator: Wernic, Henryk Translation year: 1903</p>
<p>Man kann wohl das ganze Land auf den Kopf stellen wollen — aber es ist doch schade, ein paar Zehnt Kutschpferde dafür zu opfern!</p>	<p>Można chcieć cały kraj do góry nogami wyrócić, – a przecież szkoda na to poświęcić choćby kilkadziesiąt cugantów!! ⓘ <input type="checkbox"/></p>
<p>Sieh mal einer an, da sind die dummen Viecher aufsässig geworden und glauben, dass sie die ganze Welt auf den Kopf stellen können!</p>	<p>Zbuntowały się bydłaki i myśla, że cały świat przewrócą do góry nogami. ⓘ <input type="checkbox"/></p>
<p>Auf der einen Seite sind in dem halben Jahrhundert, das an Zeit vergangen ist, seit Aleksander Ford seine „ Kreuzritter “ gedreht hat, die Welt, Europa, Deutschland und Polen komplett auf den Kopf gestellt worden.</p>	<p>Z jednej strony przez pół wieku, jakie minęło od czasu, gdy Aleksander Ford kręcił „ Krzyżaków ”, świat, Europa, Niemcy i Polska wyróciły się do góry nogami. ⓘ <input type="checkbox"/></p>
<p>Deutschland verdanke ich im Gegenzug etwas anderes, aber genauso Wichtiges und Kostbares, was meine Denkweise absolut auf den Kopf gestellt hatte.</p>	<p>Niemcom zawdzięczam natomiast coś zupełnie innego, ale tak samo ważnego i drogiego, coś, co mój sposób myślenia całkowicie postawiło na głowie. ⓘ <input type="checkbox"/></p>

Abb. 4: Darstellung der Suchergebnisse im PolGerCorp

3 Schlussbemerkung

Wie bereits anfangs angemerkt, wird das POLGERCORP nach Abschluss des Projekts auf den Internetseiten der Universitäten Mainz und Warschau kostenlos verfügbar sein und von allen Interessierten benutzt werden können. Perspektivisch stehen die Autoren des Korpus in erster Linie vor der Aufgabe, die technische Zugänglichkeit zu gewährleisten, was angesichts des institutionellen Hostings auf den Internetseiten der beiden Universitäten positiv gesehen wird. Wünschenswert wäre auch eine künftige Aktualisierung und Erweiterung des Korpus oder seiner Teile im Rahmen weiterer bi- und multilateraler Projekte, z. B. der Ausbau des Subkorpus der Rechtssprache oder die Aufnahme mehrerer Übersetzungen eines Textes, was direkte Übersetzungsvergleiche ermöglichen würde.

Korpora und Literatur

Korpora

[Clarin] Parallel corpora (2020) – <https://www.clarin.eu/resource-families/parallel-corpora> (02.11.2020)

DeReKo = Das deutsche Referenzkorpus DeReKo – <https://www1.ids-mannheim.de/kl/projekte/korpora.html?L=> (20.06.2020)

INTERCORP = INTERCORP. Projekt paralelních korpusů. – <https://intercorp.korpus.cz/> (20.06.2020)

NKJP = Narodowy Korpus Języka Polskiego – <http://nkjp.pl/> (20.06.2020)

[NKJP Tagset] 2 Tagset (2020) – <http://nkjp.pl/poliqarp/help/ense2.html> (02.11.2020)

PARASOL = ParaSol. A parallel corpus of Slavic and other languages – <http://parasolcorpus.org/> (20. Juni 2020)

POLGERCORP = Polish-German / German-Polish Parallel Corpus – <http://parasolcorpus.org/Beethoven/#!/> (20.06.2020)

POROKR = Polsko-rosyjski i rosyjsko-polski korpus równoległy – <http://pol-ros.polon.uw.edu.pl/index.php?id=01&lang=pl> (20.06.2020)

[Tree Tagger] TreeTagger – a part-of-speech tagger for many languages (2020) – <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/> (02.11.2020)

Forschungsliteratur

Fantinuoli, Claudio; Federico Zanettin (2015): "Creating and using multilingual corpora in translation studies." Claudio Fantinuoli, Federico Zanettin (eds): *New directions in corpus-based translation studies*. Berlin: Language Science Press, 1–10 – <https://langsci-press.org/catalog/view/76/64/287-1> (10.08.2020)

Gruszczyńska, Ewa; Agnieszka Leńko-Szymańska (Hg.) (2016): *Polskojęzyczne korpusy równoległe = Polish-language parallel corpora*. Warszawa: Uniwersytet Warszawski, Wydział Lingwistyki Stosowanej, Instytut Lingwistyki Stosowanej – https://rownolegle.ils.uw.edu.pl/files/2016/03/0000_Korpusy.pdf (20. Juni 2020)

- Hebal-Jezierska, Milena (Hg.) (2014): *Praktyczny przewodnik po korpusach języków słowiańskich*. Warszawa: Uniwersytet Warszawski, Wydział Polonistyki, Instytut Sławistyki Zachodniej i Południowej, Zespół Lingwistyki Korpusowej Języków Słowiańskich – https://www.researchgate.net/profile/Milena_Hebal-Jezierska/publication/274510797_Praktyczny_przewodnik_po_korpusach_jezykow_slowianskich/links/552187f10cf2f9c130528349/Praktyczny-przewodnik-po-korpusach-jezykow-slowianskich.pdf (20.06.2020)
- Łaziński, Marek (2014): “Praktyczny przewodnik po korpusach równoległych. Wiadomości wstępne. Korpus ParaSol i Korpus Polsko-Rosyjski UW.” Milena Hebal-Jezierska (Hg.): *Praktyczny przewodnik po korpusach języków słowiańskich*. Warszawa: Uniwersytet Warszawski, Wydział Polonistyki, Instytut Sławistyki Zachodniej i Południowej, Zespół Lingwistyki Korpusowej Języków Słowiańskich, 198–206 – https://www.researchgate.net/profile/Milena_Hebal-Jezierska/publication/274510797_Praktyczny_przewodnik_po_korpusach_jezykow_slowianskich/links/552187f10cf2f9c130528349/Praktyczny-przewodnik-po-korpusach-jezykow-slowianskich.pdf (20.06.2020)
- Łaziński, Marek (2020): *Wykłady o aspekcie polskiego czasownika*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego – <https://www.bijp.uw.edu.pl/files/Lazinski2020.pdf> (20.06.2020)
- Meger, Andreas (2016): “Über die Arbeit am Deutsch-Polnischen / Polnisch-Deutschen Parallelkorpus und seinem Subkorpus der Rechtssprache.” Ewa Bağlajewska-Miglus, Thomas Vogel (Hg.): *Fachsprachenunterricht – Lehren und Lernen am Beispiel des Polnischen als Fremdsprache / Nauczanie języka specjalistycznego – Nauczanie i uczenie się na przykładzie języka polskiego jako obcego*. (Polnisch als Fremd- und Zweitsprache 2.) Herzogenrath: Shaker, 48–65
- Meger, Andreas (2018): “Zur Geschichte der polnisch-deutschen Gesetzesübersetzung: Ein Überblick vom Ende des 18. Jh. bis in die Gegenwart.” *Zeitschrift für Slawistik* 63 [3]: 395–424

trans-kom**ISSN 1867-4844**

trans-kom ist eine wissenschaftliche Zeitschrift für Translation und Fachkommunikation.

trans-kom veröffentlicht Forschungsergebnisse und wissenschaftliche Diskussionsbeiträge zu Themen des Übersetzens und Dolmetschens, der Fachkommunikation, der Technikkommunikation, der Fachsprachen, der Terminologie und verwandter Gebiete.

Beiträge können in deutscher, englischer, französischer oder spanischer Sprache eingereicht werden. Sie müssen nach den Publikationsrichtlinien der Zeitschrift gestaltet sein. Diese Richtlinien können von der **trans-kom**-Website heruntergeladen werden. Alle Beiträge werden vor der Veröffentlichung anonym begutachtet.

trans-kom wird ausschließlich im Internet publiziert: <http://www.trans-kom.eu>

Redaktion

Leona Van Vaerenbergh
University of Antwerp
Arts and Philosophy
Applied Linguistics / Translation and Interpreting
O. L. V. van Lourdeslaan 17/5
B-1090 Brussel
Belgien
Leona.VanVaerenbergh@uantwerpen.be

Klaus Schubert
Universität Hildesheim
Institut für Übersetzungswissenschaft
und Fachkommunikation
Universitätsplatz 1
D-31141 Hildesheim
Deutschland
klaus.schubert@uni-hildesheim.de

Meger, Andreas; Michał Woźniak, Ruprecht von Waldenfels (2016): “Jak stworzyć korpus ‘dla wszystkich’? O pracy nad Polsko-Niemieckim i Niemiecko-Polskim Korpusem Równoległym / How to create a parallel corpus ‘for all’? About the building of the Polish-German and German-Polish Parallel Corpus.” Ewa Gruszczyńska, Agnieszka Leńko-Szymańska (Hg.): *Polskojęzyczne korpusy równoległe / Polish-language parallel corpora*. Warszawa: Uniwersytet Warszawski, Wydział Lingwistyki Stosowanej, Instytut Lingwistyki Stosowanej, 97–118 – https://rownolegle.ils.uw.edu.pl/files/2016/03/0000_Korpusy.pdf (20.06.2020)

Autor

Dr. Andreas Meger arbeitet seit 2003 als akademischer Mitarbeiter am Arbeitsbereich Polnisch des Germersheimer Fachbereiches Translations-, Sprach- und Kulturwissenschaft der Universität Mainz, wo er 2010 mit einer Arbeit über die Makro- und Mediostruktur von Neologismenwörterbüchern der Slawinen und des Deutschen promoviert wurde. Er ist u. a. Mitautor des 2007 erschienenen *Polnisch-deutschen Wörterbuches der Neologismen*. Seine didaktischen und wissenschaftlichen Interessen liegen im Bereich der Neologie und Neografie, der Fachsprachen, der Fachübersetzung (insbesondere der Rechtssprache) und der Korpuslinguistik (hier vor allem die Parallelkorpora). Seit 2018 arbeitet er an dem deutsch-polnischen Projekt “Die Entwicklung des polnischen Aspektsystems in den letzten 250 Jahren auf dem Hintergrund der Nachbarsprachen” mit (gefördert von der DFG und dem polnischen NCN), im Rahmen dessen das oben beschriebene polnisch-deutsche / deutsch-polnische Parallelkorpus (POLGERCORP) entsteht.
E-Mail: meger@uni-mainz.de

Neu bei Frank & Timme

TRANSÜD. Arbeiten zur Theorie und Praxis des Übersetzens und Dolmetschens

Herausgegeben von Prof. Dr. Klaus-Dieter Baumann, Prof. Dr. Dr. h.c. Hartwig Kalverkämper, Prof. Dr. Klaus Schubert

Annegret Sturm: **Theory of Mind in Translation.**
ISBN 978-3-7329-0492-1

Brita Dorer: **Advance Translation as a Means of Improving Source Questionnaire Translatability?** Findings from a Think-Aloud Study for French and German. ISBN 978-3-7329-0594-2

Anu Viljanmaa: **Professionelle Zuhörkompetenz und Zuhörfilter beim Dialogdolmetschen.** ISBN 978-3-7329-0719-9

Johan Franzon/Annjo K. Greenall/Sigmund Kvam/Anastasia Parianou (eds.): **Song Translation: Lyrics in Contexts.**
ISBN 978-3-7329-0656-7

Kommunikation – Partizipation – Inklusion

Herausgegeben von Jun.-Prof. Dr. Bettina M. Bock, Prof. Dr. Ulla Fix, Prof. Dr. Nathalie Mälzer

Sandra Hinckers: **Lateinische Übersetzungsreflexion in der Römischen Antike.** Von Terenz bis zur *Historia Augusta*. 210 Seiten. ISBN 978-3-7329-0614-7

Easy—Plain—Accessible

Herausgegeben von Prof. Dr. Silvia Hansen-Schirra, Prof. Dr. Christiane Maaß

Silvia Hansen-Schirra/Christiane Maaß (eds.): **Easy Language Research: Text and User Perspectives.**
ISBN 978-3-7329-0688-8

Christiane Maaß: **Easy Language – Plain Language – Easy Language Plus.** Balancing Comprehensibility and Acceptability. ISBN 978-3-7329-0691-8

Alle Bücher sind auch als E-Books erhältlich.

Barbara Ahrens/Morven Beaton-Thome/Monika Krein-Kühle/Ralph Krüger/Lisa Link/Ursula Wienen (Hg./eds.): **Interdependenzen und Innovationen in Translation und Fachkommunikation / Interdependence and Innovation in Translation, Interpreting and Specialised Communication.**
ISBN 978-3-7329-0566-9

Barbara Ahrens/Silvia Hansen-Schirra/Monika Krein-Kühle/Michael Schreiber/Ursula Wienen (Hg./eds.): **Translation – Kunstkommunikation – Museum / Translation – Art Communication – Museum.**
ISBN 978-3-7329-0716-8

schicht Translationswissenschaft Romanistik Mittelalterswissenschaft Kunstwissenschaft Altertumswissenschaft Sprachwissenschaft Fachsprachenforschung Pädagogik Philosophie Romanistik Slawistik Musikwissenschaft Literaturwissenschaft Musikwissenschaft Altertumswissenschaft Kulturwissenschaft Kommunikationswissenschaft Medienwissenschaft Kunstwissenschaft Theologie Religionswissenschaft Geschichtswissenschaft Philosophie Theaterwissenschaft Archäologie Philologie Politikwissenschaft Musikwissenschaft Romanistik Translationswissenschaft Sprachwissenschaft Sozialpädagogik Erziehungswissenschaft Slavistik Fachsprachenforschung Kunstwissenschaft Romanistik Slawistik Literaturwissenschaft Theaterwissenschaft Musikwissenschaft Altertumswissenschaft Kommunikationswissenschaft Medienwissenschaft Theologie Religionswissenschaft Geschichtswissenschaft Philosophie Theaterwissenschaft Archäologie Philologie Politikwissenschaft Soziologie Sozialpsychologie Erziehungswissenschaft Translationswissenschaft Sprachwissenschaft Fachsprachenforschung Kunstwissenschaft Philosophie Romanistik Slawistik Soziologie

F Frank & Timme