

Emmerich KELIH – Jürgen FUCHSBAUER –
Stefan Michael NEWERKLA (Hrsg.)

Lehnwörter im Slawischen:
Empirische und crosslinguistische
Perspektiven

Wien 2015

Inhaltsverzeichnis

| | |
|--|-----|
| Einleitung | 3 |
| EMMERICH KELIH (Wien) — Probleme der empirischen Lehnwortforschung: Lehnwörter im Basiswortschatz (Slowenisch) und ein frequenzbasierter Ansatz (Kroatisch)..... | 10 |
| STEFAN MICHAEL NEWERKLA (Wien) — Kernproblembereiche der Lehnwortforschung anhand von Beispielen aus dem Kontaktbereich Deutsch, Tschechisch und Slowakisch..... | 37 |
| CATHARINA STAMATA KREBS-GARIĆ (Wien) — Lehn- und Erbwortschatz des Kroatischen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ im Lichte von Etymologie und Empirie | 56 |
| SASKIA PRONK-TIETHOFF (Leiden) — The Methodology of the <i>Loanword Typology Project</i> : Opportunities for the Investigation of Loanwords in Slavic | 110 |
| STEFAN ENGELBERG, PETER MEYER (Mannheim) — Das <i>Lehnwortportal Deutsch</i> als kontaktlinguistisches Forschungsinstrument | 120 |
| JÜRGEN FUCHSBAUER (Regensburg, Wien) — Gräzisierung versus Sprachpurismus Zum Fremdgut im Wortschatz des bulgarischen Kirchenslawisch des 14. Jahrhunderts | 140 |
| ALEXANDER KUPERDYAEV (Innsbruck) — Ottoman language heritage in Macedonian: Un colosse aux pieds d'argile?..... | 150 |
| BARBARA SONNENHAUSER (Zürich) — Borrowing in context: a pragmatic perspective on Turkisms in pre-standardised Balkan Slavic..... | 173 |
| RUPRECHT VON WALDENFELS (Kraków) — Inner-Slavic contact from a corpus driven perspective..... | 196 |

Einleitung

Der vorliegende Sammelband geht zurück auf den internationalen Workshop „Lehnwörter im Slawischen: Empirische und crosslinguistische Perspektiven“, der in einem kleinen Kreis, dafür aber in umso stärker produktiver Atmosphäre vom 25. bis 26. September 2014 am Institut für Slawistik der Universität Wien stattfand. Bis auf zwei Ausnahmen werden in diesem Sammelband die damals präsentierten Beiträge abgedruckt.

Im Mittelpunkt des Workshops stand dabei die Frage der Untersuchung von Lehnwortbeziehungen von slawischen Sprachen; als Stimulus für die weiterführende theoretische, methodologische und empirische Auseinandersetzung wurde zu diesem Zweck ein in der Sprachtypologie und vergleichenden Sprachwissenschaft verankertes Projekt, nämlich *Loanwords in the world's languages*¹ von Haspelmath/Tadmor (2009a), herangezogen. Es ging darum die Möglichkeiten und Probleme auszuloten, die im Falle einer Untersuchung von slawischen Sprachen zu berücksichtigen sind. Das Projekt *Loanwords in the world's languages* ist darauf ausgerichtet innerhalb einer Sprache den jeweiligen Bestand an Lehnwörtern auf folgende Weise systematisch zu erfassen: Ausgehend von einer vorgegebenen Basisliste von ca. 1500 Begriffen (eine Art Grund- bzw. Basiswortschatz mit einer Unterteilung in über 20 semantisch-lexikalische Felder) werden die jeweiligen Lexeme identifiziert, die potenziell als Lehnwörter gelten. Ziel ist es auf diese Weise eine verlässliche Datenbasis für einen systematischen und sprachübergreifenden Vergleich von Sprachen hinsichtlich ihres Lehnwortprofils zu schaffen (vgl. dazu Haspelmath/Tadmor 2009b). Nachdem mittlerweile nach diesem Muster entsprechende Ergebnisse zu über 40 Sprachen vorliegen (vgl. dazu Haspelmath 2009, Tadmor 2009), dabei aber slawische Sprachen nur in geringem Ausmaß berücksichtigt wurden (nämlich lediglich das Niedersorbische in Bartels 2009), wurde im Rahmen des Workshops die Perspektiven der Untersuchung slawischer Sprachen nach diesem Muster diskutiert. Darüber hinaus ging es aber auch um einen aktuellen Einblick in den Stand der Lehnwortforschung in den slawischen Sprachen und um diesbezügliche, im Aufbau begriffene Projekte.

Die Untersuchung von Lehnwortbeziehungen steht seit Langem im Fokus der slawistischen, der historisch-vergleichenden Sprachwissenschaft und der Kontaktlinguistik. Für die slawistische Sprachwissenschaft bzw. die slawische Philologie ist die Auseinandersetzung mit Lehnwörtern ein integraler und wissenschaftsgeschichtlich gesehen konstituierender Beitrag. Im ersten Teil des Sammelbandes sind Arbeiten versammelt, die sich direkt in der einen oder anderen Weise mit dem Projekt *Loanwords in the world's languages* auseinandersetzen und dabei auf jeweils selektive Probleme und offene Frage eingehen.

¹ Vgl. dazu <<http://wold.cild.org/>> [aufgerufen am 01.07.2015].

Im Beitrag von **Emmerich Kelih** (Wien) „Probleme der empirischen Lehnwortforschung: Lehnwörter im Basiswortschatz (Slowenisch) und ein frequenzbasierter Ansatz (Kroatisch)“ geht es neben einer allgemeinen Vorstellung des Forschungsdesigns der Untersuchung von Lehnwörtern auf der Grundlage eines Basiswortschatzes, insbesondere auch um die Art der lexikalisch-semantischen Gliederung des zu untersuchenden Materials. Ohne Zweifel ist mit diesem Vorgehen in gewisser Weise eine neue Perspektive in der Lehnwortforschung verbunden, da bekanntermaßen gerade ein sogenannter Basiswortschatz in der Regel durch ein relativ hohes Alter bzw. durch einen geringen Anteil – so die Standardannahme der Kontaktlinguistik – von Lehnwörtern ausgezeichnet ist. Während Haspelmath/Tadmor (2009b) die Untersuchung einer bereits fix vorgegebenen Liste von etwas über 1500 Bedeutungen ansetzen, geht es dem Autor darum, alternative Möglichkeiten zur Auswahl des Basiswortschatzes aufzuzeigen. Eine solche Möglichkeit wird etwa darin gesehen – in einem ersten Schritt, bevor überhaupt noch die Erprobung an umfangreichem Material erfolgt ist – mit den aus der Glottochronologie und Lexikostatistik bekannten Swadesh-Listen zu beginnen, um auf induktivem Weg auftretende Probleme der Bestimmung von Lehnwörtern zu eruieren und zu thematisieren. Dies wird mit der Hilfe einer 200 Bedeutungen umfassenden, adaptierten Swadesh-Liste für das Slowenische durchgeführt. Darüber hinaus wird in dem Beitrag generell die Frage der Bestimmung eines sogenannten Basiswortschatzes thematisiert. Als eine durchaus lohnende Perspektive bringt der Autor einen frequenzbasierten Ansatz ins Spiel, der darauf hinausläuft, dass die Vorkommenshäufigkeit von Wortformen als Selektionskriterium herangezogen wird. Die empirische Dimension dieser Alternative wird anhand des Kroatischen exemplifiziert, indem in einem Häufigkeitswörterbuch die unterschiedlichen Arten von Lehnwörtern identifiziert werden.

Der Beitrag von **Stefan Michael Newerkla** (Wien) „Kernproblembereiche der Lehnwortforschung anhand von Beispielen aus dem Kontaktbereich Deutsch, Tschechisch und Slowakisch“ geht auf zentrale Probleme ein, die bei der Untersuchung von Lehnwortbeziehungen zu berücksichtigen sind. Zu beginnen ist generell mit der zeitlichen Tiefe dieser Beziehungen, die letztlich nur unter Berücksichtigung der jeweiligen Quellenlage adäquat eingeschätzt werden kann. So lässt sich in vielen Fällen eine eindeutige Zuordnung von lexikalischen Entlehnungen aufgrund von Regelmäßigkeiten des Lautersatzes und semantischen Entwicklungslinien bewerkstelligen. Dennoch ist oftmals eine eindeutige Einordnung der Entlehnrichtung nicht möglich. Diese Problematik betrifft insbesondere auch innerslawische Lehnwortbeziehungen, die durch eine hohe Ähnlichkeit der Dialekte in den älteren Perioden gekennzeichnet sind und somit eine Rekonstruktion der entsprechenden Entlehnungen nachhaltig erschweren. Verwiesen wird auch auf das in vielen Sprachen beobachtbare Phänomen, wonach Lehnwörter nach ihrer Entlehnung weiterhin mit der Gebersprache in Kontakt bleiben; dies führt in weiterer Folge dazu, dass Adaptionprozesse nicht auf „natürliche“, also rein innersprachliche Weise

erfolgen, sondern in gewisser Weise zeitlich verschoben und geblockt werden, was wiederum die Identifikation von Lehnwörtern bzw. Zuordnungen nachhaltig verkomplizieren kann. Die Frage von Standardisierungsprozessen, die für die heutigen slawischen Sprachen von besonderer Bedeutung sind, können auch direkte Konsequenzen für den Lehnwortbestand haben, sofern es zu bewussten puristischen Umgestaltungen der Lexik kam (in diesem Sinne ist dies nicht nur relevant für das Tschechische und das Slowakische, sondern im Grunde genommen auch für alle südslawischen Sprachen, deren lexikalisches Bild sich generell im 19. Jh. herausbildete). Im Zusammenhang damit sind neben einer bewussten Archaisierung auch Calquierungen von hoher Bedeutung. Der Autor plädiert dafür, auch diese Komponenten in das von Haspelmath/Tadmor (2009b) vorgeschlagene Untersuchungsdesign zu integrieren. Generell zeigt dieser Beitrag jedenfalls die nicht zu vernachlässigende Komplexität bei der Bestimmung von möglichen Lehnwörtern, dies sich aufgrund von unterschiedlichen Kontaminationen, mehrfachen Entlehnungen, parallelen Übernahmen und Anpassungen ergibt. Weitere Aspekte betreffen die Frage von versteckter Mehrsprachigkeit, aber auch die arealtypologische Dimension von Lehnwortbeziehungen. Zusammenfassend weist dieser Aufsatz auf den wichtigen philologisch-linguistischen Beitrag hin, den es bei der systematischen Untersuchung von Lehnwortbeziehungen, und dies nicht nur von jenen slawischer Sprachen, zu berücksichtigen gilt.

In ähnlicher Weise lässt sich der Beitrag von **Catharina Krebs-Garić** (Wien) unter dem Titel „Lehn- und Erbwortschatz des Kroatischen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ im Lichte von Etymologie und Empirie“ interpretieren. Im Fokus steht dabei jene semantisch-lexikalische Gruppe, die – sofern man sich auf Tadmor (2009: 64) bezieht – in einem Sample von über 40 Sprachen über den größten Anteil an Lehnwörtern verfügt. Dieser Eindruck lässt sich zwar in Anbetracht der akribischen Studie von Krebs-Garić durchaus bestätigen. In diesem Zusammenhang ergeben sich aber aus dem Beitrag weitere Kernprobleme für eine Untersuchung im Sinne des von Haspelmath/Tadmor (2009b) propagierten Ansatzes. Erstens sind oftmals keine Eins-zu-eins-Entsprechungen der jeweils untersuchten Bedeutungen und der entsprechenden einzelsprachlichen Äquivalente vorhanden; vielmehr lässt sich zuweilen eine – in einzelnen Fällen sogar recht beträchtliche – Anzahl von Synonymen anführen. In diesem Zusammenhang erweist sich insbesondere die Frage von Bedeutung, in welchem Subregister einer Sprache diese eruiert werden. Eine Möglichkeit besteht darin die Selektion nur auf die Standardsprache (deren Abgrenzung allerdings wiederum gesondert zu thematisieren wäre), bzw. eine bestimmte Subsprache (z. B. Fachsprache), Dialekte usw. zu begrenzen. Weiter ist aber in diesem Zusammenhang zu klären, ob bei den zu eruiierenden Äquivalenten nur jeweils in Gebrauch stehende Formen zu erfassen sind, oder ob das gesamte bekannte lexikographische Material zur Auswertung herangezogen werden soll. In Bezug auf das Kroatische stellt sich der von der Autorin angesprochene Versuch einer Erklärung für den hohen Grad an Synonymie in der religiös-theologischen Lexik durch

sprachpuristische Tendenzen durchaus als plausibel heraus, wenngleich eben der Umfang und die Tiefe der jeweiligen Datenaufnahme auch von vorrangiger Bedeutung sein können.

Eine weitere, detaillierte Auseinandersetzung mit dem Untersuchungsdesign von Haspelmath/Tadmor (2009) findet sich im Beitrag „The Methodology of the *Loanword Typology Project: Opportunities for the Investigation of Loanwords in Slavic*“ von **Saskia Pronk-Tiethoff** (Leiden). Ein genereller Kritikpunkt ergibt sich aus der im genannten Projekt verwendeten engen Definition von Lehnwörtern, die generell Calquierungen und Neologismen ausschließt. Auf diese Weise kann sich nur ein eingeschränktes, ja stark reduziertes Bild der Lehnwortbeziehungen einer Sprache ergeben. Auf der Basis eines Vergleichs des niedersorbischen und holländischen Datenmaterials gelingt es der Autorin zu zeigen, dass es bei der Interpretation der zeitlichen Tiefe zu bedeutsamen Unterschieden kommen kann. So wird bei einer Studie von schriftlich belegbarem Material, in der anderen von Daten, die mit den Mitteln der historisch-vergleichenden Sprachwissenschaft rekonstruiert wurden, ausgegangen. Die Frage der tatsächlichen Vergleichbarkeit ist jedoch nur im Falle von ähnlichen Ausgangs- und Randbedingungen gegeben. Die Analyse von germanischen Lehnwörtern im Urslawischen ergibt wiederum eine thematische Einengung von lexikalischen Entlehnungen auf bestimmte lexikalisch-semantische Bereiche. Diese Ergebnisse sind bei weiteren vergleichenden Untersuchungen der Lehnwortbeziehungen von slawischen Sprachen zu berücksichtigen. Der Beitrag ist für weiterführende Analysen aber auch deshalb von grundlegender Bedeutung, weil die Problematik von Inkonsistenzen bei der Datenaufnahme eindringlich aufgezeigt wird.

Einen aktuellen Einblick in laufende Projekte im Bereich der Lehnwortbeziehungen geben **Stefan Engelberg** und **Peter Meyer** (beide Mannheim) in ihrem Beitrag „Das Lehnwortportal Deutsch als kontaktlinguistisches Forschungsinstrument“. Es geht um einen innovativen lexikographischen Metaansatz, bei dem aus einer Vielzahl von vorhandenen Lehnwörterbüchern jeweils die deutschen Lehnwörter (u. a. in slawischen Sprachen, bislang im Polnischen und im Slowenischen) in einer online zugänglichen Datenbank zusammengeführt werden, um eine gemeinsame Datenbasis für weiterführende Untersuchungen zu bieten. Im Mittelpunkt des Beitrages steht dabei die Frage einer adäquaten onomasiologischen Struktur dieses Lehnwortportals, die eine fortgeschrittene Suche ermöglicht. Die Autoren setzen sich nicht nur kritisch mit bestehenden semantischen Klassifikationsschemata auseinander, sondern plädieren für eine induktiv zu kreierende semantische Klassifikation, die sowohl hyperonymische Relationen als auch semantische Felder beinhaltet. Eine Besonderheit des Lehnwortportales ist insbesondere in der Möglichkeit einer visuellen Datenmodellierung der semantischen Felder von Lehnwörtern bzw. auch der Repräsentation von beliebig langen Ketten von Entlehnungsbeziehungen zu sehen.

Jürgen Fuchsbauer (Regensburg) widmet sich in seinem Beitrag dem griechischen Lehngut im mittelbulgarischen Kirchenslawisch. Im Mittelpunkt steht ein spezieller Aspekt von Lehnbeziehungen, und zwar das Ersetzen von etablierten Lehnwörtern durch Lehnübersetzungen. Diese puristische Tendenz ist dem Anschein nach charakteristisch für mehrere Übersetzer, Redaktoren und Kopisten, die wohl in der ersten Hälfte des 14. Jahrhunderts in den Zentren der mittelbulgarischen Schriftkultur, nämlich am Berg Athos und in der damaligen Hauptstadt Tärnovo, wirkten. Sie kann somit in den Kontext der Reform des Kirchenslawischen im Zweiten Bulgarischen Reich gestellt werden, wobei sie einen gewissen Einblick in das Sprachdenken der Zeit bietet. Dem besagten Purismus mag das Bestreben, griechisches Sprachmaterial mit eigenen Mitteln wiederzugeben, das ja insbesondere in der Syntax klar zu Ausdruck kommt, zugrunde gelegen sein. Dabei konnten sich viele der so entstandenen Neologismen nicht durchsetzen; sie wurden bereits in der Tärnover Redaktion der Synaxarienübersetzung wieder durch die entsprechenden Lehnwörter ersetzt.

Alexander Kuperdyaev (Innsbruck) verfolgt in seinem Artikel „Ottoman language heritage in Macedonian: Un colosse aux pieds d’argile?“ einen explizit frequenz- und korpusbasierten Ansatz, um dem aktuellen Status und der Gebrauchsfrequenz von Turzismen im Makedonischen nachzuspüren. Somit lässt sich die Arbeit in den aktuellen Diskurs eines Ansatzes einreihen, den man als „usage-based“ bezeichnen kann. Ob der fehlenden Korpus-Infrastruktur für das Standardmakedonische (diese Aussage ist in Relation zu Sprachen wie Tschechisch, Russisch u. a. slawischen Sprachen zu verstehen) konzentriert sich der Autor auf die Häufigkeit von Turzismen im aktuellen journalistischen Diskurs. Ausgehend von einer Datenbasis von über 3000 auf der historischen Achse nachweisbaren Turzismen (immerhin stand das Makedonische über 600 Jahre im direkten Kontakt mit dem Türkischen, und auch aus synchroner Sicht setzt sich dieser Kontakt fort, wenngleich offenbar in einem weitaus geringeren Ausmaß) wird die heutige Verwendungshäufigkeit, getrennt nach lexikalisch-semantischen Gruppen, aber auch nach Wortarten, untersucht. Es ist nach Angaben des Autors ein quantitativer Rückgang von Turzismen feststellbar, allerdings zeigt eine abschließende Einbettung in bestehende Entlehnungsskalen (Thomason 2001), dass durchaus nach wie vor von einer relativ tiefen Verankerung von Turzismen im Makedonischen auszugehen ist.

In dem Beitrag „Borrowing in context: a pragmatic perspective on Turkisms in pre-standardised Balkan Slavic“ geht **Barbara Sonnenhauser** (Zürich) detailliert auf Turzismen im vorstandardisierten Balkanslawischen, insbesondere aus dem Zeitraum vom Ende des 17. bis zur Mitte des 19. Jh. ein. Im Fokus steht die Frage der pragmatischen Dimension der Verwendung von Turzismen und deren Einbettung in die entsprechende literarische Produktion aus dieser Zeit. Durch das Prisma von Turzismen lässt sich nach Ansicht der Autorin u. a. das soziolinguistische Gesamtbild dieser Epoche rekonstruieren, indem die metasprachliche Dimension dieser Art von Kommunikation näher betrachtet wird. Die Frage

des Eingangs von entlehntem lexikalischem Material wird im Zusammenhang des linguistischen Bewusstseins und des sprachlichen Hintergrunds der jeweiligen Autoren, aber auch unter Berücksichtigung des potenziellen Lesepublikums dieser Zeit reflektiert. Eine eingehende Auseinandersetzung mit unterschiedlichen Formen von Turzismen zeigt, dass sich zumindest drei Gebrauchstypen herauskristallisieren. Da sind zunächst integrierte Lehnwörter, die tatsächlich zum Auffüllen einer semantischen Lücke gebraucht wurden, von denen nicht integrierte Lehnwörter unterschieden werden können. Letztere scheinen besonders aus metakommunikativer und metalinguistischer Perspektive von Interesse zu sein, zumal sie im damaligen Diskurs in erklärender Funktion herangezogen wurden. Den dritten Gebrauchstyp bilden jene Lehnwörter, die in erster Linie aus stilistischen Gründen bzw. zur Erlangung eines bestimmten Effektes verwendet wurden. Insgesamt vermittelt dieser Beitrag einen umfangreichen Einblick in die spezifische Situation vor der Standardisierung der ostslawischen baltoslavischen Standardsprachen, die geprägt von einer zunehmenden Bedeutung der Volkssprache ist, allerdings eben (noch) mit einem geringen Grad an Purismus und Normativität.

Während in den meisten Beiträgen die lexikalische Komponente von Entlehnungen im Vordergrund steht, geht **Ruprecht von Waldenfels** (Krakau) in seinem Beitrag „Inner-Slavic contact from a corpus driven perspective“ auf die morphologische Dimension dieser Problemstellung näher ein, wobei die Frage der Ähnlichkeit bzw. der Distanz von slawischen Sprachen zur Diskussion steht. Sie wird anhand von ausgewählten verbalen Präfixen und nominalen Suffixen in zur Verfügung stehenden Paralleltextkorpora analysiert. Mit der Hilfe eines vorgeschlagenen Distanzmaßes und einer damit einhergehenden Visualisierung in der Form von Baumgraphen lässt sich die jeweilige Nähe und Gruppierung der einzelnen slawischen Sprachen anhand ausgewählter Parameter lokalisieren. Im untersuchten Material zeigt sich, dass in der Regel nicht die traditionelle (genetisch bzw. areal determinierte) Gliederung der slawischen Standardsprachen reproduziert wird, sondern vielmehr für verbale Präfixe bzw. nominale Suffixe jeweils unterschiedliche Ergebnisse erzielt werden. Insbesondere erweist sich vor diesem Hintergrund die Zuordnung des Slowenischen und des Bulgarischen zu den südslawischen Sprachen als problematisch, zumal im Falle des Slowenischen eine (durchaus argumentierbare) Nähe zum Tschechischen bzw. Slowakischen und im Falle des Bulgarischen eine Nähe zum Russischen zu beobachten ist. Die präsentierte Methode (Verwendung von Paralleltexten, Berechnung von Distanzmaßen und einhergehende Visualisierung) sowie die erzielten Resultate zeugen vom generellen Potenzial für weitere Untersuchungen von Entlehnungen in den slawischen Sprachen.

Die versammelten Beiträge sind in ihrer Gesamtheit eine umfassende Sammlung einer Vielzahl von Problemen und Perspektiven, die mit der Lehnwortforschung slawischer Sprachen verbunden sind. Ruft man die ursprüngliche Intention in Erinnerung, die darin bestand, aus slawistischer, aber auch als

allgemein philologisch-linguistischer Position Stellung zu aktuellen Ansätzen der Lehnwortforschung zu beziehen, so zeigt sich – ohne dem Inhalt an dieser Stelle zu weit vorgreifen zu wollen –, dass die Erstellung eines empirisch fundierten Lehnwortprofils für die slawischen Sprachen vor dem Hintergrund der Vielzahl von zu beachtenden Problemen als ein nur langfristig zu erreichendes Ziel gesehen werden kann. Zumindest werden im Buch aber bereits eine Reihe von mögliche Fallen und Sackgassen auf dem Weg dorthin aufgezeigt.

Die Herausgeber,
Wien, im Juli 2015

Verwendete Literatur:

- BARTELS, HAUKE (2009): Loanwords in Lower Sorbian, a Slavic language of Germany. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 304–323.
- HASPELMATH, MARTIN (2009): Lexical borrowing: Concepts and issues. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 35–54.
- HASPELMATH, MARTIN; TADMOR, URI (eds.) (2009a): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter.
- HASPELMATH, MARTIN; TADMOR, URI (2009b): The Loanword Typology project and the World Loanword Database. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 1–34.
- TADMOR, URI (2009): Loanwords in the world's languages: Findings and results. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 55–75.
- THOMASON, SARAH GREY (2001): *Language contact*. Edinburgh: Edinburgh University Press.

Emmerich Kelih (Wien)

**Probleme der empirischen Lehnwortforschung: Lehnwörter im Basiswortschatz
(Slowenisch) und ein frequenzbasierter Ansatz (Kroatisch)**

Abstract

The present paper discusses in detail the project *Loanwords in the world's languages*, initiated by Haspelmath/Tadmor (2009a, b). It mainly focuses on the proposed methodological design of the crosslinguistic analysis of lexical borrowings in the world's languages. The main idea is to survey the types of loanwords in the core vocabulary of languages on the basis of a fixed list of 1,460 lexical meanings. After the discussion of the main theoretical issues and the methodological design, two different approaches of the determination of the core vocabulary are proposed. One approach is the reduction of the 1,460 lexical meanings according to the generally known Swadesh-list, which consists of only 200 lexical meanings. The empirical dimension of the analysis of loan words is shown by means of an adapted Swadesh-list for Slovene. The second approach is a frequency based one and implies the analysis of the lexical material provided by frequency dictionaries. The analysis of the 1,500 most frequent lexical items of a balanced Croatian frequency dictionary shows a surprisingly high number of loanwords. It appears that both approaches can be regarded as useful for the empirical examination of loan relations.

1. Einleitung

Der vorliegende Beitrag¹ setzt sich detailliert mit dem von Haspelmath/Tadmor (2009a) initiierten Projekt *Loanwords in the world's languages* auseinander. Im gegebenen Zusammenhang wird es darum gehen zu diskutieren, inwiefern und ob das in diesem Projekt verwendete Forschungsdesign für die Untersuchung slawischer, insbesondere südslawischer Sprachen verwendet werden kann bzw. ob und in welcher Form Adaptationen und Modifikationen durchzuführen sind. Das Projekt *Loanwords in the world's languages* ist darauf ausgerichtet, innerhalb einer Sprache den jeweiligen Bestand an Lehnwörtern systematisch zu erfassen. Ausgehend von einer vorgegebenen Basiswortliste mit circa 1500 Bedeutungen (eine Art Grund- bzw. Basiswortschatz mit einer Unterteilung in über 20 unterschiedliche semantisch-lexikalische Gruppen) werden für die zu untersuchenden Sprachen die jeweiligen Lexeme identifiziert, die als Lehnwörter gelten. Ziel ist es, auf diese Weise eine verlässliche Datenbasis für einen systematischen und sprachübergreifenden Vergleich von Sprachen hinsichtlich ihres Lehnwortprofils zu erhalten.

Nachdem mittlerweile nach diesem Muster entsprechende Ergebnisse zu über 40 Sprachen vorliegen (vgl. dazu Details in Haspelmath 2009, Tadmor 2009), dabei aber slawische Sprachen nur in einem

¹ Für wertvolle Kommentare und Anmerkungen ist an dieser Stelle S. M. Newerkla, J. Fuchsbauer und C. Krebs-Garić herzlichst zu danken.

geringen Ausmaß berücksichtigt wurden (vgl. dazu die Analysen für das Niedersorbische in Bartels 2009), werden im gegebenen Rahmen die Möglichkeiten und Perspektiven der Untersuchung slawischer Sprachen nach diesem Muster ausgelotet. Im Detail wird es um die Frage gehen, ob es brauchbare Alternativen zur Untersuchung der umfangreichen 1500 Bedeutungen umfassenden Wortliste gibt. Konkret werden zwei alternative Möglichkeiten diskutiert:

1. die Verwendung einer auf 200 Bedeutungen reduzierten Basiswortschatzliste in Form einer für slawische Sprachen modifizierten Swadesh-Liste (Details dazu in Kap. 4.1.),
2. die Untersuchung und Eruiierung von Lehnwörtern in häufig vorkommenden Lexemen innerhalb einer Sprache, wobei in diesem Zusammenhang auf Häufigkeitswörterbücher Bezug genommen werden kann, die ihrerseits auf umfangreichen, auf den synchronen Bestand einer Sprache ausgerichteten Korpora basieren.

Die Untersuchung und Eruiierung von Lehnwörtern in einem reduzierten Basiswortschatz wird anhand des Slowenischen (vgl. Kap. 4.1.) demonstriert, während die Untersuchung von Lehnwörtern in Häufigkeitswörterbüchern anhand des Kroatischen erfolgt (vgl. Kap 4.2.).

2. Empirische Lehnwortforschung: Haspelmath/Tadmor (2009a)

Eine systematische Untersuchung zur Verteilung von Lehnwörtern in den Sprachen der Welt geht, wie bereits einleitend erwähnt, auf das von Haspelmath/Tadmor (2009a) initiierte Projekt *Loanwords in the world's languages* zurück. Dieses Projekt bietet ein einheitliches Untersuchungsdesign², welches für die Untersuchung unterschiedlicher Sprachen angewandt wird und somit nicht nur eine Vergleichbarkeit der Verteilung von Lehnwörtern in unterschiedlichen semantisch-lexikalischen Gruppen gewährleistet, sondern überhaupt erst das Erstellen eines individuellen Lehnwortprofils einer Sprache ermöglicht.

Im Detail ist deren Vorgehen folgendes: Ausgehend von einer Liste mit 1460 lexikalischen Bedeutungen, die auf Englisch vorliegt, werden in einem ersten Schritt für jede zu untersuchende Sprache die entsprechenden Übersetzungsäquivalente angeführt. In einem nächsten Schritt wird sodann vermerkt, ob das entsprechende Lexem ein Lehnwort ist oder nicht. Eine Besonderheit dieses Ansatzes liegt darin, dass die 1460 lexikalischen Bedeutungen keine ungeordnete Menge darstellen, sondern diese in 22

² Einige problematische Eckpunkte dieses Projektes können an dieser Stelle nicht ausführlich besprochen werden. Diese betreffen vor allem die heterogene Sprachenauswahl, die Untersuchung von Standardsprachen und dialektalen Formen, die eingeschränkte Vergleichbarkeit von Sprachen aufgrund von vorhandenen biologischen, geographischen und kulturellen Unterschieden (damit zusammenhängend die fehlenden Übersetzungsäquivalente und die beobachtbare fehlende semantische Kongruenz), die in einigen Fällen als arbiträr erscheinende Zuordnung von Bedeutungen zu lexikalisch-semantischen Gruppen u. a. m. Auffällig ist auch die Tatsache, dass in diesem Projekt von einem sehr engen Verständnis eines Lehnworts ausgegangen wird, und z. B. entlehnte Morphemstämme nur unter bestimmten Umständen als Lehnwort interpretiert werden (vgl. dazu Haspelmath/Tadmor 2009b: 13 und 15 bzw. ähnlich Haspelmath 2009: 37).

lexikalisch-semantische Klassen unterteilt werden, wie sie bereits in dem Wörterbuch von Buck (1949), der den Wortschatz von ausgewählten indogermanischen Sprachen vergleichend gegenübergestellt hat, vorgenommen wurde. Haspelmath/Tadmor (2009b: 7) haben diese Liste geringfügig modifiziert³ und zwei weitere lexikalisch-semantische Gruppen hinzugefügt, nämlich eine Gruppe „Moderne Welt“ und „diverse Funktionswörter“. Die einzelnen Gruppen umfassen – vgl. dazu die Tab. 1 mit einer Übersicht – unterschiedliche Bereiche, wie die physikalische Welt, die Tierwelt, Körperbezeichnungen, Arten der Bewegung, Begriffe aus dem Bereich Jagd und Kriegsführung usw.

Tab. 1 Lexikalisch-semantische Gruppe und Anzahl von Bedeutungen nach Haspelmath/Tadmor (2009b: 7)

| Nr. | Lexikalisch-semantische Gruppe | Anzahl von Bedeutungen |
|-----|--|------------------------|
| 1 | physische Welt | 75 |
| 2 | Verwandtschaft | 85 |
| 3 | Tiere | 116 |
| 4 | Körper | 159 |
| 5 | Essen und Trinken | 81 |
| 6 | Kleidung und Pflege | 59 |
| 7 | Haus(halt) | 47 |
| 8 | Landwirtschaft und Vegetation | 74 |
| 9 | grundlegende Handlungen und Technologien | 78 |
| 10 | Bewegung | 82 |
| 11 | Besitz | 46 |
| 12 | räumliche Relationen | 75 |
| 13 | Quantität | 38 |
| 14 | Zeit | 57 |
| 15 | Sinneswahrnehmungen | 49 |
| 16 | Gefühle und Werte | 48 |
| 17 | Kognition | 51 |
| 18 | Sprechen und Sprache | 41 |
| 19 | soziale und politische Beziehungen | 36 |
| 20 | Kriegsführung und Jagd | 40 |
| 21 | Recht | 26 |
| 22 | Religion und Glaube | 26 |

³ Haspelmath/Tadmor (2009b) haben diese Basiswortschatzliste nicht direkt aus Buck (1949) übernommen, sondern diese beruht auf der im Rahmen der *Intercontinental Dictionary Series (IDS)* erstellten Liste, die von Mary Ritchie Key und Bernard Comrie betreut wird. Vgl. dazu <<http://lingweb.eva.mpg.de/ids/>> [aufgerufen am 25.02. 2015]. Das Projekt *IDS* umfasst derzeit über 200 Sprachen mit einer Liste von etwas über 1200 Bedeutungen. Demgegenüber wird die von Haspelmath/Tadmor (2009b) vorgestellte Liste um über 200 Bedeutungen in unterschiedlichen lexikalisch-semantischen Gruppen ergänzt. Die Adaption und Modifikation betrifft, neben geringfügigen Änderungen der Benennung von lexikalisch-semantischen Gruppen, vor allem die Einführung der Gruppe „Moderne Welt“, die 57 Bedeutungen (z. B. Bezeichnungen für ‘Radio’, ‘Bus’, ‘Krankenhaus’ etc.) enthält, aber aufgrund der Heterogenität der darin zusammengefassten Konzepte kaum als eine lexikalisch-semantische Klasse aufgefasst werden kann. Ähnliches gilt für die hinzugefügte Gruppe „diverse Funktionswörter“.

| | | |
|---------------|-------------------------|-------------|
| 23 | moderne Welt | 57 |
| 24 | diverse Funktionswörter | 14 |
| Gesamt | | 1460 |

Diese Auswahl der lexikalisch-semantic Gruppen zielt in erster Linie darauf ab, den Basiswortschatz einer Sprache zu erfassen. Wenngleich die Auswahlkriterien nicht weiter reflektiert und diskutiert werden, handelt es sich hierbei offenbar um als wichtig und zentral empfundene Begriffe innerhalb einer Sprache. Auf diesen Aspekt, die Auswahlkriterien und einige damit verbundene Probleme wird in Kap. 4 detailliert einzugehen sein. Hinsichtlich der in Tab. 1 angeführten lexikalisch-semantic Klassen fällt aus quantitativer Sicht insbesondere Folgendes auf: Es dominieren eindeutig die Bereiche Körper/Körperteile, Tiere, Verwandtschaftsbezeichnungen und Bewegungsarten, womit es zu einer eindeutigen Ungleichverteilung der einzelnen lexikalisch-semantic Klassen kommt, die aber nicht weiter begründbar ist.

Abgesehen von diesem problematischen Teilaspekt ist aber dennoch die globale Relevanz dieses Untersuchungsdesigns bei der Untersuchung von Lehnwörtern hervorzuheben. In gewisser Weise ist damit ein Perspektivenwechsel verbunden, indem eben ein sogenannter Basiswortschatz, dessen Bestimmung im Detail zwar nicht problemlos ist, zum Ausgangspunkt genommen und sodann systematisch die einzelnen lexikalisch-semantic Bereiche hinsichtlich vorhandener Lehnwörter durchkämmt werden. Damit gelingt es in erster Linie die vielfach in der Linguistik angenommene, bislang aber nicht genügend reflektierte Annahme in Frage zu stellen, dass der Basiswortschatz einer Sprache besonderes resistent gegenüber Entlehnungen ist (Swadesh 1952, Thomason/Kaufman 1988: 69ff, Hock/Joseph 1996: 253ff, Thomason 2001: 71–72, Embleton 1986: 71 und 100ff, McMahon 1994: 204 u.v.a.m.).

Es können an dieser Stelle nicht alle Ergebnisse der bisherigen Untersuchungen von über 40 Sprachen (unterschiedlicher typologischer Zugehörigkeit, Untersuchung von sowohl Schriftsprachen als auch Dialekten) referiert werden, die im Handbuch von Haspelmath/Tadmor (2009a) zu finden sind. Ein zentrales Ergebnis ist aber, dass der Basiswortschatz von Sprachen in einem sehr unterschiedlichen Ausmaß durch Lehnwörter ausgedrückt wird. Die eruierten quantitativen Ausmaße geben einen ersten Einblick in die mögliche Variation. Während in einer in der Slowakei gesprochenen Varietät des Romani der Basiswortschatz einen Lehnwortanteil von über 60% aufweist, kann für das Mandarin-Chinesische nur ein Anteil von 1,2% Lehnwörtern ausgemacht werden, wobei selbst dieser Anteil umstritten ist, da einige Etymologien nicht eindeutig sind. Das Obersorbische, die einzige im Projekt untersuchte slawische Sprache, ist laut Tadmor (2009: 57) mit einem knapp über 20% liegenden Anteil an Lehnwörtern als ein „durchschnittlicher Entlehner“ zu bezeichnen. Insofern lässt sich bereits zum gegenwärtigen Zeitpunkt

festhalten, dass nicht pauschal von einer Resistenz⁴ des Basiswortschatzes gegenüber Entlehnungen gesprochen werden kann, sondern vielmehr von einer unterschiedlichen Entlehnungsaktivität einer Sprache auszugehen ist. Darüber hinaus ist der Anteil an Lehnwörtern im Basiswortschatz als ein intersubjektiv ermittelbarer Gradmesser für die Tiefe und die Intensität von Sprachkontakten anzusehen.

Das zweite wichtige Ergebnis, welches sich aus den bisher vorliegenden Ergebnissen ableiten lässt, ist die Möglichkeit, innerhalb von Sprachen lexikalisch-semantische Präferenzzonen für Entlehnungen zu eruieren. Hinsichtlich der Verteilung von Lehnwörtern auf die genannten lexikalisch-semantischen Klassen gilt, wie es Tadmor (2009: 64) treffend formuliert hat, dabei Folgendes:

However, different languages display a remarkable degree of consistency which [sic!] regard to which fields are more or less affected by borrowing. While there are certainly cross-linguistic differences, most languages tend to borrow more words into similar fields, and the same fields turn up again as the ones most resistant to borrowing.

Es lässt sich beobachten, dass jeweils bestimmte lexikalisch-semantische Klassen für Entlehnungen besonders zugänglich sind, während andere Klassen aber besonders resistent zu sein scheinen. Dennoch lassen sich ausgehend von den untersuchten 41 Sprachen einige für Entlehnungen favorisierte lexikalisch-semantische Bereiche eruieren.

So erweisen sich als führend hinsichtlich des Lehnwortanteils (vgl. Tadmor 2009: 64) die lexikalisch-semantischen Klassen (1) Religion und Glaube, (2) Kleidung und Pflege und (3) Haus(halt). Diese Ergebnisse lassen sich dadurch erklären, dass die Sprecher genau für diese Bereiche ein erhöhtes Bedürfnis nach der Übernahme von Lehnwörtern aufweisen und die betreffenden Bezeichnungen einer erhöhten Dynamik (im Sinne eines stetigen Austauschs) ausgesetzt sind und offenbar dazu tendieren, aus jeweils anderen Sprachen bzw. Kulturen übernommen zu werden. Hinsichtlich der Gruppe Religion und Glaube zeigt sich offenbar die in vielen Fällen beobachtbare sprachübergreifende Relevanz der theologisch-religiösen Lexik, die in der Regel nicht nur innerhalb einer Sprache, sondern in mehreren Sprachen bzw. Kulturen gleichzeitig in jeweils adaptierter oder modifizierter Form auftritt.

Als tatsächlich lehnwortresistente lexikalisch-semantische Bereiche des Basiswortschatzes erweisen sich nach Tadmor (2009: 64) folgende: (1) Sinneswahrnehmungen, (2) räumliche Relationen und (3) der menschliche Körper, inklusive der Bezeichnung von Körperteilen. Diese Bereiche werden entweder gar nicht bzw. nur in einem sehr geringen Ausmaß durch Lehnwörter ausgedrückt. Als zentraler Kernbestand von in der Regel nicht entlehnter Lexik gelten vor allem Funktionswörter mit deiktischer

⁴ Im gegebenen Zusammenhang wird darunter verstanden, dass der Basiswortschatz einer Sprache generell nicht durch Lehnwörter ausgedrückt wird. Zenner/Speelman/Geeraerts (2014: 78) unterscheiden plausibler Weise drei unterschiedliche Spielarten von Lehnwortresistenz: (1) Resistenz gegenüber Ersetzung einer indigenen durch eine entlehnte Form, (2) Resistenz gegenüber einer Koexistenz von indigenen und entlehnten Formen und (3) Resistenz gegenüber der Koexistenz, aber mit einer sich durchsetzenden entlehnten Form.

Funktion, insbesondere in Bezug auf zeitliche (*heute, gestern, übermorgen*) aber auch räumliche Relationen (*oben, unten, draußen*). Angesichts dieser Ergebnisse lässt sich doch eine stringente innere Gliederung des Basiswortschatzes erkennen, da es wohl tatsächlich einen engen Bereich gibt, der fast ausschließlich mit indigener Lexik ausgedrückt wird.

Auch wenn diese ersten, durchaus überzeugenden Resultate vorliegen, wird im Folgenden dennoch neuerlich die Frage aufgeworfen, auf welche Art und Weise ein sogenannter Basiswortschatz einer Sprache eruiert wird bzw. werden die entsprechenden Auswahlkriterien kritisch zu reflektieren sein. Schließlich ist es genau die Auswahl der zu untersuchenden Bedeutungen, die das Resultat hinsichtlich des Anteils von Lehnwörtern präjudizieren kann.

3. Bestimmung und Gruppierung des Basiswortschatzes

Ein generelles Problem – abgesehen von vielen in diesem Zusammenhang auftauchenden Detailfragen – ist die Frage der Bestimmung und der Auswahl des Basiswortschatzes für weiterführende Lehnwortuntersuchungen. Der Begriff *Basiswortschatz* teilt das Schicksal vieler linguistischer Termini und ist nicht eindeutig definiert und konkurriert darüber hinaus mit ähnlichen Begriffen wie *Grundwortschatz*, *lexikalisches Minimum*, *Kernwortschatz*, *core vocabulary* u. a. m. Darunter wird in der Regel ein lexikalisches Minimum im Sinne eines reduzierten Wortschatzes verstanden, welches in der Lage ist, grundlegende kommunikative Bedürfnisse eines Sprechers abzudecken. Insofern ist es daher nicht verwunderlich, dass die Diskussion um die Auswahl eines Basiswortschatzes in erster Linie in einem sprachdidaktischen Kontext (Mindestwortschatz für L2-Lerner, geeigneter Wortschatz für L1-Lerner auf unterschiedlichen Kompetenzstufen etc.) geführt wird (vgl. Bußmann 2008: 297, Krohn 1992, Schnörch 2002), indem aus kommunikativ-pragmatischen Gründen eben bestimmte Sphären als besonders wichtig erachtet werden.

Als ein immer wieder angeführtes Kriterium für die Auswahl von Lexemen für einen Basiswortschatz wird u. a. die Häufigkeit ihres Auftretens in bestimmten Registern einer Sprache genannt. D. h. häufig vorkommende Lexeme werden als lexikalischer Kernbestand⁵ einer Sprache angesehen, wengleich bereits in Krohn (1992) gezeigt werden konnte, dass ein häufigkeitsbasierter Ansatz in vielen Fällen Lexeme, die aus pragmatisch-kommunikativer Sicht als wichtig und grundlegend angesehen werden können, nicht erfasst werden. Begründbar ist dieses Phänomen durch die Abhängigkeit

⁵ Das Häufigkeitskriterium ist komplementär zum Konzept der Textdeckung zu sehen. Darunter versteht man in der Regel unterschiedliche lexikometrische Kenngrößen, die Auskunft über das Vorkommen von Lexemen in einem Text bestimmter Länge geben. Auf der Basis von lexikostatistischen Befunden erweist sich, dass eine im Grunde relativ beschränkte Anzahl von hochfrequenten Wortformen in Texten sehr oft vorkommt und somit bereits einen großen Teil eines Textes abdeckt. (Vgl. dazu die gut lesbare Übersicht in Tuldava 1998: 72ff.). Dies ist aber nur ein selektiver deskriptiver Aspekt, der wenig über ein tatsächliches Textverständnis durch einen Leser aussagt, aber bei Konzepten der Textverständlichkeit zu berücksichtigen ist.

des häufigkeitsbasierten Ansatzes von der jeweiligen Auswahl der zu untersuchenden Register bzw. Textsorten und darüber hinaus – technisch gesprochen – trivialerweise von der zugrunde gelegten Stichprobengröße, d. h. dem Umfang der untersuchten Textmenge.

Dennoch erscheinen sowohl ein kommunikativ-pragmatischer Ansatz als auch ein häufigkeitsbasierter Ansatz bzw. auch eine Kombination aus diesen beiden Ansätzen als durchaus legitim. Während die konkreten Kriterien der jeweiligen Auswahl umstritten bleiben, scheint aber zumindest hinsichtlich der Anzahl von Lexemen, die ein Grundwortschatz umfassen soll, mehr oder weniger Einigkeit zu bestehen. Wie Krohn (1992: 115) aus dem Vergleich von bestehenden Grundwortschätzen des Deutschen, Französischen und Englischen eruieren konnte, umfasst ein Basiswortschatz einer Sprache in etwa 1500 Lexeme, wenngleich erhebliche Schwankungen sowohl nach oben als auch nach unten beobachtbar sind.

Es erweist sich somit hinsichtlich der von Haspelmath/Tadmor (2009b) verwendeten Liste, dass zumindest die Anzahl der untersuchten Bedeutungen in etwa dem üblichen Ausmaß eines Basiswortschatzes entspricht, die Auswahl der Bedeutungen aber weder durch häufigkeitsbasierte Ansätze noch durch kommunikativ-pragmatische Faktoren motiviert ist, sondern vielmehr subjektiv von als für die menschliche Lebenswelt wichtig empfundenen Bedeutungen ausgegangen wird. In diesem Sinne ist die von Haspelmath/Tadmor (2009b) verwendete Liste eine massive Erweiterung der 200 bzw. 100 Bedeutungen umfassenden Swadesh-Liste, die darauf abzielt jeweils „[...] concepts and experiences common to all human groups [...]“ (Swadesh 1950: 157) bzw. „[...] the fundamental everyday vocabulary of any language [...]“ zu erfassen, deren Auswahl an Bedeutungen aber letztendlich subjektiv und intuitiv bleibt.

Die vorgenommene lexikalisch-semantische Gruppierung (z. B. physische Welt, Verwandtschaft, Tiere, Körper, Essen und Trinken usw.) und die Zuordnung von einzelnen Bedeutungen zu diesen Bereichen erweist sich als ein noch größeres Problem, als dass hier ein Häufigkeitskriterium nur in eingeschränkter Weise von Bedeutung wäre und ob der Vielzahl von Möglichkeiten sich letztlich, wie auch Buck (1949: xiii) anführt, die Suche nach einer idealen Klassifizierung bzw. Zuordnung als hoffnungslos („the ideal is hopeless“) erweist, weswegen aber die Vorteile einer semantischen (Vor-)Klassifizierung nicht in Frage zu stellen sind. In diesem Zusammenhang ist auch der Vorwurf einer gewissen Arbitrarität bei den einzelnen Zuordnungen und Gruppierungen kaum zu entkräften. Es ist wohl die konkrete Auswahl der zu untersuchenden Einheiten der jeweiligen Forschungsfrage anzupassen bzw. im Einzelfall zu reflektieren.

4. Alternativen: Reduzierter Basiswortschatz vs. frequenzbasierter Ansatz

Im Folgenden geht es um die Diskussion von zwei alternativen Vorschlägen hinsichtlich der Untersuchung von Lehnwörtern in einem jeweils selektiven Ausschnitt einer Sprache. In einem Fall wird eine Modifikation der sogenannten Swadesh-Liste mit etwa 200 Bedeutungen herangezogen, die insgesamt als ein auf ein äußerstes Minimum reduzierter Basiswortschatz anzusehen ist. Es werden einige Probleme der Verwendung dieser für das Slowenische modifizierten Swadesh-Liste für Lehnwortuntersuchungen diskutiert.

Im zweiten Fall geht es um die Frage, ob und in welcher Form Häufigkeitswörterbücher für die Bestimmung von Lehnwörtern herangezogen werden können, d. h. es wird von einem frequenzbasierten Ansatz der Eruiierung des Basiswortschatzes ausgegangen, welcher als Grundlage für die Eruiierung der Lehnwörter dient. Diese Frage wird anhand des Kroatischen diskutiert.

Diese zwei Alternativen zu dem von Haspelmath/Tadmor (2009b) vorgeschlagenen Weg sind abschließend dahingehend zu reflektieren, welchen Informationsgehalt sie hinsichtlich der *Lehnwortaktivität* einer Sprache liefern.

4.1. Lehnwortschatz in Swadesh-Liste(n) – Fallbeispiel Slowenisch

Die Analyse eines Basiswortschatzes ist für die vergleichende und kontrastive Linguistik keine methodologische Neuerung (vgl. dazu Hymes 1971: 254), sondern vielmehr deren integraler Bestandteil. Besondere Bekanntheit hat die Analyse des Basiswortschatzes in der sogenannten Glottochronologie bzw. Lexikostatistik, als dessen Begründer Morris Swadesh gilt, in den fünfziger Jahren des 20. Jh. erlangt. Swadesh hat für seine vielfältigen Untersuchungen zur Datierung des absoluten Alters einer Sprache, des Übereinstimmungsgrades der Lexik und der Bestimmung des genetischen Verwandtschaftsgrades von Sprachen (vgl. dazu Dyen 1964 und die detaillierte methodologische Beschreibung des glottochronologischen Ansatzes in Gudschinsky 1956) eine Wortliste erstellt, die einen Basiswortschatz mit etwas über 200 Wortformen enthält, der sich aus unterschiedlichen thematischen Gebieten, wie etwa der Tier- und Pflanzenwelt oder wichtige, den Menschen betreffende Bereichen (Körper, körperliche Zustände, soziale Beziehungen, Personalpronomina, als wichtig empfundene Verben usw.) zusammensetzt (Details dazu Swadesh 1950, 1952, Swadesh 1971).

Die zentrale Annahme bei der Erstellung dieses Basiswortschatzes ist, dass darin eine besonders alte und kaum Änderungen unterliegenden lexikalische Schicht enthalten sein soll, die für vergleichende Analysen, für die Bestimmung der Ähnlichkeit und der genetischen Verwandtschaft von Sprachen eine verlässliche Datengrundlage bildet. Die ursprüngliche Liste von Swadesh verfügt über 200 Bedeutungen, während die neuere Liste nur mehr über 100 Bedeutungen verfügt (vgl. dazu Swadesh 1952, Swadesh 1955 und den Nachdruck der Liste in Swadesh 1971: 283–284). Diese sogenannten Swadesh-Listen sind

in weiterer Folge immer wieder modifiziert⁶ worden, wenngleich deren Zusammensetzung und Verwendung nicht frei von Kritik⁷ geblieben ist.

Als eines der Hauptargumente wird angeführt, dass die Liste vor allem den Wortschatz (west)europäischer Sprachen und Kulturen wiedergibt und eine Übersetzung in bestimmte (vor allem nicht indogermanische) Sprachen nur schwer bzw. gar nicht möglich erscheint (vgl. dazu Hoijer 1956: 53, H. Lunt im Kommentar zu Dyen (1964: 247–252)). Daher ist es generell eine Illusion von einer universell gültigen Liste von Basislexemen auszugehen. Als problematisch werden darüber hinaus die simplizistische Annahme einer 1:1 Korrespondenz zwischen einer gesuchten Bedeutung und deren Entsprechung (vgl. Slaska 2005, die ein alternatives sprecherzentriertes Feststellen der Äquivalente propagiert) bzw. der unklare Status beim Umgang mit vorhandenen Synonymen angesehen.

Um jedoch auf die hier verfolgte Fragestellung nach dem Vorkommen von Entlehnungen in den unterschiedlichen Swadesh-Listen zurückzukommen, ist festzuhalten, dass die Liste zumindest von der Intention so konzipiert ist, dass darin keine Entlehnungen enthalten sein sollten, da diese bei der Bestimmung der genetischen Ähnlichkeit von Sprachen zu einer Verzerrung der Verhältnisse führen. Es ist aber aus der Literatur gut bekannt (vgl. dazu insbesondere Haarmann (1990), der für jede lexikalisch-semanticke Gruppe der Swadesh-Listen eine Vielzahl von Entlehnungen aus vielen Sprachen anführen kann), dass in den Swadesh-Listen sehr wohl Entlehnungen vorkommen können.⁸ Genau dieser Aspekt soll im Folgenden anhand des Slowenischen untersucht werden, um einen ersten Einblick über das Ausmaß an Lehnwörtern in einem reduzierten Basiswortschatz zu erhalten.

⁶ Ein Beispiel für eine derartige Liste für das Russische und Litauische findet sich in Dyen (1964: 244-246). Einen Überblick über neueste Entwicklungen auf diesem Gebiet geben Bakker et al. (2009), die u. a. auf Projekte verweisen, in denen derartige Wortlisten für bis zu 3500 Sprachen erstellt werden. Dies ist ein Bestandteil des *The Automated Similarity Judgment Program (ASJP)*. Details unter <<http://asjp.cild.org/>> [aufgerufen am 18.03.2015].

⁷ An der Glottochronologie, deren Methoden immer elaborierter werden, wird u. a. bemängelt, dass der Basiswortschatz aufgrund von darin enthaltenen Lehnwörtern nicht uneingeschränkt für die Datierung der Herausbildung von Einzelsprachen herangezogen werden kann. Man ist daher dazu übergegangen, die Anzahl von zu untersuchenden Wortformen sukzessive zu reduzieren (vgl. dazu Atkinson/Gray 2006 und Embleton 2000: 152), sodass mittlerweile zum Teil nur mehr von 15 Bedeutungen ausgegangen wird. Es stellt sich allerdings die berechnete Frage, warum nicht zumindest versucht wird zuerst potenzielle Lehnwörter zu identifizieren und a priori von der Untersuchung auszuschließen. Die Problematik der Entlehnungen für die glottochronologische Methode wurde bereits erkannt und Embleton (1986: 100ff.) schlägt die Integration von sprachspezifischen Entlehnungsraten vor. Aber auch dies erweist sich laut Heggarty (2010: 304) ebenfalls als problematisch, da somit das Aufdecken bislang unbekannter Verwandtschaft erschwert wird, da z. B. Sprachen ja durchaus über den gleichen Lehnwortschatz verfügen können.

⁸ Man ist mittlerweile in der Glottochronologie dazu übergegangen, die Swadesh-Liste neuerlich zu modifizieren und zu dichotomisieren, indem zwischen einer äußerst konservativen und einer weniger konservativen Liste unterschieden wird. Vgl. dazu McMahon et al. (2005) und Heggarty (2010) mit einem Überblick über den aktuellen Stand der Arbeit mit jeweils auf unterschiedliche Art und Weise modifizierten Swadesh-Listen.

Als Ausgangsbasis wird eine Basiswortschatzliste herangezogen, die auf Mel’nyčuk (1966) zurückgeht bzw. in der Monographie zur historischen Phonologie von Carlton (1991) zu finden⁹ ist. Diese Wortlisten, die für alle slawischen Standardsprachen¹⁰ und das rekonstruierte Urslawisch vorliegen, sind generell als eine Adaption der ursprünglichen Swadesh-Liste anzusehen. Im Folgenden wird diese Liste als SMC-Liste (Swadesh-Mel’nyčuk-Carlton) bezeichnet. Die SMC-Liste enthält, wie die folgende Auflistung zeigt, sowohl lexikalisch-semantische als auch grammatische Subgruppen: (1) häufig gebrauchte Adjektive, (2) häufige Tiere und Vögel, (3) häufige Pflanzen, (4) häufig gebrauchte Verben, (5) Verwandtschaftsbezeichnungen, (6) Natur, Werkzeuge, Jagd, (7) Ernährung (8) Körperteile (menschliche, tierische). Die 212 untersuchten Wortformen sind im Folgenden aufgelistet. Im Rahmen der Eruierung von potenziellen Lehnwörtern innerhalb dieses Basiswortschatzes erwies es sich als notwendig, einzelne Einträge aus Carlton (1991: 334-349) zu aktualisieren. Einige der dort angeführten Formen sind aus heutiger Sicht nicht mehr als Bestandteil des synchron verwendeten und gesprochenen Standardslowenischen anzusehen, das aber bewusst im Fokus der vorgenommenen Untersuchung steht. Als formales Kriterium dafür wird das Vorkommen der zu untersuchenden Wortformen in den verfügbaren einsprachigen Wörterbüchern des Slowenischen (SSKJ-1 bzw. SSKJ-2)¹¹ herangezogen. Dies gilt auch für die Extrahierung von veralteten bzw. archaischen Wortformen, die in der Regel in diesen Wörterbüchern durch entsprechende Deskriptoren gekennzeichnet sind. Im Folgenden werden die zu untersuchenden Wortformen der slowenischen SCM-Liste, gegliedert nach den vorgeschlagenen Subgruppen, auf eventuell vorhandene Lehnwörter geprüft. Im Anschluss daran ist der Informationsgehalt dieser Wortliste hinsichtlich der Entlehnungsaktivität des Slowenischen zu bewerten.

4.1.1. Häufig gebrauchte Adjektive

Die SMC-Liste beinhaltet folgende 24 Lemmata: *bel* (‘weiß’), *cel* (‘ganz’), *čist* (‘sauber’), *črn* (‘schwarz’), *črljen* (‘rot’), *dolg* (‘lang’), *dober* (‘gut’), *gorek* (‘bitter’), *kisel* (‘sauer’), *krasen* (‘prächtig’), *ljub* (‘lieb’), *mali* (‘klein’), *mehak* (‘weich’), *mlad* (‘jung’), *poln* (‘voll’), *zdrav* (‘gesund’), *slab* (‘schlecht’, ‘schwach’), *star* (‘alt’), *tuj* (‘fremd’), *velik* (‘groß’), *zelen* (‘grün’), *zel* (‘böse’), *živ*

⁹ Die in Carlton (1991: 334-349) enthaltenen und von Mel’nyčuk (1966) übernommenen orthographischen Fehler wurden für die vorliegende Analyse korrigiert.

¹⁰ Eine 100 Wortformen umfassende Liste für slawische Sprachen (Russisch, Ukrainisch, Weißbrussisch, Slowakisch, Tschechisch, Polnisch, Kaschubisch, Ober- und Niedersorbisch, Polabisch, Altkirchenslawisch, Bulgarisch, Makedonisch, Serbokroatisch, Slowenisch) findet sich in Fodor (1961), der dieses Wortmaterial für die zeitliche Bestimmung der einzelsprachlichen Entwicklung der slawischen Sprachen verwendet. Darüber hinaus findet sich bei Fodor (1961) eine ausführliche Diskussion der Problematik der Erstellung derartiger Wortlisten, wie etwa die fehlende semantische Entsprechung, die Frage des Umgangs mit Synonymen, die stilistische Markierung einiger Wortformen usw. So verweist er z. B. auf eine Vielzahl von möglichen Ausdrücken für das Lexem *Arbeit* innerhalb der ostslawischen Sprachen.

¹¹ Beide Wörterbücher sind online zu finden und entsprechen den jeweiligen Druckausgaben, sodass es berechtigt ist, im Folgenden ohne Anführung der Seitenzahlen zu zitieren. Details dazu im Literaturverzeichnis.

(‘lebendig’) und *žolt* (‘gelb’). Unter diesen häufig gebrauchten¹² Adjektiven finden sich zwei Einträge, die eindeutig aus heutiger Sicht als Archaismen zu identifizieren sind: *črljen* (‘rot’) ist weder im SSKJ-1 noch im SSKJ-2 belegt und findet sich nur bei Pleteršnik (1894–1895) mit der angeführten Bedeutung. Die heute allgemein verwendete Form *rdeč* (‘rot’) kann lt. Snoj (2003: 608) nicht als Lehnwort angesehen werden. Ebenfalls als Archaismus erweist sich *žolt* (‘gelb’), welches im heutigen Slowenischen durch *rumen* (‘gelb’) wiedergegeben wird. Somit lässt sich festhalten, dass innerhalb der häufig gebrauchten Adjektive kein einziges (offensichtlich erkennbares) Lehnwort festzustellen ist.

4.1.2. Häufige Tiere und Vögel

In der Rubrik „Häufige Tiere und Vögel“ finden sich insgesamt 37 Einträge: *jagnje*¹³ (‘Lamm’), *jajce* (‘Ei’), *bober* (‘Biber’), *čebela* (‘Biene’), *bik* (‘Stier’), *črv* (‘Wurm’), *jelen* (‘Hirsch’), *jeseter* (‘Stör’), *gos* (‘Gans’), *jež* (‘Igel’), *kobila* (‘Stute’), *konj* (‘Pferd’), *krava* (‘Kuh’), *koza* (‘Ziege’), *kura* (‘Huhn’), *medved* (‘Bär’), *mravlja* (‘Ameise’), *labod* (‘Schwan’), *orel* (‘Adler’), *osa* (‘Wespe’), *ovca* (‘Schaf’), *ogor* (‘Meeraal’), *pes* (‘Hund’), *prase*¹⁴ (‘Schwein’), *riba* (‘Fisch’), *skot* (‘Vieh’), *svinja* (‘Schwein’), *sraka* (‘Rabe’), *tele* (‘Kalb’), *tur* (‘Auerochse’), *veper* (‘Wildeber’), *volk* (‘Wolf’), *vol* (‘Ochse’), *vrana* (‘Krähe’), *žrebe* (‘Fohlen’), *zmija*¹⁵ (‘Schlange’) und *zver* (‘Raubtier’).

In dieser Rubrik lassen sich folgende Einträge als eindeutige Archaismen identifizieren: *zmija* (‘Schlange’) ist durch *kača*, *veper* (‘Wildeber’) durch die analytische Form *divji merjasec* zu ersetzen. *Skot* in der Bedeutung ‘Vieh’ ist nur mehr bei Pleteršnik (1894–1895) zu finden und somit ebenfalls als veraltet anzusehen (im heutigen Slowenischen lautet die adäquate Form *živina*). In dieser lexikalisch-semanticen Gruppe ist von besonderem Interesse – bedingt durch die Auswahl –, dass auf den ersten Blick ebenfalls nur sehr wenige Lehnwörter zu finden sind. In diesen Fällen ist von einer Entlehnung auszugehen: Lt. Snoj (2003: 240) wäre *jeseter* (‘Stör’) als ein innerslawisches Lehnwort aufzufassen, welches aus anderen (bei Snoj nicht weiter spezifizierten) slawischen Sprachen übernommen wurde. Hinsichtlich von *tur* (‘Auerochse’) ist eine sekundäre Entlehnung aus einer anderen slawischen Sprache sehr wahrscheinlich, wenngleich *tur* z. B. in älteren alpenlawischen Namen (*Turjak* als Toponym) nachweisbar ist. Nicht eindeutig einzuschätzen ist *ogor* (‘Meeraal’), sodass im gegebenen Kontext auch *jegulja* (‘Aal’) in Betracht zu ziehen wäre. Nimmt man *jegulja* als Basis, so ist dieses als Lehnwort zu

¹² Ob es sich hier tatsächlich um häufig gebrauchte Adjektive handelt, kann in Frage gestellt werden. Z. B. hat *zel* im heutigen Slowenischen eine sehr niedrige Frequenz. Allerdings spricht Carlton (1991: 334) in diesem Zusammenhang von *commonly used adjectives*, so dass auch von üblich, gewöhnlich verwendeten Adjektiven gesprochen werden könnte, wenngleich auch diese Umschreibung mit der Vorkommenshäufigkeit korrelieren sollte.

¹³ Bei Carlton (1991: 336) fälschlicherweise *jagne*.

¹⁴ Die heutige stilistisch neutrale Form muss *prašič* lauten, während *prase* im heutigen Kontext nur mehr in übertragener Bedeutung im Sinne von ‘Saukerl’ verwendet wird.

¹⁵ Die von Carlton (1991: 337) angeführte Form *zmija* ist nicht im SSKJ-1 belegt, dort findet sich ausschließlich die Form *kača*.

erachten, da es auf das Istro-Romanische *angvela* (Snoj 2003: 238) zurückgeführt werden kann. Insgesamt ist wiederum von einem relativ geringen Lehnwortanteil in der lexikalisch-semantischen Gruppe „Tiere und Vögel“ auszugehen (3 Lehnwörter aus einer Gesamtmenge von 37).

4.1.3. Häufige Pflanzen

Innerhalb der 34 Lemmata umfassenden lexikalisch-semantischen Gruppe „Häufige Pflanzen“ mit *jablan* (‘Apfelbaum’), *jaboda* (‘Erdbeere’), *breza* (‘Birke’), *bob* (‘Bohne’), *bukev* (‘Buche’), *bez* (‘Holunder’), *čresnja* (‘Kirsche’), *česnjak* (‘Knoblauch’), *drevo* (‘Baum’), *dob* (‘Eiche’), *ječmen* (‘Gerste’), *jela* (‘Tanne’), *jelša* (‘Erle’), *gaber* (‘Weißbuche’), *gruška* (‘Birne’), *klen* (‘Feldahorn’), *koper* (‘Dill’), *cvet* (‘Blüte’), *les* (‘Wald’), *lipa* (‘Linde’), *lan* (‘Leinen’), *malina* (‘Himbeere’), *oreh* (‘Nuss’), *oves* (‘Hafer’), *proso* (‘Hirse’), *pšenica* (‘Weizen’), *rž* (‘Roggen’), *sliva* (‘Zwetschke’, ‘Pflaume’), *smreka* (‘Fichte’), *trava* (‘Gras’), *vrba* (‘Weide’), *želod* (‘Eichel’), *žito* (‘Getreide’) sind in einem ersten Schritt die entsprechenden Archaismen zu identifizieren.

Darunter fallen *jablan* ‘Apfelbaum’ mit heute ausschließlich weiblicher Endung *jablana*, für ‘Holunder’ ist ausschließlich *bezeg* in Gebrauch, die aktuelle Form für ‘Kirsche’ lautet *č(r)esnja*, der ‘Knoblauch’ ist im Standard ausschließlich *česenj* und nicht *česnjak*, wie von Carlton (1991: 338) angegeben. Hinsichtlich von ‘Eiche’ hat sich im Slowenischen eine Dublette herausgebildet: *dob* gilt mittlerweile als veraltet und wird stilistisch neutral durch *hrast* ausgedrückt. Aus Sicht des gegenwärtigen Sprachgebrauchs muss *gruška* ‘Birne’ durch *hruška* und *jela* ‘Tanne’ durch *jelka* ersetzt werden. Die ersetzten Formen unterscheiden sich somit in erster Linie in phonologischer und in wenigen Fällen auch in morphologischer Hinsicht, allerdings bleibt die Wortwurzel in allen Fällen dieselbe. An eindeutigen Lehnwörtern lässt sich in dieser semantisch-lexikalischen Gruppe *bukev* ‘Buche’ identifizieren, welches üblicherweise als Übernahme aus dem Germanischen angesehen wird (Snoj 2003: 64) und einer der ältesten Lehnwortschichten der slawischen Sprachen überhaupt angehört (vgl. dazu ERHSJ I: 230–231).

Ähnliches gilt auch für *čresnja*, welches auf das Vulgärlateinische *ceresia* zurückführbar ist (Snoj 2003: 85). Die Etymologie für *gaber* ‘Weißbuche’ – eine in vielen slawischen Sprachen belegbare Form – ist nicht eindeutig geklärt, sodass unter Umständen eine Entlehnung nicht ausgeschlossen ist (Snoj 2003: 164). Für *hruška* wird aufgrund der unterschiedlichen und nicht im Detail erklärbaren Lautung in den slawischen Sprachen die Möglichkeit einer Entlehnung aus einer unbestimmten östlichen Sprache zugelassen (vgl. Snoj 2003: 213). Insgesamt ergibt sich, dass in dieser Rubrik, inklusive nicht eindeutig geklärter Fälle, von vier Lehnwörtern auszugehen ist und somit ein ähnlicher Anteil wie in der Gruppe Bereich „Häufige Tiere und Vögel“ festzustellen ist.

4.1.4. Häufig gebrauchte Verben

In Anbetracht des in der Lehnwortforschung bekannten Phänomens, wonach Verben weitaus weniger häufig entlehnt werden als z. B. Substantive, ist für diese Gruppe a priori eine geringe Lehnwortaktivität zu vermuten. Die Gruppe besteht aus folgenden 24 Verben: *bežati* ('fliehen'), *biti* ('sein'), *boleti* ('wehtun'), *biti* ('schlagen'), *česati* ('kämmen'), *čuti* ('wachen'), *dati* ('geben'), *držati* ('halten'), *jahati* ('reiten'), *hoteti* ('wollen'), *iti* ('gehen'), *kleti* ('fluchen'), *krmiti* ('füttern'), *mleti* ('mahlen'), *moliti* ('beten'), *mučiti* ('foltern'), *pečti* ('backen'), *sejati* ('säen'), *tresti* ('schütteln'), *tkati* ('weben'), *variti* ('brauen', 'schweißen'), *voziti* ('fahren'), *vezati* ('binden') und *živeti* ('leben'). Alle Verben in der Liste sind im heutigen Standardslowenischen im Gebrauch, sodass keine Adaptionen und Änderungen notwendig sind. Innerhalb dieser Gruppe lässt sich kein einziges Lemma als Lehnwort identifizieren. Dieser Befund deckt sich somit mit der in der Sprachkontaktforschung oftmals diskutierten Annahme, dass es zu einer Entlehnung von Verben in der Regel nur bei einem sehr intensiven Sprachkontakt kommt (vgl. Moravcsik 1978, Matras 2009: 134).

4.1.5. Verwandtschaftsbezeichnungen

Die SCM-Liste umfasst insgesamt 19 Lemmata: *baba* ('Weib', 'Frau'), *brat* ('Bruder'), *človek* ('Mensch'), *ded* ('Großvater'), *dete* ('Kleinkind'), *dever* ('Schwager', 'Mannesbruder'), *hči* ('Tochter'), *ljudje* ('Menschen'), *mati* ('Mutter'), *mož* ('Mann'), *nevesta* ('Braut'), *oče* ('Vater'), *sestra* ('Schwester'), *sin* ('Sohn'), *sveker* ('des Mannes Vater'), *tast* ('Schwiegervater'), *vnuk* ('Enkel'), *zet* ('Schwiegersohn') und *žena* ('(Ehe)Frau'). In dieser Liste sind *dever* und *sveker* eindeutig als Archaismen zu identifizieren (laut SSKJ-1) und im heutigen Slowenischen durch *svak* bzw. *tast* zu ersetzen. Bei Pleteršnik (1894–1895) wird *dever* in der Bedeutung 'Mannesbruder' bzw. 'Beistand und Zeuge der Braut' geführt; *sveker* mit der Bedeutung 'des Mannes Vater' bzw. 'der Schwiegervater der Frau'. Diese Verwandtschaftsbezeichnungen sind im heutigen Slowenischen nicht mehr im Gebrauch, allerdings lassen sich dafür zumindest Belege im 18. und 19. Jh. finden. Bei diesen Lemmata handelt es sich um einen gemeinslawischen Wortschatz und die slowenischen Formen liefern keine Hinweise auf Lehnwörter. Die Zuordnung von *ljudje* zu dieser Gruppe ist eindeutig arbiträr.

4.1.6. Natur, Werkzeuge, Jagd

Der inhaltlich heterogene Bereich Natur, Werkzeuge, Jagd umfasst 32 Einträge: *breg* ('Ufer', 'Hang'), *brana* ('Egge'), *cep* ('Spalte', 'Dreschflegel' u. a. m.), *čad* ('Rauch, Dunst'), *čoln* ('Boot'), *dan* ('Tag'), *dol* ('Tal'), *dom* ('Heim'), *dež* ('Regen'), *dim* ('Rauch'), *duri* ('Tür'), *dvor* ('Hof'), *jezero* ('See'), *gora* ('Berg'), *zvezda* ('Stern'), *kamen* ('Stein'), *leto* ('Jahr'), *mesec* ('Monat'), *noč* ('Nacht'), *ogenj* ('Feuer'),

os ('Achse'), *plug* ('Pflug'), *polje* ('Acker', 'Feld'), *reka* ('Fluß'), *rosa* ('Tau'), *sneg* ('Schnee'), *sonce*¹⁶ ('Sonne'), *veter* ('Wind'), *voda* ('Wasser'), *voz* ('Wagen'), *zemlja* ('Erde') und *zlato* ('Gold'). Aus heutiger Sicht sind die bei Pleteršnik (1894–1895) belegten *cep* 'Spalte, Dreschflegel' und *dol* 'Tal' als Archaismen zu bezeichnen. Die Form *duri* 'Tür' ist ebenfalls als veraltet anzusehen, wenngleich dieses Lexem in den slowenischen Dialekten bis heute weit verbreitet ist. Die heutige standardsprachliche Form dafür lautet *vrata*. Die Gruppe „Natur, Werkzeuge und Jagd“ ist sehr breit angelegt und erwartungsgemäß – der Bereich von konkreten Gegenständen unterliegt sehr oft Entlehnungen – können zumindest zwei Lehnwörter identifiziert werden, wenngleich die Etymologie im Einzelfall wiederum umstritten ist. Zu nennen sind folgende Lemmata: *plug* 'Pflug' als umstrittenes Lehnwort aus dem Gotischen (ERHSJ II: 690), aber es kann in diesem Fall auch nicht ausgeschlossen werden, dass es sich hierbei im Germanischen um eine Entlehnung aus dem Urslawischen handelt (vgl. Snoj 2003: 529). Problematisch ist die Etymologie von *zlato* ('Gold'). Dazu weitere Details in Kap. 4.2.

4.1.7. Ernährung

Die Gruppe Ernährung besteht aus 17 Einträgen und ist hinsichtlich der vorkommenden Wortarten als eine Mischung anzusehen, da sowohl Nomen als auch Verben und Adjektive zu finden sind: *dojiti* ('stillen'), *jesti* ('essen'), *hleb*¹⁷ ('Laib'), *kvass* ('Germ'), *loj* ('Talg'), *maslo* ('Butter'), *med* ('Honig'), *mleko* ('Milch'), *meso* ('Fleisch'), *moka* ('Mehl'), *piti* ('trinken'), *pivo* ('Bier'), *saló* ('Speck'), *sir* ('Käse'), *sit* ('satt'), *testo* ('Teig') und *vino* ('Wein').

In dieser Gruppe sind keine Anpassungen an aktuellere Formen des Standardslowenischen notwendig, da alle Lemmata in dieser Form im Gebrauch sind. In dieser Gruppe gelten eindeutig als Lehnwörter: *hleb* – im Slowenischen ausschließlich 'Laib' – gilt als altes (im Grunde aber bereits im Urslawischen vorhandenes) Lehnwort aus dem Germanischen (Snoj 2003: 206), *pivo* wird in der slowenistischen Linguistik als innerslawisches Lehnwort mit vermutlicher Übernahme aus dem Tschechischen interpretiert (Snoj 2003: 518) und *vino* als Lehnwort aus dem Lateinischen. Somit lassen sich in dieser lexikalisch-semanticen Gruppe drei von 34 Lemmata als Lehnwörter identifizieren, was einem ca. 8%igen Anteil gleichkommt.

4.1.8. Menschliche und tierische Körperteile

¹⁶ Die von Carlton (1991: 345) angeführte Form *solnce* mit historisch-etymologischer Schreibweise ist für das heutige Standardslowenische nicht mehr vorgesehen.

¹⁷ Üblicher ist im Slowenischen die Diminutivform *hlebec*.

Der Bereich menschliche und tierische Körperteile¹⁸ besteht aus 26 Lemmata: *brada* ('Kinn', 'Bart'), *obrv* ('Augenbraue'), *čelo* ('Stirn'), *čeljust* ('Kiefer'), *črevo* ('Darm'), *dlesen* ('Zahnfleisch'), *dlan* ('Handfläche'), *glava* ('Kopf'), *jezik* ('Zunge'), *kost* ('Knochen'), *koža* ('Haut'), *kri* ('Blut'), *noga* ('Fuß', 'Bein'), *noht* ('Nagel'), *nos* ('Nase'), *oko* ('Auge'), *laket* ('Ellbogen'), *palec* ('Daumen', 'große Zehe'), *peta* ('Ferse'), *prst* ('Finger'), *roka* ('Hand', 'Arm'), *srce* ('Herz'), *telo* ('Körper'), *volna* ('Wolle'), *vime* ('Euter') und *zob* ('Zahn'). Hierbei wurden bereits folgende Adaptionen durchgeführt: Die bei Carlton (1991: 347) angeführte Form *dlesna* ('Zahnfleisch') gilt als veraltet, dies gilt auch für *nohat* ('Nagel'). Die bei Carlton (1991: 349) angeführte Form *lakat* ('Ellbogen') lässt sich nur in Pleteršnik (1894–1895), aber weder im SSKJ-1 noch SSKJ-2 belegen. Es handelt sich hierbei in allen Fällen um einen genuinen slawischen Erbwortschatz, ohne einen erkennbaren Lehnwortschatz.

4.1.9. Zwischenresultate

Nach erfolgter Analyse der SMC-Liste sind folgende Resultate festzuhalten: Der reduzierte Basiswortschatz ist nicht als lehnwortresistent zu bezeichnen, wenngleich der Anteil an Lehnwörtern mit 5% insgesamt relativ gering ausfällt (12 Lehnwörter aus einer Gesamtmenge von 212 untersuchten Formen). Die Bereiche, die Lehnwörter aufweisen, sind Tier- und Pflanzenbezeichnungen und der Bereich von Natur, Werkzeugen, Jagd; zu einem geringeren Ausmaß sind auch Lexeme mit einem allgemeinen Bezug zur Ernährung davon betroffen. Festzuhalten ist, dass es sich bei allen identifizierten Lehnwörtern ausschließlich um Substantive handelt, was in Hinblick auf propagierte Entlehnungsskalen als Kennzeichen für einen relativ oberflächlichen Sprachkontakt zu deuten wäre. Zu beachten ist allerdings, dass es sich beim untersuchten Basiswortschatz um – so die gängige Annahme – eine relativ alte lexikalische Schicht handeln sollte, sodass in erster Linie das Entlehnungsverhalten des Slowenischen aus historischen Perioden erfasst wird. Es gilt aber auch festzuhalten, dass einige Lehnwörter entweder hinsichtlich ihrer Etymologie nicht eindeutig geklärt sind oder aber dass es sich dabei um innerslawische Lehnwörter handelt, die mitunter relativ spät ins Slowenische übernommen worden sind.

Zusammenfassend kann daher festgehalten werden, dass es sich bei vielen der identifizierten Lehnwörter um gemeinslawisches Lehngut handelt, d. h. um Lehnwörter, die vermutlich bereits vor der einzelsprachlichen Herausbildung des Slowenischen in das Urslawische gekommen sind und sehr alten Lehnwortschichten angehören. Somit erhält man bei Verwendung der SMC-Liste nur relativ wenig Auskunft über Lehnwortbeziehungen des Slowenischen seit dem Beginn der einzelsprachlichen Entwicklung. Aus dieser Perspektive ist die SMC-Liste hinsichtlich des Informationsgehalts zur

¹⁸ Der harschen Kritik an der Zusammenfassung von menschlichen und tierischen Körperteilen zu einer Gruppe in Haarmann (1990: 6) ist zuzustimmen. Sofern bei der Erstellung eines Basiswortschatzes von einer konzeptuellen Nähe zur menschlichen Lebenswelt ausgegangen wird, so wären die Bereiche von Mensch und Tier sinnvollerweise zu trennen.

Lehnwortaktivität des Slowenischen etwas geringer einzustufen als die von Haspelmath/Tadmor (2009b) vorgeschlagene Basiswortschatzliste mit ca. 1500 Bedeutungen, die u. a. auch neuere und moderne Lebensbereiche umfasst. Abschließend ist des Weiteren festzuhalten, dass selbst innerhalb eines reduzierten Basiswortschatzes nicht von einer absoluten Resistenz gegenüber Entlehnungen gesprochen werden kann, wenngleich dies nicht nur von der Auswahl der Lemmata (die ja im Einzelfall schwer begründbar ist) abhängt, sondern eben auch von dem jeweiligen quantitativen Umfang des Basiswortschatzes, der untersucht wird.

4.2. Lehnwörter innerhalb von Häufigkeitswörterbüchern – Fallbeispiel Kroatisch

Als ein grundlegendes Kriterium bei der Bestimmung eines Basiswortschatzes kann die Vorkommenshäufigkeit von Lexemen herangezogen werden. Die Häufigkeit ist in erster Linie auf das Anwendungsbedürfnis eines Sprechers zurückzuführen und somit ist die Frequenz eindeutig als eine syntagmatische Eigenschaft eines Sprachsystems sowie in direkter Abhängigkeit von dem zugrunde gelegten Textkorpus zu sehen. Die Auswahl eines Textkorpus steht aber wiederum in einem engen Zusammenhang zum Konzept der Repräsentativität, welches darauf abzielt, einen repräsentativen Ausschnitt aus einer Sprache (Standardsprache, Umgangssprache usw.) zu erhalten. Der übliche Weg, einen gewissen Grad an Repräsentativität zu erreichen, wird üblicherweise darin gesehen, möglichst umfangreiches Textmaterial auszuwerten („the more, the better approach“). Zusätzlich spielt aber bei der Erstellung eines großen repräsentativen Korpus in der Regel auch die balancierte Auswahl von unterschiedlichen Registern (Textsorten, Funktionalstile u. Ä. m.) eine große Rolle.

Die Problematik, welche Register als besonders geeignet erscheinen, lässt sich aber nicht ohne Weiteres eindeutig beantworten. Bei der Erstellung von umfangreichen Textkorpora bzw. im slawistischen Kontext besonders relevanten sogenannten Nationalkorpora wird in der Regel versucht, von einer adäquaten Mischung von unterschiedlichen Registern und Textsorten auszugehen, die es erlaubt, zumindest dem Anspruch auf Repräsentativität (in Bezug auf die ausgewählten Funktionalstile) gerecht zu werden. Pionierleistungen auf diesem Gebiet inklusive eines hohen Grades an methodologischer Reflexion gehen in diesem Zusammenhang auf die tschechische Korpuslinguistik zurück, die u. a. auch auf Lesergewohnheiten, die Verbreitung von Literatur in Bibliotheken, d. h. also auch soziologische Informationen zurückgreift, um auf diese Art und Weise die Auswahl an Texten bei der Korpuserstellung zu begründen (vgl. dazu Králík/Šulc 2005, Kučera 2002: 247). Abgesehen von der komplexen Problematik der Textauswahl für die Erstellung von repräsentativen Korpora einer Sprache ist aber den vielfältigen Fortschritten der Korpuslinguistik in den letzten Jahren die Erstellung von umfangreichen Häufigkeitswörterbüchern zu verdanken (vgl. dazu u. a. z. B. Čermák et al. 2007, Ljaševskaja/Šarov

2009, Moguš/Bratanić/Tadić 1999), die in der Regel auf sehr umfangreichen korpuslinguistischen Auswertungen basieren (vgl. dazu detaillierter Čermák/Křen 2005).

Genau an dieser Stelle kann nunmehr wiederum ein Bezug zur Lehnwortforschung hergestellt werden. Ohne Zweifel spielt die Bestimmung der Vorkommenshäufigkeit von Lehnwörtern eine große Rolle, zumal auf diese Weise der Usus und der jeweilige Gebrauch eines Lehnwortes erfasst wird (vgl. generell zur Bedeutung der Frequenz in der Lehnwortforschung Weinreich ³1964: 57, Poplack/Sankoff 1984, Zenner/Speelman/Geeraerts 2014 u. v. m.). Im Folgenden wird es aber nicht um die Vorkommenshäufigkeit von Lehnwörtern an sich gehen, sondern um die Frage, welche und wie viele Lehnwörter innerhalb der hochfrequenten Lexik einer Sprache identifiziert werden können. Auf diese Weise sollen die Möglichkeiten und die Relevanz einer frequenzbasierten Auswahl des Basiswortschatzes aufgezeigt und diskutiert werden. In dem Ansatz von Haspelmath/Tadmor (2009b) spielt die Vorkommenshäufigkeit nur eine sekundäre Rolle, da – sofern überhaupt vorhanden – die Häufigkeit von potenziellen Lehnwörtern a posteriori eruiert wird. Bei der Anwendung eines frequenzbasierten Ansatzes bei der Auswahl des Basiswortschatzes ist darüber hinaus zu untersuchen, welche Art von Lehnwörtern (Herkunft, Alter, Register) auf diese Weise eruiert werden können.

Die Identifikation von Lehnwörtern im hochfrequenten Bereich der Lexik wird im Folgenden anhand des Kroatischen demonstriert. Es steht mit Moguš/Bratanić/Tadić (1999) das Häufigkeitwörterbuch *Hrvatski Čestotni Rječnik* (im Folgenden HČR) zur Verfügung, welches es erlaubt, den verfolgten Ansatz systematisch zu verfolgen. Der Häufigkeitsbereich wird a priori auf die 1500 häufigsten Lemmata beschränkt, die im HČR zu finden sind. Diese Anzahl entspricht in etwa derjenigen, die in Diskussionen um den quantitativen Umfang eines lexikalischen Minimums bzw. Basisgrundwortschatzes immer wieder genannt wird.

Darüber hinaus liefert der HČR wertvolle Informationen darüber, in welchen Funktionalstilen die jeweiligen Lemmata vorkommen. Es werden im HČR folgende Funktionalstile unterschieden: Dramen (D), journalistische Texte (N), Prosa (P), Gedichte (S) und Fachtexte aus Lehrbüchern (U). Für die Erstellung des HČR wurden die in das Korpus eingehenden Wortformen nach vorangehender (manueller) Disambiguierung lemmatisiert, sodass von einer durchaus validen linguistischen Datenbasis ausgegangen werden kann.

Vor der Besprechung der Ergebnisse hinsichtlich der in den ersten 1500 häufigsten Lemmata vorkommenden Lehnwörter ist aber ein genauer Blick auf die Konsistenz und die Zusammensetzung des für das HČR herangezogenen Korpus zu werfen. Es wurde versucht, aus jedem der oben genannten Funktionalstile den jeweils gleichen Anteil (je ca. 200.000 Wortformen) heranzuziehen, um zumindest aus quantitativer Perspektive eine Balance zu erhalten. Hinsichtlich der Zeitspanne ist Textmaterial aus der Periode seit den 30er bis in die 70er Jahre des 20. Jh. erfasst (vgl. HČR 1999: 6).

Im Folgenden findet sich eine Auswertung aller identifizierten Lehnwörter,¹⁹ die innerhalb der ersten 1500 häufigsten Lemmata (vgl. dazu HČR 1999: 19–29) auftreten. Diese werden im Folgenden ihrer Ranghäufigkeit nach angeführt. Die Zahl vor dem Lehnwort ist der jeweilige Rang im HČR (Rang 1 = häufigstes Lemma, 2 = zweithäufigstes etc.). Auf die Nennung der absoluten Häufigkeit von Lehnwörtern wird im gegebenen Zusammenhang verzichtet, da sie nicht von vorrangiger Bedeutung ist. Darüber hinaus werden immer zuerst die Lehnwörter aufgeführt, die in allen untersuchten Funktionalstilen (D-N-P-S-U) vorkommen, und im Anschluss daran diejenigen, die nur in einer eingeschränkten Kommunikationssphäre auftreten.

Als erstes ist eine Gruppe von identifizierbaren Lehnwörtern zu besprechen, die man bedingt und tentativ als Internationalismen²⁰ bezeichnen kann und deren Morphemstamm als eine Entlehnung aus dem Lateinischen bzw. Griechischen anzusehen ist und darüber hinaus in ähnlicher Form jeweils in mehreren Sprachen zu finden ist. Zum Teil ist bei den im Folgenden angeführten Lehnwörtern der Grad der orthographischen und morphologischen Adaptionen bereits sehr weit fortgeschritten, allerdings lassen sich die Stämme relativ problemlos als Lehnwörter aus dem Lateinischen, Griechischen usw. identifizieren. Auf eine etymologische Angabe der potenziellen Gebersprache wird im Folgenden bewusst verzichtet.

Folgende Lehnwörter bzw. Internationalismen lassen sich in den 1500 häufigsten Lemmata des HČR eruieren, die in allen Funktionalstilen (D-N-P-S-U) vorkommen: (160) *organizacija*, (289) *plan*, (323) *sistem*, (357) *problem*, (369) *politika*, (386) *masa*, (418) *tekst*, (423) *motor*, (460) *energija*, (483) *papir*, (527) *grupa*, (536) *električan*, (559) *objekt*, (589) *doktor*, (593) *automobil*, (596) *milijun*, (620) *organ*, (636) *program*, (640) *situacija*, (650) *kapetan*, (689) *revolucija*, (705) *dr.* (als Abkürzung für *doktor*), (739) *kemijski*, (743) *materijalan*, (754) *film*, (793) *cigareta*, (809) *magnetski*, (821) *statistički*, (858) *linija*, (873) *banka*, (875) *metal*, (905) *talijanski*,²¹ (956) *karta*, (967) *tip*, (984) *lak*, (1017) *ideja*,

¹⁹ Lt. den Verfassern des HČR (1999: 11) wurden zwar Fremdwörter nicht in das Wörterbuch aufgenommen, sehr wohl aber phonologisch wie auch morphologisch adaptierte Lehnwörter, womit der HČR für die hier genutzten Zwecke als geeignet erscheint.

²⁰ Die Definition eines Internationalismus ist umstritten und Gegenstand einer lang anhaltenden Diskussion (vgl. dazu Kolwa 2003, Braun 1990). In Glück (2010: 301f.) wird ein Internationalismus allgemein als Entlehnung, die in sehr vielen Sprachen Usus ist, z. B. *Taxi*, *Hotel*, *Computer* usw. definiert. Diese sind für die internationale Verständigung sehr nützlich. Eine, etwas andere, sehr allgemeine Definition für Internationalismus lautet: „[...] sprachliche Einheiten, die ausdrucksseitig und inhaltsseitig (gegebenenfalls nur strukturell und inhaltsseitig) übereinstimmend oder ähnlich in mehr als einer Sprache vorkommen [...]“ (vgl. Bergmann 1995: 272, zitiert nach Schaefer 2003: 86). Diese Definition ist für nah verwandte Sprachen wie Slowenisch und Kroatisch nur eingeschränkt von Bedeutung. In diesem Sinne ist eine Spezifizierung auf einen griechisch-lateinischen Ursprung (vgl. dazu Kolwa 2003: 18) durchaus ein sinnvolles Kriterium. Weitere Definitionsprobleme von Internationalismen betreffen die Anzahl von Sprachen, in denen ein Internationalismus verbreitet sein muss und die Frage der semantischen Übereinstimmung und der Grad der Ähnlichkeit (sowohl in Bezug auf die Form als auch auf den Inhalt).

²¹ Die Einschätzung von *talijanski* als Internationalismus kann durchaus kritisch hinterfragt werden, da es aufgrund der Aphärese von einem Laien nicht ohne Weiteres erkannt wird und hier eventuell berechtigterweise durchaus von

(1050) *evropski*, (1058) *ton*, (1067) *klasa*, (1115) *avion*, (1151) *engleski*, (1168) *minuta*, (1193) *poezija*, (1220) *ministar*, (1234) *reakcija*, (1273) *scena*, (1280) *francuski*, (1300) *karakter*, (1327) *turist*, (1363) *kriza*, (1373) *tema*, (1421) *muzika*, (1433) *bomba*, (1452) *moderan* und (1473) *partija*.

Demgegenüber finden sich folgende Internationalismen nur in einigen wenigen Funktionalstilen: (279) *ekonomski*, (358) *funkcija*, (361) *materijal*, (367) *turistički*, (417), *politički*, (432) *rezultat*, (459) *atom*, (470) *element*, (471) *metoda*, (487) *proces*, (508) *temperatura*, (511) *agencija*, (544) *turizam*, (551) *tehnički*, (558) *hotel*, (606) *nacionalan*, (622) *aktivnost*, (629) *socijalistički*, (648) *akcija*, (693) *delegacija*, (706) *kultura*, (730) *interes*, (744) *magnetski*, (824) *vagon*, (829) *struktura*, (833) *konferencija*, (874) *industrija*, (886) *faktor*, (891) *američki*, (898) *komitet*, (924) *organizam*, (931) *furnir*, (952) *transformator*, (964) *kulturan*, (965) *praksa*, (975) *dokument*, (981) *republika*, (1032) *sekretar*, (1041) *analiza*, (1060) *elektron*, (1093) *federacija*, (1104) *kredit*, (1117) *faza*, (1164) *direktor*, (1165) *general*, (1174) *armija*, (1200) *elemenat*, (1215) *delegat*, (1238) *tehnika*, (1248) *realan*, (1254) *kapacitet*, (1266) *dolar*, (1270) *kvaliteta*, (1279) *baza*, (1306) *socijalan*, (1343) *organizirati*, (1346) *sindikata*, (1361) *karakteristika*, (1372) *spektar*, (1377) *biološki*, (1379) *fizički*, (1437) *kristal*, (1440) *opera*, (1442) *plastičan*, (1451) *komunist*, (1470) *komisija* und (1497) *industrijski*.

Auf der Basis dieser Daten gelangt man bereits zu einem ersten Zwischenresultat: Insgesamt lassen sich auf diese Weise innerhalb des auf 1500 Lemmata beschränkten Basiswortschatzes 123 Internationalismen²² eruieren, davon 53 Lehnwörter, die in allen und 70 Lehnwörter, die nur in wenigen Funktionalstilen vorkommen. Somit wird deutlich, dass innerhalb der 1500 häufigsten Lemmata Internationalismen dominieren, die in einer eingeschränkten (fachsprachlichen) Kommunikationssphäre vorkommen. Die eruierten Internationalismen sind aber bei genauerer Betrachtung in der Regel keiner bestimmten Fachsprache zuzuordnen, sondern sind insgesamt als weitverbreitete Internationalismen im Kroatischen zu sehen, die nicht als in irgendeiner Weise besonders markiert erscheinen. Vielmehr handelt es sich dabei um einen Lehnwortschatz aus den Bereichen Politik, Technik, Kultur, Wissenschaft und Verwaltung, der im Kroatischen von derartigen lateinisch-griechischen Internationalismen dominiert ist.

An weiteren Lehnwörtern lassen sich aus der Liste der 1500 häufigsten Lemmata des HČR folgende extrahieren: (176) *soba* ('Zimmer'), welches lt. ERHSJ (III: 299) im Kroatischen als Lehnwort aus dem Althochdeutschen über die Mittlersprache Ungarisch übernommen wurde, (184) *boja* ('Farbe') als Turzismus (vgl. ERHSJ I: 182), *stroj* ('Maschine') als Bohemismus (ERHSJ III: 347), (308) *staklo* ('Glas') (ebenso (1004) *staklen* als Derivat) eine vermutliche alte Übernahme aus dem Gotischen

einem romanischen Lehnwort auszugehen wäre. Die serbische Form *italijanski* wäre dem gegenüber aber eindeutig als Internationalismus zu identifizieren.

²² Nicht als Internationalismus wurde (762) *društveno-politički* ('gesellschaftlich-politisch') aufgefasst, da zwar *politički* als Lehnwort zu sehen ist, nicht aber das Kompositum als Ganzes.

(allerdings handelt es sich hierbei um eine Übernahme aus dem Urslawischen)²³ (ERHSJ III: 325), (322) *dinar* ('Dinar') als Lehnwort aus dem Romanischen bzw. aus dem Griechischen (ERHSJ I: 405–406), (341) *roba* ('Ware') ein Romanismus (ERHSJ III: 151), (377) *vino* ('Wein') (vgl. auch (1447) *vinograd* ('Weingarten')) als bereits im Altkirchenslawischen belegbares Lehnwort (ausführlicher ERHSJ III: 594–595 und insbesondere Snoj 2003: 822), ebenso (384) *škola* (ebenso die Adjektivableitung (1296) *školski*) (vgl. dazu ERHSJ III: 272), (403) *knjiga* ('Buch') (auch (909) *književnost* und (957) *književan* als Ableitungen) als gemeinslawisches Lehnwort mit unterschiedlichen etymologischen Auslegungen, in der Regel aber als Übernahme aus dem Persischen (vgl. dazu ERHSJ II: 109–110) verstanden, (457) *sigurno* ('sicher') (und auch (1123) *sigurnost*) als Entlehnung aus dem Romanischen (Details dazu siehe ERHSJ III: 233–234), (498) *miris* ('Duft') (ebenso das davon abgeleitete Verb (1190) *mirisati*) gilt als Gräzismus (ERHSJ II: 428–429), (565) *cilj* ('Ziel') als Lehnwort aus dem Mittelhochdeutschen (Snoj 2003: 73), (614) *krevet* ('Bett') lt. ERHSJ (II: 192) ein Balkanturzismus mit griechischem Ursprung, (677) *para* ('Geld') ebenfalls ein Balkanturzismus mit persischem Ursprung (ERHSJ II: 604), (688) *hajde* ('auf!', 'los!') ein Balkanturzismus (ERHSJ II: 649), (830) *crkva* ('Kirche') mit bekanntlich umstrittener Etymologie, d. h. entweder als Lehnwort aus dem Althochdeutschen oder aber aus dem Gotischen (Details dazu vgl. ERHSJ I: 275) interpretierbar, (845) *barem* ('wenigstens') (auch Variante (918) *bar*) als ein Turzismus mit persischem Ursprung (ERHSJ I: 109), (869) *čelik* ('Stahl') als Turzismus (ERHSJ I: 304), (888) *ruža* ('Rose') als Entlehnung aus dem Deutschen bzw. Lateinischen (Snoj 2003: 631), (906) *cipela* ('Schuh') als Lehnwort aus dem Ungarischen (*cipellő*), aber mit einer lateinischen (*zipellus*) Wurzel (ERHSJ I: 268), (933) *komad* ('Stück') ein Balkangräzismus (ERHSJ II: 131), (944) *sapun* ('Seife') ein am Balkan weiterverbreiteter Romanismus (ERHSJ III : 202–203), (971) *biskup* ('Bischof') wird im ERHSJ (I: 157) als Gräzismus und weniger als Latinismus interpretiert²⁴, (997) *ulje* ('Öl') als Lehnwort aus dem Lateinischen (ERHSJ III: 543), (1011) *rakija* ('Schnaps') als balkanischer Turzismus mit arabischem Ursprung (ERHSJ II: 102), (1081) *makar* ('wenn auch', 'wenn gleich') als allgemein verbreiteter Balkanismus mit italienischer Wurzel (ERHSJ II: 352), (1094) *kuhinja* ('Küche') als Entlehnung aus dem Althochdeutschen (ERHSJ II: 224), (1119) *košulja* ('Hemd') als allgemein verbreiteter lateinischer Balkanismus (ERHSJ II: 58), (1130) *križ* ('Kreuz') als sehr wahrscheinliche Entlehnung aus dem Lateinischen (ERHSJ II: 201–202), (1201) *kralj* ('König') als alte Entlehnung und etymologisch vermutlich zurückführbar auf Karl den Großen (Snoj 2003: 317), (1217) *kutija* ('Schachtel') ein Balkanturzismus mit griechischem Ursprung (ERHSJ II: 248), (1265) *boca* ('Flasche')

²³ Der Umgang mit bereits im Urslawischen bzw. Altkirchenslawischen vorhandenen Lehnwörtern Bedarf einer gesonderten Abhandlung. Im Grunde genommen liefern diese Lehnwörter nur eingeschränkte Informationen über die lexikalische Entwicklung des Kroatischen.

²⁴ Demgegenüber wird im Slowenischen *škof* als Entlehnung aus dem Althochdeutschen interpretiert, wenngleich hier das Deutsche insbesondere als Mittler aus dem Lateinischen auftritt (vgl. Snoj 2003: 728–729).

als romanisches Lehnwort (ERHSJ I: 177), (1269) *kip* ('Statue') lt. Snoj (2003: 272) ein Lehnwort aus dem Ungarischen, welches aber auf das Alttürkische zurückgeführt werden kann, (1305) *puška* ('Gewehr') als Entlehnung aus dem Bairischen (vgl. Snoj 2003: 595, ERHSJ III: 71–72), für (1474) *pukovnik* ('Oberst') (ERHSJ I: 70–71) und (1481) *gomila* ('Haufen, Menge') werden unterschiedliche etymologische Auslegungen angeboten. Im Falle von *pukovnik* ist überhaupt der Status als Lehnwort zu klären, da hierbei nur der Stamm als bereits im Urslawischen bekanntes Lehnwort interpretiert werden kann. Für *gomila* ist lt. ERHSJ (I: 588) die etymologische Herkunft unklar.

Darüber hinaus sind folgende mehr oder weniger etymologisch umstrittene Lehnwörtern anzuführen: *olovka* ('Bleistift') < *olovo* ('Blei') (vgl. dazu die unterschiedlichen etymologischen Auslegungen in ERHSJ II: 554), *klupa* ('Bank', 'Sitzvorrichtung') wird bei ERHSJ (I: 103) auf einen slawischen Stamm zurückgeführt, während Snoj (2003: 283) auf eine mögliche Wurzel aus dem Altpreußischen verweist, aber beide sind sich einig, dass die jeweiligen etymologischen Erklärungen nicht zufriedenstellend sind; ähnlich unklar sind die Verhältnisse für *lanac* ('Kette'), welches als mögliche alte Entlehnung aus dem Deutschen interpretiert werden kann (ERHSJ II: 266) und für *tuđ* ('fremd'), welches u. a. auf ein Lehnwort aus dem Gotischen zurückzuführen ist, aber lt. ERHSJ (III: 519) auf eine gemeinsame indogermanische Wurzel verweist.

In Abhängigkeit der Interpretation von (130) *bog* ('Gott') als einem sehr alten Lehnwort (im Altkirchenslawischen belegt und mit etymologischen Wurzeln im Altpersischen bzw. Altindischen) ergibt sich zusammen mit den entsprechenden Ableitungen (852) *bogat* ('reich'), (853) *božji* ('göttlich') und (117) *zbog* ('wegen') eine kompakte Gruppe mit einer alten urslawischen Lehnwortschicht (vgl. dazu ERHSJ I: 178–181, Snoj 2003: 48–49). Als problematisch erweisen sich aus etymologischer Sicht die Bezeichnungen für folgende Metalle: (851) *zlato* ('Gold') und (815) *željezo* ('Eisen'), beides im Kroatischen hochfrequente Wortformen (ebenso (568) *zlatan* ('goldig')). Im Falle von *zlato* wird eine Entlehnung nicht ausgeschlossen bzw. u. a. auf Gemeinsamkeiten mit dem Gotischen bzw. Germanischen verwiesen (ERHSJ III: 656–657, Snoj 2003: 855), während bei *željezo* (ERHSJ III: 676–677) unklare Lautverhältnisse bei baltoslawischen Formen zu verzeichnen sind, sodass auch von einer Entlehnung auszugehen ist. Anzumerken ist aber, dass sowohl *bog* als auch *zlato* und *željezo* als Teil der gemeinlawischen Lexik aufzufassen sind und somit im Grunde keine Informationen über die Sprachkontakte des Kroatischen im eigentlichen Sinne liefern, sondern über Kontakte vor der einzelsprachlichen Entwicklung der slawischen Sprachen.

Nach dieser Aufzählung von Lehnwörtern, die innerhalb der 1500 häufigsten Lemmata des Kroatischen identifiziert werden konnten, kann nunmehr zu einer inhaltlichen Interpretation übergegangen werden. Zu beginnen ist mit den quantitativen Verhältnissen: Zählt man alle als Lehnwörter identifizierten Wortformen, inkl. der Gruppe mit einer umstrittenen Etymologie, so ergibt

sich eine Anzahl von 176 Lehnwörtern, davon sind 122 Internationalismen und 54 weitere Lehnwörter. Dies bedeutet auf die Gesamtmenge der untersuchten Lemmata gerechnet einen Gesamtanteil von gerundet 12%. Ob des Fehlens von vergleichbaren Untersuchungen kann die Bedeutung dieses Wertes einstweilen nicht näher spezifiziert werden. Dennoch kann der Befund zumindest als bemerkenswert bezeichnet werden, denn es gilt zu berücksichtigen, dass die hochfrequente Lexik einer Sprache immer aus einer Vielzahl von Synsemantika bzw. Funktionswörtern besteht, die für die grammatische Organisation von Texten zuständig sind. Es sind aber gerade Synsemantika, die in Regel als relativ resistent gegenüber Entlehnungen angesehen werden bzw. nur im Falle eines sehr intensiven Sprachkontaktes entlehnt werden.

Für das Kroatische stellt sich aber heraus, dass in erster Linie Autosemantika entlehnt wurden, wie dies u. a. anhand der Vielzahl von eruierten Internationalismen abzulesen ist, bei denen es sich zum größten Teil um Nomen, Adjektive und Verben handelt. Als weiteren Befund gilt es festzuhalten, dass aus der Liste der 1500 häufigsten Lemmata Lehnwörter zu eruieren sind, die durchaus auch in der gesprochenen Umgangssprache zu finden sind (z. B. *soba, boja, boca, rakija, sapun, lanac, košulja, cipela, ruža, makar, hajde* usw.). D. h. der frequenzbasierte Ansatz bei der Bestimmung des Basiswortschatzes erweist sich in diesem Fall als durchaus tragfähige Möglichkeit, um im Usus verankerte Lehnwörter zu eruieren. Neben Internationalismen und allgemein in der kroatischen Umgangssprache verbreiteten Lehnwörtern lassen sich aus den 1500 häufigsten Lemmata des Weiteren gemeinslawische Lehnwörter (z. B. u. a. *knjiga, bog, bogat, zlato, željezo, vino, staklo*, partiell *križ, kuhinja* u. a. m.) eruieren, die allerdings – wie bereits erwähnt – im Grunde genommen (die Frage von semantischen Verschiebungen und Bedeutungsveränderungen ist an dieser Stelle einstweilen auszublenden) keine Auskunft über die Sprachkontakte des Kroatischen geben, sondern vielmehr auf ältere Sprachkontakte des Urslawischen verweisen. Allerdings ist die Frage – insbesondere für das Slowenische und das Kroatische durchaus von hoher Relevanz – im Einzelfall genau zu überprüfen, ob und welche Formen aus dieser Schicht tatsächlich eine bestimmte Form von Kontinuität aufweisen und nicht etwa im 19. Jh. im Rahmen der bekannten puristischen Aktivitäten bzw. einer bewussten Hinwendung zu „slawischen Modellen“ reaktiviert wurden.

Somit lässt sich abschließend festhalten, dass das Heranziehen von Häufigkeitswörterbüchern trotz einer im Detail anzubringenden Kritik (Frage der Zusammensetzung des für die Erstellung eines Häufigkeitswörterbuches herangezogenen Korpus, Lemmatisierung, Disambiguierung, Umfang des untersuchten Korpus) durchaus als ein geeignetes Vorgehen erscheint, um jeweils unterschiedliche Lehnwortschichten innerhalb einer Sprache aufzudecken. Damit ergibt sich auch eine durchaus lohnenswerte „usage-based“ Perspektive auf die lexikalische Struktur und Lehnwörter von slawischen Standardsprachen.

5. Zusammenfassung

Als Ausgangspunkt für die Diskussion einiger Probleme aus dem Bereich der aktuellen empirischen Lehnwortforschung wurde das von Haspelmath/Tadmor (2009a) initiierte Projekt herangezogen. Das grundsätzliche methodologische Vorgehen, welches aus der Untersuchung eines abgeschlossenen Basiswortschatzes mit ca. 1500 Bedeutungen besteht und als Grundlage für das Eruiere der jeweiligen Lehnwörter herangezogen wird, kann insgesamt als innovativ bezeichnet werden, wenngleich im Detail bei zukünftigen Untersuchungen weitere Fragen zu reflektieren sind (Auswahl des Basiswortschatzes, Frage des Umgangs mit vorhandenen Synonymen, Fehlen bzw. Vielzahl von Übersetzungsäquivalenten u. v. m.). Im vorliegenden Beitrag wurden in diesem Zusammenhang zwei ausgewählte Problemkomplexe im Detail diskutiert.

Einerseits geht es um die Frage der Auswahl des zu untersuchenden Basiswortschatzes, der mit der Begrenzung von ca. 1500 Bedeutungen und der vorgelegten semantisch-lexikalischen Gliederung in 24 Gruppen als arbiträr erscheinen kann. Die besondere Herausforderung ist die weitverbreitete Annahme, wonach es einen Basiswortschatz einer Sprache gibt, der sich als besonders resistent gegenüber Entlehnungen erweist. Die Kriterien der Erstellung dieses Basiswortschatzes sind aber weder durch kommunikativ-pragmatische Faktoren noch durch einen frequenzbasierten Ansatz begründet, sondern orientieren sich mehr oder weniger an einer intuitiven Auswahl. Wie gezeigt werden konnte, geht das Vorgehen von Haspelmath/Tadmor (2009b) direkt auf einen älteren Ansatz aus dem Bereich der in der Linguistik durchaus umstrittenen Glottochronologie zurück. In diesem Bereich wird als Vergleichsbasis der Bestimmung der genetischen Ähnlichkeit von Sprachen in der Regel mit den sogenannten Swadesh-Listen gearbeitet. Im vorliegenden Beitrag wurde eine derartige für das Slowenische adaptierte Swadesh-Liste (in Anlehnung an die beteiligten Autoren als SCM-Liste bezeichnet) mit etwas über 200 Bedeutungen herangezogen, um der Frage nachzuspüren, welche und wie viele Lehnwörter innerhalb dieser eruiert werden können. Auf der Basis dieses reduzierten Basiswortschatzes lässt sich in der Tat kaum pauschal von einer Entlehnungsresistenz sprechen. Vielmehr können wir – wenngleich in einem quantitativ sehr eingeschränkten Ausmaß – im Slowenischen in der SCM-Liste einige Lehnwörter identifizieren, die sich durch ein relativ hohes Alter auszeichnen und z.T. als Erbwörter aus dem Urslawischen anzusehen sind. Die Frage, ob die SCM-Liste im Vergleich zur umfangreichen Datenbasis von 1500 Bedeutungen den gleichen Informationsgehalt hinsichtlich der Lehnwortaktivität einer Sprache aufweist oder nicht, lässt sich – solange keine Analyse dieser Liste vorliegt – nicht eindeutig beantworten. Es ist aber davon auszugehen, dass ohne Zweifel eine umfangreichere Wortliste die Wahrscheinlichkeit für das Vorkommen von Lehnwörtern erhöhen wird. Als zentrales Resultat bleibt aber in jedem Fall festzuhalten, dass der Basiswortschatz (sei es nun mit 200 oder 1500 Bedeutungen) eben nicht generell als

entlehnungsresistent bezeichnet werden kann, eine Reduzierung des untersuchten Basiswortschatzes auf wenige Einheiten aber nur bedingt als empfehlenswert erscheint.

Als eine konzeptuell anderweitig ausgerichtete Alternative wird die Anwendung eines frequenzbasierten Ansatzes bei der Auswahl des Basiswortschatzes vorgeschlagen. Das Vorgehen besteht in diesem Fall aus der Analyse der hochfrequenten Lexik einer Sprache, die als Basis dient, um innerhalb deren die entsprechenden Lehnwörter zu eruieren. Dieser Ansatz, der sich zumindest hinsichtlich der Auswahlkriterien als intersubjektiv nachvollziehbar erweist, steht aber in direkter Abhängigkeit davon, welches Textmaterial für die Bestimmung der hochfrequenten Lexik einer Sprache herangezogen wird. Ein Zugang zu dieser Problematik besteht aus der Analyse von Häufigkeitwörterbüchern, die auf – insbesondere in den letzten Jahren – für viele slawische Sprachen erstellten umfangreichen, „repräsentativen“ Nationalkorpora basieren. Eine Analyse der 1500 häufigsten Lemmata (deren Begrenzung auf 1500 als arbiträr erscheinen mag, aber zumindest durch sprachdidaktische Vorgaben hinsichtlich eines Grundwortschatzes zu begründen ist) aus dem kroatischen Häufigkeitwörterbuch (HČR) hinsichtlich vorkommender Lehnwörter fördert generell interessante Befunde zu Tage. Wie gezeigt werden konnte, beträgt der Anteil von Lehnwörtern etwas über 10%, wobei der eruierte Lehnwortschatz von unterschiedlicher Zusammensetzung ist (Internationalismen, Lehnwörter, die in der mündlichen Umgangssprache verwendet werden, und Lehnwörter, die aus älteren historischen Perioden des Kroatischen stammen bzw. überhaupt wiederum als gemeinslawisches Lehngut zu qualifizieren sind). Diese ersten tentativen Resultate für das Kroatische können dahingehend interpretiert werden, dass die Eruierung von Lehnwörtern in einem frequenzbasierte Basiswortschatz durchaus als valide alternative Methode der empirischen Untersuchung des Lehnwortprofils einer Sprache erscheint.

6. Literatur

- ATKINSON, QUENTIN D.; GRAY, RUSSELL D. (2006): How old is the Indo-European language family? Progress or more moths to the flame? In: Forster, Peter; Renfrew, Colin (eds.): *Phylogenetic methods and the prehistory of languages*. Cambridge: McDonald Institute for Archaeological Research, 91–109.
- BAKKER, DIK ET AL. (2009): Adding typology to lexicostatistics: A combined approach to language classification. In: *Linguistic Typology* 13, 1, 169–181.
- BARTELS, HAUKE (2009): Loanwords in Lower Sorbian, a Slavic language of Germany. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 304–323.
- BRAUN, PETER (1990): Internationalismen – Gleiche Wortschätze in europäischen Sprachen. In: Braun, Peter; Schaefer, Burkhard; Volmert, Johannes (eds.): *Internationalismen. Studien zur interlingualen Lexikologie und Lexikographie* (= Reihe germanistische Linguistik 102). Tübingen: Niemeyer, 13–33.
- BUCK, CARL DARLING (1949): *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago: University of Chicago Press.

- BUßMANN, HADUMOD (ed.) (2008): *Lexikon der Sprachwissenschaft*. Vierte, durchgesehene und bibliographisch ergänzte Auflage. Stuttgart: Kröner.
- CARLTON, TERENCE R. (1991): *Introduction to the phonological history of the Slavic Languages*. Columbus, Ohio: Slavica.
- ČERMÁK, FRANTIŠEK ET AL. (2007): *Frekvenční slovník mluvené češtiny*. Praha: Univerzita Karlova v Praze.
- ČERMÁK, FRANTIŠEK; KŘEN, MICHAL (2005): New Generation Corpus-Based Frequency Dictionaries: The Case of Czech. In: *International Journal of Corpus Linguistics* 10, 4, 453–467.
- DYEN, ISIDORE (1964): On the validity of comparative lexicostatistics. In: Lunt, Horace G. (ed.): *Proceedings of the ninth international congress of linguistics*. Cambridge, Mass., August 27–31, 1962. London – The Hague – Paris: Mouton, 238–252.
- EMBLETON, SHEILA (2000): Lexicostatistics/Glottomorphology: from Swadesh to Sankoff to Starostin to future horizons. In: Renfrew, Colin; McMahon, April; Trask, Larry (eds.): *Time depth in historical linguistics. Volume 1*. Cambridge: McDonald Institute for Archaeological Research, 143–165.
- EMBLETON, SHEILA (1986): *Statistics in historical linguistics* (= Quantitative Linguistics 30). Bochum: Brockmeyer
- ERHSJ I, II, III = Skok, Petar (1971–1974): *Etimologijski rječnik hrvatskoga ili srpskoga jezika*. Zagreb: JAZU.
- FODOR, ISTVÁN (1961): The Validity of Glottomorphology on the Basis of the Slavonic Languages. In: *Studia Slavica* 7, 295–346.
- GLÜCK, HELMUT (ed.) (2010): *Metzler Lexikon Sprache*. 4. aktualisierte und überarbeitete Auflage. Stuttgart – Weimar: Metzler.
- GUDSCHINSKY, SARAH C. (1956): The ABC's of Lexicostatistics (Glottomorphology). In: *Word* 12, 175–210.
- HAARMANN, HARALD (1990): "Basic" Vocabulary and Language Contacts: The Disillusion of Glottomorphology. In: *Indogermanische Forschungen* 95, 1–37.
- HASPELMATH, MARTIN (2009): Lexical borrowing: Concepts and issues. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 35–54.
- HASPELMATH, MARTIN; TADMOR, URI (eds.) (2009a): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter.
- HASPELMATH, MARTIN; TADMOR, URI (2009b): The Loanword Typology project and the World Loanword Database. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 1–34.
- HEGGARTY, PAUL (2010): Beyond lexicostatistics: How to get more out of 'word list' comparisons. In: *Diachronica* 27, 2, 301–324.
- HOCK, HANS HENRICH; JOSEPH, BRIAN D. (1996): *Language history, language change, and language relationship. An introduction to historical and comparative linguistics* (= Trends in Linguistics. Studies and Monographs 93). Berlin – New York: Mouton De Gruyter.
- HOIJER, HARRY (1956): Lexicostatistics: a Critique. In: *Language* 32, 49–60.

- HYMES, DELL (1971): Morris Swadesh: From the first Yale school to world history. In: Sherzer, Joel (ed.): Swadesh, Morris (1971): *The origin and diversification of language*. Chicago: Aldine Atherton, 228–270.
- KOLWA, ANDREA (2003): Zur Geschichte der Internationalismen-Forschung. In: Braun, Peter; Schaeder, Burkhard; Volmert, Johannes (eds.): *Internationalismen II. Studien zur interlingualen Lexikologie und Lexikographie* (= Reihe germanistische Linguistik 246). Berlin: de Gruyter, 13–21.
- KRÁLÍK, JAN; ŠULC, MICHAL (2005): The Representativeness of Czech Corpora. In: *International Journal of Corpus Linguistics* 10, 3, 357–366.
- KROHN, DIETER (1992): *Grundwortschätze und Auswahlkriterien. Metalexikographische und fremdsprachendidaktische Studien zur Struktur und Funktion deutscher Grundwortschätze* (= Göteborger germanistische Forschungen 34). Göteborg: Acta Univ. Gothoburgensis.
- KUČERA, KARL (2002): The Czech National Corpus: Principles, Design, and Results. In: *Literary and Linguistic Computing* 17, 2, 245–257.
- LJAŠEVSKAJA, OL'GA N.; ŠAROV, SERGEJ. A. (eds.) (2009): *Častotnyj slovar' sovremennogo ruskogo jazyka (na materialach Nacional'nogo korpusa ruskogo jazyka)*. Moskva: Azbukovnik. <<http://dict.ruslang.ru/freq.php>> [aufgerufen am 01.07.2015]
- MATRAS, YARON (2009): *Language contact*. Cambridge: Cambridge Univ. Press.
- MCMAHON, APRIL; HEGGARTY, PAUL; MCMAHON, ROBERT; SLASKA, NATALIA (2005): Swadesh Sublists and the benefits of borrowing: an Andean case study. In: *Transactions of the Philological Society* 103, 2, 147–170.
- MCMAHON, APRIL M. (1994): *Understanding language change*. Cambridge: Cambridge Univ. Press.
- MEL'NYČUK, OLEKSANDR S. (ed.) (1966): *Vstup do porivnjal'no-istoričnogo vyvčennja slov'jans'kych mov*. Kyjiv: Naukova Dumka.
- MOGUŠ, MILAN; BRATANIĆ, MAJA; TADIĆ, MARKO (1999): *Hrvatski čestotni rječnik*. Zagreb: Školska knjiga.
- MORAVCSIK, EDITH A. (1978): Language Contact. In: Joseph H. Greenberg (ed.): *Universals of human language. Volume 1: Method and Theory*. Stanford: Stanford University Press, 93–122.
- PLETERŠNIK, MAKS (1894–1895): *Slovensko-nemški slovar*. Ljubljana: Knezoškofijstvo. <<http://www.fran.si/136/maks-pletersnik-slovensko-nemski-slovar>> [aufgerufen am 01.07.2015]
- POPLACK, SHANA; SANKOFF, DAVID (1984): Borrowing: the synchrony of integration. In: *Linguistics* 22, 99–135.
- SCHAEDER, BURKHARD (2003): Neuerlicher Versuch einer theoretischen und methodischen Grundlegung der Internationalismen-Forschung. In: Braun, Peter; Schaeder, Burkhard; Volmert, Johannes (eds.): *Internationalismen II. Studien zur interlingualen Lexikologie und Lexikographie* (= Reihe germanistische Linguistik 246). Berlin: de Gruyter, 71–107.
- SCHNÖRCH, ULRICH (2002): *Der zentrale Wortschatz des Deutschen. Strategien zu seiner Ermittlung, Analyse und lexikografischen Aufarbeitung* (= Studien zur deutschen Sprache 26). Tübingen: Narr.
- SNOJ, MARKO (2003): *Slovenski etimološki slovar. 2. pregladana in dopolnjena izdaja*. Ljubljana: Modrijan.
- SSKJ-1 = *Slovar slovenskega knjižnega jezika* (1970–1991). Ljubljana: Slovenska Akademija Znanosti in Umetnosti, Inštitut za slovenski jezik. <<http://bos.zrc-sazu.si/sskj.html>> [aufgerufen am 01.07.2015]

- SSKJ-2 = *Slovar Slovenskega Knjižnega Jezika*. (2014): *Prva knjiga A-Pa & Druga knjiga Pe-Ž*. Ljubljana: Cankarjeva založba. <<http://www.sskj2.si/>> [aufgerufen am 01.07.2015]
- SLASKA, NATALIA (2005): Lexicostatistics away from the armchair: handling people, props and problems. In: *Transactions of the Philological Society* 103, 2, 221–242.
- SWADESH, MORRIS (1950): Salish internal relationships. In: *International Journal of American Linguistics* 16, 4, 155–167.
- SWADESH, MORRIS (1952): Lexico-Statistic Dating of Prehistoric Ethnic Contacts. With Special Reference to North American Indians and Eskimos. In: *Proceedings of the American Philosophical Society* 96, 4, 452–463.
- SWADESH, MORRIS (1955): Towards greater Accuracy in Lexicostatistic Dating. In: *International Journal of American Linguistics* 21, 121–137.
- SWADESH, MORRIS (1971): What is Glottochronology? In: Sherzer, Joel (ed.): Swadesh, Morris (1971): *The origin and diversification of language*. Chicago: Aldine Atherton, 271–284.
- TADMOR, URI (2009): Loanwords in the world's languages: Findings and results. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 55–75.
- THOMASON, SARAH GREY (2001): *Language contact*. Edinburgh: Edinburgh University Press.
- THOMASON, SARAH GREY; KAUFMAN, TERRENCE (1988): *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- TULDAVA, JUHAN (1998): *Probleme und Methoden der quantitativ-systemischen Analyse* (= *Quantitative Linguistics* 59). Trier: Wissenschaftlicher Verlag Trier.
- WEINREICH, URIEL (³1964): *Languages in contact: findings and problems*. London: Mouton.
- ZENNER, ELINE; SPEELMAN, DIRK; GEERAERTS, DIRK (2014): Core vocabulary, borrowability and entrenchment: A usage-based onomasiological approach. In: *Diachronica* 31, 1, 74–105.

Stefan Michael Newerkla (Wien)

Kernproblembereiche der Lehnwortforschung anhand von Beispielen aus dem Kontaktbereich Deutsch, Tschechisch und Slowakisch

Abstract

The study of loan words in the contact area of German, Czech and Slovak has led to various questions, which are fundamental for any analysis of loan relations. In our article we will discuss some of the core issues and illustrate them with concrete examples. Apart from the (often rather unclear) temporal depth and the complexity of the examination of intra-Slavic loan relations, we make reference to the (also often not so clear) frequency of loanwords and describe key problem areas of standardization processes in the context of lexical borrowings. Among other things we mention parallel acquisitions, the spatial distribution of linguistic commonalities and cultural similarities, the possibility of subsequent contamination, multiple borrowings or successive adjustments of loanwords to the developments in the donor language, various semantic modifications within a lexical-semantic field and possible blurring of findings due to hidden multilingualism.

1. Einleitung

Sprachliche Kontakte zwischen dem Deutschen und westslawischen Varietäten bestehen bekanntermaßen schon seit mehr als einem Jahrtausend. Davon zeugt unter anderem auch der bedeutende Anteil von lexikalischem Lehnwort im Wortschatz des Tschechischen und Slowakischen. Wissenschaftliche Untersuchungen aus dem 19. Jh. und der ersten Hälfte des 20. Jh. widmeten sich dabei vor allem der allgemeinen Problematik der ältesten Lehnbeziehungen zwischen dem Slawischen und Germanischen bzw. Deutschen (insbesondere auf dem Gebiet der Lehnwörter)¹ oder aber untersuchten den slawischen Wortschatz im Rahmen ihrer etymologischen Wörterbücher wie etwa E. Berneker (1908–1914), A. Brückner (1926–1927) und F. Miklosich (1886). Seither kam es nicht nur zur Veröffentlichung zahlreicher moderner etymologischer Wörterbücher diverser slawischer Sprachen wie z. B. ERSJ, ÈSBM, ESJS, ESSJ, ÈSSJa, ESUM, Machek (1957, ²1968, ³1971), Rejzek (2001, ²2012, ³2015), Skok (1971–1974), Snoj (²2003) u. a. m., sondern auch zur Publikation wichtiger Einzelstudien zur Identifizierung von germanischen Lehnwörtern im Bereich des Slawischen.² Darüber hinaus sei auf die Überblicksdarstellungen der Forschungen zu den germanisch-slawischen Lehnbeziehungen von

¹ So z. B. die Studien von A. Brückner (1929), H. Hirt (1893), V. Kiparsky (1934), K. Knutsson (1929, 1938), V. Machek (1940), A. Matzenauer (1870), F. Miklosich (1867), S. Mladenov (1909), A. Stender-Petersen (1927), K. Štrekelj (1904), C. C. Uhlenbeck (1890, 1893), M. Vasmer (1932, 1934) u. a. m.

² So z. B. jener von G. Bellmann (1971), G. Holzer (1990, 2001, 2013), V. V. Martynov (1963, 1969) und G. Thomas (1997).

H. Birnbaum (1984) und G. Bellmann (2004) sowie die umfangreichen Literaturangaben in S. M. Newerkla (2004, ²2011) verwiesen. Letzterer bringt auch eine ausführliche und kritisch-analytische Übersicht über die bislang erschienen Studien zur Problematik der deutschen Lehnwörter im Tschechischen und Slowakischen, auf die wir hier aus Platzgründen lediglich Bezug nehmen können (Newerkla ²2011: 10–16). Neben einigen fehlerhaften Etymologien kommt es in diesen Studien bisweilen zu Fehleinschätzungen, insbesondere was das Alter mancher deutscher Lehnwörter betrifft. Mitunter berücksichtigen sie die dialektale Differenzierung des Deutschen zu wenig und erwecken so den Eindruck, es handle sich um Übernahmen aus dem schriftsprachlichen Deutschen, selbst wenn die lautliche Gestalt der entsprechenden tschechischen und slowakischen Lexeme dem zuwiderläuft. Nicht zuletzt stellt sich auch die Frage der deutschen Vermittlung von Lehnwörtern aus anderen Sprachen als problematisch dar: Erwägen ältere Publikationen diese teils in zu geringem Ausmaß, so wird sie heute bisweilen auch in einigen Fällen angenommen, wo genauso gut direkte Entlehnungen möglich sind.

2. Regularitäten des Lautersatzes und nachvollziehbare semantische Entwicklungslinien

Für die sorgfältige Analyse einer Wortschatzschicht ist es zunächst wichtig, die Erkenntnisse von lexikografischen Werken zu berücksichtigen, die den Wortschatz älterer Sprachepochen zum Gegenstand haben, da wir ohne sie in vielen Fällen kein Auslangen finden können. In Bezug auf das Tschechische geht es hier vor allem um den *Staročeský slovník* (Gebauer 1903–1916, StčS 1968 ff.), hinsichtlich des Slowakischen um den *Historický slovník slovenského jazyka* (HSSJ 1991 ff.) sowie die jeweils diesen Wörterbüchern zugrunde liegenden lexikalischen Archive. Die akribische lexikografische und semantische Arbeit der Mitarbeiter in der Abteilung für Sprachentwicklung des Instituts für tschechische Sprache der Tschechischen Akademie der Wissenschaften in Prag (Oddělení vývoje jazyka ÚJČ AV ČR) sowie der Abteilung für die Geschichte des Slowakischen des nach Ľudovít Štúr benannten Sprachwissenschaftlichen Instituts der Slowakischen Akademie der Wissenschaften in Bratislava (Oddelenie dejín slovenčiny Jazykovedného ústavu Ľudovíta Štúra SAV) ist dabei nicht zuletzt auch für die Aufdeckung neuer und plausibler Etymologien einiger Lehnwörter aus dem Deutschen von entscheidender Bedeutung.

Forschungen haben dabei deutlich bestätigt, dass lexikalische Entlehnungen zum überwiegenden Teil durch erkennbare Regularitäten des Lautersatzes und nachvollziehbare semantische Entwicklungslinien bezeugt werden können (vgl. u. a. Bellmann 2004: 3229 und Newerkla ²2011: 47–54). Für die früheste Zeit der Entlehnungen aus dem Urgermanischen bzw. Gotischen ist es bisweilen dennoch schwierig zu entscheiden, ob angesichts der oft erst in geringem Maße ausgebildeten Kontraste

zwischen beiden indogermanischen Sprach(grupp)en nicht überhaupt Urverwandtschaft vorliegt. Auch kann die Entlehnrichtung aufgrund der spärlichen Beleglage sehr häufig bis ins Mittelalter herauf umstritten sein.³ So war etwa noch bis vor einiger Zeit die Entlehnungsrichtung zwischen altschech. *búda*, altpoln. *buda* ‘Hütte, Bude’ im Kontrast mit mhd. *buode* strittig. Tschechische Etymologen⁴ nahmen hier bis zum Ende des 20. Jh. die Entlehnung des mittelhochdeutschen Ausdrucks aus dem Altschechischen bzw. Altpolnischen an. Die westslawischen Wörter sind jedoch vielmehr Entlehnungen aus mhd. *buode*,⁵ und das vermutlich aus mitteldeutschen Dialekten, in denen zu diesem Zeitpunkt die Monophthongierung von mittelhochdeutsch *-uo-* bereits eingetreten war.⁶ Diese Entlehnrichtung beweisen die Beleglage, die Verbreitung des Wortes im Germanischen und Slawischen, die ausschließlich westslawische Verbreitung der Verben tschech. *budovat*, slowak. *budovať*, poln. *budować* ‘bauen, errichten’, die Entstehung des altschechischen Verbs *budovati* (*sě*) ‘ein Lager aufschlagen, zelten’, bei dem es sich offenbar um ein Denominativum von altschech. *búda* handelt und nicht umgekehrt, sowie die weder lautlich noch semantisch nachvollziehbare Falschzuordnung der westslawischen Verben zu russ. *оборудовать* ‘ausrüsten, ausstatten, einrichten’, die Machek (1971 [1997]: 62) zur Stützung seiner These vom heimischen Ursprung der Verba vorgenommen hatte. Eine Beeinflussung durch tschech. *bouda* ‘Hütte, Bude’ ist einzig beim deutsch-schlesischen Ausdruck *Baude* ‘Bude, Bauernhof im Riesengebirge, Berggasthof’ wahrscheinlich. (Newerkla 2004: 159, ²2011: 154–155, 2014a: 66–68).

Die Ergebnisse unserer eingehenden Untersuchung der Rezeption von deutschen Lehnwörtern im Tschechischen und Slowakischen (Newerkla 2004, ²2011) weisen aber noch stärker in dieselbe Richtung, wie sie schon für die slawischen Lehnbeziehungen zum Germanischen bzw. Deutschen festgestellt wurde. So ist nicht nur die interne phonologische Ähnlichkeit von Ausgangsvokal und Zielvokal im Rahmen einer Entlehnung groß, sondern auch die Konsonanten von Lehnwörtern werden mit hoher Konstanz durch solche in der Zielsprache ersetzt, deren artikulatorische Merkmale den zu substituierenden Ausgangskonsonanten am nächsten sind. Mag eine solche Feststellung auf den ersten Blick trivial erscheinen, ist sie es angesichts der zunächst unübersichtlichen Situation aus diachroner Sicht aufgrund des Ineinandergreifens von innersprachlicher Lautentwicklung und tatsächlichen Übernahmemechanismen keineswegs. Denn sie unterstreicht, dass eine vorgeschlagene Etymologie umso

³ G. Holzer hat die regelmäßigen Substitutionen bei der Eindeutschung slawischer Lautungen in Österreich bis ins 12. Jh. schematisch dargestellt. Am Beispiel von slaw. *měša* ‘Messe’ und slaw. *oplatě* ‘Hostie, Oblate’, die über althochdeutsche Vermittlung auf lat. *missa* bzw. lat. *oblatum* zurückgehen (Newerkla ²2011: 132–133, 145), zeigt er, dass sich seine Schemata teilweise auch in umgekehrter Richtung verwenden lassen (Holzer 2001: 93).

⁴ Wie z. B. Holub/Kopečný (1952: 80), Holub/Lyer (⁴1992: 98), Machek (³1971 [1997]: 61f.) und SSČ (1994: 35). Falsch ist im Übrigen auch der Eintrag im Tschechischen Sprachatlas ČJA 2 (1997), 114 zu tschech. *budka*, das als Diminutivum zu urslaw. **buda* gedeutet wird.

⁵ Diese Entlehnrichtung führt von Seiten der tschechischen Etymologen zuletzt auch Rejzek (2001: 87, ³2015: 94) an.

⁶ Schon seit 1100 im Ostmitteldeutschen, aber nie im Oberdeutschen.

weniger wahrscheinlich ist, je größer die lautliche Diskrepanz zwischen dem angenommenem Quellwort der Ausgangssprache und dem Lehnwort in der Zielsprache zum Zeitpunkt der postulierten Übernahme ist. Das bereits von Holzer (1990: 62–63) eingeforderte wahrscheinlichkeitstheoretische Kriterium der Anzahl der vorhergesagten Phoneme in Verbindung mit dem Kriterium des Bedeutungsumfanges muss also stets berücksichtigt werden.

Unter Berücksichtigung dieses Kriteriums gelang es uns einige bislang vorgeschlagene Etymologien für manche Lexeme des Tschechischen und Slowakischen in Frage zu stellen und durch eine nachvollziehbare und plausible Erklärung der Herkunft zu ersetzen. Für tschech. *cimbuří*, slowak. *cimburie* ‘Zinne, Mauerzacke’ wurde etwa bis vor einigen Jahren noch unisono behauptet,⁷ es handle sich hierbei um eine Entlehnung und Metathese aus mhd. *burc-zinne* bzw. nhd. *Burgzinne*. Die alttschechischen Formen des Lexems und ihre semantischen Entwicklungsstränge widerlegen jedoch zweifelsfrei diese Auslegung: So liegt als tschechischer Erstbeleg für das Wort alttschech. *cibořie* (später *cimbořie*, *cimbuřie*) in der ursprünglichen Bedeutung ‘baldachinartige Krönung, Epistylum’ vor, das lautlich und semantisch mhd. *zibōrje* ‘Säulenhäuschen, baldachinartige Krönung, Hostiengefäß’ entspricht und damit auf lat. *cibōrium* ‘Hostiengefäß, Altarbaldachin’ und gr. *κιβώριον* ‘Fruchtgehäuse der ägyptischen Bohne’ zurückgeführt werden kann (Newerkla 2004: 164, ²2011: 159).

Ähnlich abwegig wurde lange Zeit die Herkunft von tschech. *prejt* ‘Füllung von Leber- und Blutwürsten, Brät’ und tschech. fachsprachlich *prát* ‘Brat, Wurstfülle, Wurstmasse’ erklärt. So gingen die Etymologen für tschech. *prejt* von einer Übernahme aus nhd. *Brei* unter Anfügung von *-t* wie im tschechischen gemeinsprachlichen Ausdruck *policajt* ‘Polizist’ aus.⁸ In tschech. *prát* sah Machek (³1971 [1997]: 483) überhaupt „zcela nové umělé přetvoření“ („eine gänzlich neue künstliche Umbildung“). In Anbetracht der mit den tschechischen Ausdrücken gleichbedeutenden Wörter *Brat* im österreichischen Deutsch und nhd. *Brät* (bzw. entsprechender bairisch-dialektaler Formen wie *prāt* und *prejt*) besteht auch in diesem Fall kein Zweifel an der tatsächlichen Etymologie dieser Lexeme. (Newerkla 2004: 334–335, ²2011: 326–327).

Die oberhalb genannten lautlichen Varianten machen jedoch zugleich deutlich, dass Übernahmen aus dem Deutschen oft parallel aus verschiedenen deutschen Nachbarmundarten beziehungsweise regional unterschiedlich gefärbten Varietäten erfolgten und sich erst im Laufe der weiteren Entwicklung der entlehnenden Sprache die eine oder andere Form des Lehnworts als überregionaler Standard durchsetzte. Die Vielfalt an möglichen lautlichen Varianten illustrieren dabei besonders anschaulich die

⁷ Vgl. Eisner (1946 [1992]: 394, 1948 [1996]: 39), Janko (1916: 411), Holub/Kopečný (1952: 85), Holub/Lyer (⁴1992: 107), Machek (³1971 [1997]: 86), Rejzek (2001: 106), Rudolf (1991: 189–192) und Schneeweis (1912: 24). In der neuesten Auflage hat dies Jiří Rejzek bereits richtig gestellt (Rejzek ³2015: 114).

⁸ Vgl. Eisner (1946 [1992]: 400, 1948 [1996]: 215), Holub/Lyer (⁴1992: 361, 370), Machek (³1971 [1997]: 483) und Rejzek (2001: 499). In der neuesten Auflage hat dies Jiří Rejzek bereits richtig gestellt (Rejzek ³2015: 550).

rund zweihundert alttschechischen Belege für tschech. *hejtman* ‘Hauptmann’, das uns im Alttschechischen in den Varianten *hajp(t)-/hajt-/hajtp-/haj-/hant-/hat-/haup-/haupt-/houpt-/haut-/hejt-/hét-/heupt-/hút-/hýtman* begegnet. (Newerkla 2004: 272–273, ²2011: 265).

3. Phänomen der Sensibilität für Wörter fremder Herkunft

Nicht einmal die scheinbar so zuverlässigen Grenzziehungen eines *Terminus a quo* und *Terminus ante quem* sind für Lehnwörter in jedem Fall eindeutige Beurteilungskriterien. So ist für die deutschen Lehnwörter im Tschechischen und Slowakischen ein Phänomen in Betracht zu ziehen, das André de Vincenz (1986) schon vor bald drei Jahrzehnten für die Entwicklung der deutschen Lehnwörter im Polnischen beschrieb, nämlich das Phänomen der Sensibilität für Wörter fremder Herkunft.⁹ Im Tschechischen und Slowakischen lassen sich einige Entlehnungen aus dem Deutschen identifizieren, die nach ihrer Übernahme weiterhin mit der Gebersprache in Kontakt blieben, das heißt, dass sich Sprecher des Tschechischen und Slowakischen der deutschen Herkunft dieser Wörter noch Jahrzehnte oder gar Jahrhunderte später bewusst waren oder aber erneut wurden. Dies führte mitunter dazu, dass die bestehende phonologische Form eines deutschen Lehnworts im Tschechischen bzw. Slowakischen nicht der sprachlichen Entwicklung der Nehmersprache folgte, sondern mit der Zeit durch eine andere ersetzt wurde, die der späteren phonologischen Gestalt des deutschen Quellworts und damit dem Ausdruck in der Gebersprache näher stand. Es gibt also Lehnwörter, die nach ihrer Entlehnung zwar phonemisch, morphemisch usw. in das System des tschechischen beziehungsweise slowakischen Wortschatzes integriert wurden, andererseits jedoch mit dem deutschen Ausgangswort in Kontakt blieben, was sich in einer nachträglichen Angleichung an dessen phonemische Fortsetzung niederschlug. Ein offenkundiges Beispiel ist slowak. *farba* analog zu nhd. *Farbe* im Gegensatz zu älterem slowak. bzw. dialektal *barva* (wie tschech. *barva*) aus mhd. *varwe* (Newerkla 2004: 156–157, ²2011: 151–152).¹⁰ Angesichts diffuser Beleglagen ist es jedoch nicht immer einfach zu entscheiden, ob es sich in einem bestimmten Fall um ein Beispiel für das soeben beschriebene Phänomen handelt oder ob lediglich eine Neuentlehnung vorliegt.

⁹ Dieses Phänomen wurde übrigens schon Anfang des 20. Jh. von Otto Jespersen (⁹1938 [1952]: 104–105) für Lehnwörter im Englischen festgestellt.

¹⁰ Weitere Beispiele sind u. a. tschech. und slowak. *barchet* ‘Barchent’, bei Jungmann (1835, I: 71) tschech. *barchan*, *barkan*, alttschech. *barchan*, historisch slowak. *barchan*, *borchan*, *barhan*, *borhan* (Newerkla 2004: 156, ²2011: 151); tschech. und histor. slowak. *galgant* ‘Galgant(wurzel); *Alpinia officinarum*’ neben tschech. *galgán*, slowak. *galgan*, bei Jungmann (1835, I: 638) tschech. *galgan*, *galgán*, *kalkán*, alttschech. *galban*, *galgán*, *galgan*, *kalkan*, *kalgan*, historisch slowak. *galgan* (Newerkla 2004: 175–176, ²2011: 170); tschech. gemeinsprachlich *štrozok*, *štrúzok*, slowak. alltagssprachlich *štrozak* ‘Strohsack’, aber alttschech. *strožak*, *strožek*, histor. slowak. *strožiak*, *strožok*, *strožliak* (Newerkla 2004: 232, ²2011: 225).

Das Phänomen selbst ist jedoch bis in die Gegenwart zu beobachten. So orientieren sich etwa HSSJ und SSN bei der Festlegung der Stichwörter für Übernahmen aus dem Deutschen offenbar auch an den jeweils standardsprachlichen Formen der deutschen Quellwörter, so etwa in HSSJ (1991, I: 124) histor. slowak. *bergfervalter* ‘Verwalter eines Bergwerks’ (belegt ist jedoch lediglich *pergfrbaltr*), *bericht* ‘Bericht’ (belegt sind jedoch *pericht*, *periht*), *bergman* ‘Bergmann’ (im 17. Jh. lediglich histor. slowak. *pergmon*, *perkmon*, *pergman*; die Form *bergman* ist erst seit der zweiten Hälfte des 18. Jh. belegt) oder in SSN (1994, I: 411) slowak. dialektal *durchšlág* (neben *durchšlag*, *durchšlóg*, *duršlag*, *dušlág*, *dršliak*, *dršlák* u. a. m.

4. Problematik des Purismus im Rahmen der Kodifikation

Während alle bisher genannten Problembereiche nur insofern Auswirkungen auf das Untersuchungsdesign des von Martin Haspelmath und Uri Tadmor initiierten Projekts *Loanwords in the World's Languages* (Haspelmath/Tadmor 2009) haben, als sich die Qualität der etymologischen Arbeit unmittelbar auf die Qualität des erhobenen Datenmaterials niederschlägt, so bleiben weitere Problemfelder überhaupt außen vor, was zu einem falschen oder nur bedingt gültigen Bild der Anzahl und Verwendung von Lehnwörtern führen kann. Gerade in Sprachen wie Tschechisch und Slowakisch, wo es über Jahrhunderte zu massiven Umgestaltungen des Wortschatzes aus puristischen Gründen kam, ist es oft nicht ganz einfach zu bestimmen, ob ein Wort nun überhaupt ein Lehnwort ist oder nicht. Während das Slowakische von den Kodifikatoren seit seiner relativ späten Kodifikation bis dato vor vermeintlichen und tatsächlichen Bohemismen bewahrt bzw. von diesen befreit wird, verwendet die Bevölkerung selbst diese sogenannten Bohemismen in der alltäglichen Kommunikation flächendeckend.¹¹ Im Tschechischen ist wiederum die Abgrenzung zum Deutschen hin eine „Konstante der tschechischen Sprachdiskussionen“ (Berger 2009: 144) seit den bekannten mahnenden Worten des tschechischen Reformators Jan Hus zu Beginn des 15. Jh. in seinem *Výklad viery, desatera božieho prikázanie a modlitby Páně*, also der Auslegung des Glaubensbekenntnisses, des Dekalogs und des Gebets des Herrn. Darin rügt Hus in Anlehnung an das Verbot von Mischehen bei Nehemia 13,23–27 seine Landsleute wegen ihrer mit zahlreichen deutschen Lehnwörtern gespickten Sprache mit folgenden Worten (zitiert nach HusVýklB 58a in Newerkla²2011: 9):

Také majú sě postaviti [knížata, páni, rytíři, vladykové, měšťané], aby česká řeč nehynula [...]: též nynie hodni by byli mrskanie Pražané i jiní Čechové, jenž mluvie

¹¹ Selbst über den engeren Kreis der slowakistischen Fachwelt hinaus sind Beispiele wie breit verwendetes *prádlo* statt des kodifizierten Ausdrucks *bielizeň* ‘Wäsche’ bzw. durchwegs im Usus gebräuchliches *nula celá* anstelle der kodifizierten Verbindung *nula celých* ‘null Komma fünf’ gut bekannt. Vgl. dazu weiter führend auch Newerkla (2009, 2011).

odpoly česky a odpoly německy, říekajíc: [...] hantuch za ubrusec, šorc za zástěrku, knedlík za šišku, renlík za trérožku, pancieř za kruněř, hunškop za koňský náhlavek, marštale za konnici, mazhaus za vrchní sieň, trepky za chody, mantlík za plášťiek, hausknecht [ms. hawz[knecht] za domovní pacholek, forman za vozataj [...]. // („Sie [Fürsten, adlige Herren, Ritter, Wladikas, Bürger] sollen auch aufbegehren, auf dass die tschechische Sprache nicht aussterbe [...]: so hätten nun auch jene Einwohner Prags und andere Böhmen die Peitsche verdient, die halb tschechisch und halb deutsch reden, indem sie sagen: [...] hantuch statt ubrusec, šorc statt zástěrka, knedlík statt šiška, renlík statt trérožka, pancieř statt kruněř, hunškop statt koňský náhlavek, marštale statt konnicě, mazhaus statt vrchní sieň, trepky statt chody, mantlík statt plášťiek, hausknecht statt domovní pacholek, forman statt vozataj [...].“ (Newerklá 2011: 9).

Jason Pontius geht sogar soweit, den bis in die Gegenwart nicht überwundenen Abstand zwischen der kodifizierten Standardsprache und dem tatsächlich verwendeten gemeinsprachlichen Usus insbesondere in Böhmen direkt dieser puristischen Grundhaltung anzulasten: „The connecting link between Czech pseudo-glossia and anti-Germanism is the long tradition in the Czech lands of linguistic purism in the literary language, which has contributed directly to the stratification in the modern language.“ (Pontius 1997: 102). Dabei merkt er auch hier an, dass diese sprachreinigenden Bemühungen von der breiten Bevölkerung nicht im selben Maße mitgetragen wurden, was zur Vergrößerung des Abstands beitrug: „[...] in fact Czech linguistic stratification results from the different attitudes toward German held on the one hand by the codifiers of Czech; and on the other by the average Czech, less concerned with the maintenance of a pure national language than with the needs of everyday communication.“ (Pontius 1997: 102). Daraus zieht er den Schluss: „It is possible to argue that the linguistic stratification which characterizes modern Czech has its roots in the codification of the modern language, and in a pervasive anti-German bias which reaches even to the level of segmental phonology.“ (Pontius 1997: 107).

In besonders großem Stil wurden im Tschechischen vom Beginn des 19. Jh. bis in die erste Hälfte des 20. Jh. Lehnwörter aus dem Deutschen durch – oft nach deutschem Muster calquierte – Neubildungen, wiederbelebte Archaismen und lautlich angepasste Entlehnungen aus anderen slawischen Sprachen ersetzt. Die Auswirkungen eines solchen Purismus auf den Wortschatz sind zweifellos auch von theoretischem Interesse, da – wie der Tübinger Slawist Tilman Berger treffend bemerkt – „letztlich eine direkte Beeinflussung des Tschechischen durch das Deutsche durch eine Art von Beeinflussung abgelöst wurde, die weniger sichtbar ist, aber genauso intensiv verlaufen ist.“ (Berger 2009: 147), wobei dieser für bestimmte Sprachsituationen nicht untypische Fall weder im Untersuchungsdesign von Haspelmath und Tadmor (2009) noch im Modell von Thomason und Kaufman (1988) entsprechend berücksichtigt wird. Hinzu kommt, dass viele dieser ersetzten Lehnwörter aus dem Deutschen gar nicht aus dem Tschechischen verschwanden, sondern lediglich in unterschiedlichem Maße verdrängt oder auf bestimmte Sprachschichten beschränkt wurden, an sich aber weiterhin gebräuchlich sind.

Vor dem Hintergrund ihrer möglichen Integration oder aber ihrer Verdrängung bzw. ihrem nachfolgenden Ersatz hat schon Günter Bellmann (1971: 28–31) vorgeschlagen, Lehnwörter grob in zwei Gruppen zu gliedern, und zwar in jene lexikalischen Einheiten, die eine grundlegende Kode-Erweiterung bedeuten,¹² und jene, die unmittelbar nach der Übernahme als (annähernd) synonyme lexikalische Dubletten zu bereits bestehenden Bezeichnungen Verwendung finden.¹³ Berger (2009: 148) unterscheidet mit Bezug auf das Tschechische zumindest zwischen drei Fällen, nämlich solchen, wo das Lehnwort aus dem Deutschen in gesprochener Sprache weiterhin merkmallos bleibt und der puristische Ersatz im Wesentlichen auf schriftliche Verwendungen beschränkt bleibt (wie etwa bei *strúdl* im Vergleich zu *závin* ‘Strudel’), solchen, in denen das deutsche Lehnwort weitgehend verdrängt und gewissermaßen nur noch als „Archaismus“ belegbar ist (wie etwa bei *firhánky* im Vergleich zu *záclona* ‘Vorhang’), und solchen, in denen die Lehnwörter eine zusätzliche „expressive“ Konnotation annehmen, während in den merkmallosen Verwendungen die neu gebildeten Ausdrücke vorherrschen (wie etwa bei expressivem *frajer* ‘Liebhaber, Geck’ im Vergleich zu neutralem *nápadník* ‘id.’).

Wie komplex so eine semantische Diversifizierung ablaufen kann und wie dabei mehrere beschriebene Phänomene (wie zum Beispiel auch das Phänomen der Sensibilität für Wörter fremder Herkunft) ineinandergreifen, zeigt das Beispiel von tschech. und slowak. *šnek*. Altschech. *šnek* bedeutete ursprünglich ‘Gefäß in Schneckenform, Wendeltreppe’ (gebucht seit dem 14. Jh. u. a. bei Comestor). Als ‘Wendeltreppe’ ist auch älteres slowak. *šnek* erstmals 1563 in Hlohovec (Freistadt, Galgóc) belegt. Diese Formen gehen auf frühnhd. oberdt. *schneck(e)* < mhd. *snēcke* männlichen Geschlechts und mit der Bedeutung ‘gewundenes, spiralförmiges Gebilde; Schnecke, Wendeltreppe’ zurück. Im Laufe der semantischen Weiterentwicklung kam es nun bei tschech. *šnek* einerseits zum Verlust der Bedeutung ‘Wendeltreppe’ und zu einer Beschränkung auf die Bedeutung ‘Spirale, Schneckenwelle’ in der Fachterminologie der Maschinenbauer, andererseits wurde im 17. Jh. unter deutschem Einfluss die Bedeutung ‘Schnecke mit Schneckenhaus, Weinbergschnecke’ aktiviert, die bis heute im alltagssprachlichen Stil des Standardtschechischen Verwendung findet. Diese Form steht seither in Konkurrenz zum unmarkierten tschech. Standardausdruck *hlemýžď*, der im Altschechischen nicht nur eine ‘Schnecke mit Schneckenhaus’, sondern auch eine ‘Schildkröte’ (mitteltschech. *šilkruta*, heute

¹² Zum ersten Typ zählen also in erster Linie Lehnwörter, die ihre Übernahme einer unmittelbar notwendig gewordenen Anpassung an die kommunikativen Erfordernisse der außersprachlichen Realität im Rahmen des Sprach- und Kulturkontakts verdanken, also neue Bezeichnungen für bis dahin in einem bestimmten Milieu unbekannt Gegenstände, Güter, Organisations- und Rechtsformen, Haltungen, Einstellungen, Erkenntnisse in Wissenschaft, Technik usw.

¹³ Der zweite Typ trifft auf jene Gruppe von Entlehnungen zu, für die zum Zeitpunkt der Übernahme bereits eigene Entsprechungen in den jeweiligen Sprachen existieren. Vor einer weiteren Differenzierung kommt es also für gewisse Zeit zu einer annähernden Synonymie zwischen den bestehenden und den in diese Sprache neu transferierten lexikalischen Einheiten, also zu einer Konkurrenzsituation zwischen althergebrachten und neu übernommenen Bezeichnungen, die in eine Substitution des Altworts, negative Integration der Neuentlehnung oder semantische Diversifizierung des Altworts und/oder der Neuentlehnung mündet.

tschech. *želva*) bezeichnen konnte. Im Slowakischen war die Entwicklung zunächst ähnlich, doch wurde dort das Wort *šnek* stärker an den Rand des Systems gedrängt. Es verlor frühzeitig die Bedeutung ‘Wendeltreppe’, auch konnte sich die im 17. bzw. 18. Jh. angenommene Bedeutung ‘Schnecke mit Schneckenhaus’ nie gegen indigenes slowak. *slimák*¹⁴ behaupten. So ist slowak. *šnek* heute nur im Substandard als Ausdruck des technischen Fachjargons in der Bedeutung ‘Spirale, Schneckenwelle’ verankert. (Newerkla 2004: 306–307, ²2011: 298–299).

Besonders schwierig wird eine Klassifikation bzw. Einordnung von eigentlichen Lehnwörtern dann, wenn am Ende der Entwicklung ein Lexem steht, bei dem weder die Wortwurzel noch die Form der Wortbildung an eine Entlehnung erinnern. Dies ist beispielsweise beim tschechischen und slowakischen Ausdruck *sporák* ‘Herd’ der Fall, denn die Wortwurzel **spor-* ‘ausgiebig etc.’ ist allslawisch, die Derivation mittels *-ák* typisch tschechisch und slowakisch. Dennoch handelt es sich hier lediglich um eine nachträgliche Slawisierung von tschech. gemeinsprachlich *šporhert*, *špolhert*, slowak. dialektal *šparhert*, *šporhert*, *šporhelt*, eines ursprünglichen Lehnworts aus dem Bairischen *šporhert* bzw. Austriazismus *Sparherd*, der sich in dialektalen und volkssprachlichen Varianten auch im Polnisch-Schlesischen des Teschener Gebiets (*szperok*),¹⁵ Slowenischen (*šporget*), Kroatischen und Serbischen (*šporet*, *šparet*, *šporhert*) sowie Ungarischen (*sparhert*, *spór*) belegen lässt. (Newerkla 2004: 340, 341–342, ²2011: 332, 333).

Ein anderes Beispiel ist das tschechische Wort *hřbitov* für Friedhof, bei dem zunächst ebenfalls nichts mehr an eine Entlehnung erinnert. Bei näherer Analyse ergibt sich jedoch, dass es im Altschechischen nicht nur diese, sondern daneben noch die ältere Form *břitov* gab, die als solche direkt auf mhd. *vrīthof* ‘eingefriedeter Raum um eine Kirche, Friedhof’ zurückgeht und erst unter Anlehnung an tschech. *(po)hřbít* ‘begraben, beerdigen’ zu altschech. *hřbitov* wurde. (Newerkla 2004: 184, ²2011: 178).

5. Problematik der Beleglage (Bedeutung der Erstbelege)

Wie auch das soeben erwähnte Beispiel verdeutlicht, sind jedenfalls das Quellenstudium und die Auswertung der ältesten vorliegenden Texte von unmittelbarer Bedeutung zur Bestimmung von Lehnwörtern. Es ist zwar nicht jedes Lehnwort gleich nach seiner Übernahme schriftlich fixiert worden, somit kann das Fehlen eines frühen Belegs allein auch nicht als Beweis für die Übernahme in späterer Zeit ins Treffen geführt werden. Ein Hinweis ist es jedoch allemal. Vor allem legen umgekehrt gerade frühe Belege sehr wohl eine frühe Übernahme nahe, und viele davon sind lange übersehen worden. Im

¹⁴ Interessanterweise bedeutet tschech. *slimák* ‘Nacktschnecke’. Die Nacktschnecke heißt auf Slowakisch *slizniak*.

¹⁵ Bezüglich der sprachlichen Besonderheiten dieses Gebiets siehe u. a. Menzel/Hentschel (2003) und Černá-Willi (2012).

Rahmen unserer Arbeiten hat sich etwa gezeigt, dass oft sehr alte Belege für Lehnwörter existieren, die von tschechischen Etymologen bislang nach lautlichen und kulturhistorischen Gesichtspunkten allein für jüngere Entlehnungen aus dem Neuhochdeutschen gehalten wurden, so z. B. tschech. und slowak. *baldrián* ‘Baldrian, Valeriana officinalis’, für das rund 25 alttschechische Belege (als *paldrián*, *paldriám*, *paldrán*, *valdrián*, *baldrián*) vorliegen (Newerkla 2004: 156, ²2011: 151).¹⁶

Doch auch in semantischer Hinsicht helfen die Belege aus ältester und alter Zeit bei der Klärung der tatsächlichen Übernahmeprozesse. Während tschechische Etymologen¹⁷ für tschech. *cimbál*, slowak. *cimbal* ‘Zimbal, Zimbel, Glockenspiel’ bis dato oft eine direkte Entlehnung aus lat. *cymbalum* bzw. gr. *κύμβαλον* ‘Zimbel, Metallbecken’ annehmen, hat die Brünner Etymologin Helena Karlíková (1999) schließlich nachgewiesen, dass die alttschech. Wortformen *cymbala*, *cymbal* ‘Zimbal, Zimbel’ aus semantischen Gründen auf ahd. *cymbala* ‘Tongerät mit Glöckchen, Glockenspiel’ aus mittellat. *cymbala* ‘Glocke’ zurückgehen müssen. (Newerkla 2004: 127–128, ²2011: 123, 2014a: 70). Auch uns gelang es mittels semantischer Vergleiche die deutsche Vermittlung für tschech. *pukla*, *pukle*, histor. slowak. *pukla* ‘Metallbeschlag, erhabene Metallverzierung (am Schild, Buch, Pferdegeschirr)’ sowie mährisch-schlesisch dialektal *pukel* ‘Buckel auf dem Rücken, Beule’ nachzuweisen und angesichts der Übereinstimmung der alttschech., hist. slowak. und mhd. Belege sowie der verwandten Formen wie mittellat. *bokel* ‘Metallbeschlag am Schild, Spange, Schnalle’, engl. *buckle* ‘Schnalle, Spange’, mittellat. *buccula* ‘Schildbuckel, eigtl. Backe, Backenstück am Helm’, altfranz. *boucle* ‘halb- oder erhabener Metallbeschlag in der Mitte des Schildes’ und franz. *boucle* ‘Schnalle, Spange, Schlaufe, Öse’ den Ausdruck als Wanderwort zu identifizieren. (Newerkla 2004: 212, ²2011: 205, 2014a: 71).

6. Problematik von areallinguistischen und sprachgeographischen Einflüssen

Ein Phänomen, das nur in einem vermittelten Zusammenhang mit Entlehnungen aus dem Deutschen steht, aber hier dennoch Erwähnung finden sollte, sind sprachliche Konvergenzprozesse, also die Teilhabe an arealen Gemeinsamkeiten in Sprache und Kultur, wie sie sich durch die Geschichte an vielen Orten nicht nur in Europa entwickelt haben. Natürlich geht es dabei unter anderem auch um gemeinsam bewahrte Entlehnungen in den Sprachen der ehemaligen Habsburgermonarchie, die zugleich das Deutsche in Österreich von jenem etwa in Deutschland unterscheiden (z. B. bei Ausdrücken wie *Matura* im Gegensatz zu *Abitur*).¹⁸ Im Einklang mit den wissenschaftlichen Erkenntnissen der Mehrsprachigkeitsforschung etwa von Kurt Braunmüller (Braunmüller/House 2009) oder Pieter Muysken (2000) bezüglich der Herstellung

¹⁶ Für weitere Beispiele siehe Newerkla (2010: 37–38).

¹⁷ So Holub/Kopečný (1952: 85), Holub/Lyer (⁴1992: 107), Machek (³1971 [1997]: 92), Rejzek (2001: 106, ³2015: 114) und SSČ (1994: 42).

¹⁸ Ausführlicher dazu Newerkla (2004: 79–89, ²2011: 76–86; 2007a, 2007b, 2014b).

einer weitgehenden lexikalischen Parallelführung der beteiligten Kodes in langwährenden Kontaktsituationen lassen sich für das altösterreichische Kontaktareal solche konvergenten Prozesse aber auch auf anderen Sprachebenen festmachen.¹⁹

Doch zurück zur Frage der lexikalischen Entlehnungen: Überraschenderweise hatte George Thomas von der Universität Hamilton bei der kontrastiven Untersuchung der deutschen Lehnwörter im Tschechischen, Slowakischen, Slowenischen und Kroatischen der Habsburgermonarchie festgestellt, dass nach der statistischen Auswertung seines Korpus der Einfluss des Deutschen gesamt gesehen auf das slowenische Lexikon am größten gewesen sei. Die Zahl der deutschen Lehnwörter in den nicht standardsprachlichen Varietäten wäre jedoch im Tschechischen und jene der stilistisch unmarkierten Lehnwörter in der slowakischen Standardsprache verhältnismäßig am höchsten (Thomas 1997: 343–344). Besonders interessant ist in diesem Zusammenhang seine Beispielsliste von allen vier Sprachen gemeinsamen deutschen Lehnwörtern, wobei zugleich die Übereinstimmung ihrer slawischen Äquivalente in allen vier Sprachen nur in rund 16 % der Fälle gegeben ist (Thomas 1997: 341–349). Dieser Umstand unterstreicht selbstredend die Rolle des Deutschen als *Lingua franca* des Habsburgerreiches.

All diese Erkenntnisse sind darüber hinaus eine weitere Bestätigung für die Erkenntnis der Soziolinguistik, dass verschiedene Sprachen unter politischem und sozialem Einfluss eine Sprechergemeinschaft ausbilden. Symptomatisch ist dafür übrigens ebenso der lange Zeit wenig beachtete Umstand der spezifischen linguistischen Klassifikation der Gebärdensprachen dieses Gebiets. So gibt es etwa seit 1780 die Familie der Österreichisch-ungarischen Gebärdensprachen, die letztlich auf die *Langue des signes française* (1752), also die Französische Gebärdensprache zurückzuführen sind. Zu diesen Österreichisch-ungarischen Gebärdensprachen zählen dabei die Österreichische, die Tschechische, die Slowakische und die Ungarische Gebärdensprache (Wittmann 1991: 282), aus der 1806 auch die Russische und daraus wieder die Bulgarische Gebärdensprache hervorgingen (Wittmann 1991: 285), sowie die Jugoslawischen Gebärdensprachen (ab 1840), also die Serbische, Kosovarische, Kroatische, Makedonische und Slowenische Gebärdensprache (Wittmann 1991: 286). Die Deutsche Gebärdensprache wird hingegen nur der kleinen Gebärdensprachenfamilie der deutschen Gebärdensprachen zugerechnet,

¹⁹ In diesem Zusammenhang ist anzumerken, dass in der 53. Kuratoriumssitzung des österreichischen Fonds zur Förderung der wissenschaftlichen Forschung (FWF) am 4. Mai 2015 der Spezialforschungsbereich (SFB) F 60-G23 „German in Austria (DiÖ). Variation – Contact – Perception“ nach langer Vorbereitungszeit (Ausarbeitung des Vorantrags seit 13. April 2013; Einreichung des Vorantrags zum 29. August 2013; Bewilligung zur Ausarbeitung des Vollartrags in der 47. Sitzung des Kuratoriums des FWF am 3. März 2014; Ausarbeitung des Vollartrags zum 30. September 2014; Review Hearing am 23. Februar 2015) bewilligt wurde. Als Sprecherin und Sprecher des SFB, dessen offizieller Projektstart mit 1. Jänner 2016 festgelegt wurde, treten Alexandra N. Lenz vom Institut für Germanistik der Universität Wien und Stephan Elspaß vom Institut für Germanistik der Universität Salzburg in Erscheinung. Unsereins fungiert im Rahmen des SFB als Projektleiter für das Teilprojekt 05 „German in the context of the other languages of the Habsburg State“ und das Teilprojekt 06 „German and Slavic Languages in Austria: Aspects of Language Contact“.

zu der auch die Polnische Gebärdensprache und eventuell auch die Israelische Gebärdensprache gehören, wobei letztere aber Anteile von Österreichisch-ungarischen Gebärdensprachen aufweist (Wittmann 1991: 282, 283).²⁰

Jedenfalls kam es infolge der engen Beziehungen zwischen den Sprachen auf dem Gebiet der Habsburgermonarchie auch zu einer allgemeinen Annäherung der Wortinhalte. Der Prager Germanist Emil Skála (1992: 176, 2000: 80) verwies in diesem Zusammenhang gerne auf rund 150 Verben der Fortbewegung, die im Deutschen und Tschechischen weitgehend in den Wortinhalten übereinstimmen. Das mitteleuropäische Areal der Sprachen der ehemaligen Habsburgermonarchie²¹ ist jedoch nur eines von mehreren Spracharealen, an dem das Deutsche beteiligt ist bzw. war. So stellt die Wiener Finnougristin Johanna Laakso (2004: 171–177) in Bezug auf das Baltikum für die Diminutivbildung im Estnischen auf *-ke(ne)* und im Niederdeutschen auf *-ken* überzeugend areale Bezüge her.

Wie in der dialektologischen Forschung Isoglossenbündel bestimmte Dialektgebiete hervortreten lassen, so heben gebündelte Sprachkontakterscheinungen bestimmte Kontaktareale voneinander ab. In manchen Fällen kommt es dabei sogar zur Deckung von kleinräumigen Dialekt- und Kontaktarealen, etwa wenn bei Dublettformen die Wortgeographie der Lehnwörter in den Kontakt-dialekten oftmals eine Art Ausdehnung der in den gegenüberstehenden Kontaktvarietäten bestehenden Verhältnisse ist. Rudolf Šrámek (1998: 302–303) macht diesbezüglich auf das aufschlussreiche Beispiel der Genusgrenze von deutsch *die Butter* gegenüber bairisch dialektal *der Butter* und die überraschende Fortsetzung dieser Isoglosse auf tschechischer Seite als Kontaktentlehnungen *putra* gegenüber *putr* aufmerksam.²²

7. Problematik der versteckten Mehrsprachigkeit

Die Entlehnung von Lehnwörtern wird häufig unter dem Gesichtspunkt des Sprachkontakts von Bilingualen bzw. von Sprechergruppen zweier Sprachen betrachtet, ohne dabei zu bedenken, dass wir es in der Regel mit Situationen der offenen oder gar verdeckten Mehrsprachigkeit zu tun haben.²³ Wirke etwa in früheren Jahrhunderten Latein als Sakral- und Bildungssprache sowohl auf das Deutsche als auch das Tschechische und Slowakische bis zu einem gewissen Grad normbildend, so ist heute die Vorbildrolle des Englischen offenkundig.

²⁰ Die Situation in der Schweiz ist noch komplexer, die Deutschschweizer Gebärdensprache (DSGS) zerfällt in weitere Dialekte, gilt auch in Liechtenstein und unterlag zuletzt einer dynamischen Entwicklung (Boyes Braem/Haug/Shores 2012).

²¹ Vgl. dazu Kurzová (1996) und Thomas (2008).

²² Vgl. auch Kloferová (1996: 49–56).

²³ Jeder Mensch ist mehrsprachig. Vgl. dazu auch Busch (2013).

Doch auch andere Sprachen entfalteten ihren Einfluss auf dem Gebiet der Habsburgermonarchie, wobei deren Spuren ungleich schwieriger aufzudecken sind, etwa die Rolle des Romani bzw. Romanes²⁴ und insbesondere die Problematik des Knaanischen und später Jiddischen. Mit Knaanisch, auf Tschechisch *lešon Kena'an*, ist jene von Juden verwendete westslawische Sprache gemeint, die bis zum ausgehenden Mittelalter vor allem auf dem Gebiet der böhmischen Länder gesprochen wurde. Inwieweit diese Sprache überhaupt vom Altschechischen verschieden war oder mit den Worten von Lenka Uličná einen eigenen religiösen Interdialekt bzw. Religiolekt darstellte, ist Gegenstand der Diskussionen der Forscher (Bláha/Dittmann/Uličná 2013, Uličná 2011). Jedenfalls liegen frühe hebräische Texte mit Glossen in Altschechisch, Altfranzösisch und Jiddisch vor, so etwa Abraham ben Azriels *Arugat ha-Bosem* (= Beet der Balsampflanzen, beendet ca. 1234) oder Jicchak ben Mošes *Or Zarua* (= Das gesäte Licht, beendet ca. 1246). Allmählich wurde Knaanisch gänzlich von Jiddisch verdrängt, nachdem ab dem 12. bis ins 14. Jh. Jiddisch sprechende Juden aus dem Rheingebiet nach Böhmen und Mähren wanderten.²⁵ Einige Ausdrücke haben aber darüber hinaus Bestand gehabt, was gemeinhin zu wenig berücksichtigt wurde.

Ein anschauliches Beispiel dafür stellt tschech. *hampejz* 'Freudenhaus, Spelunke' dar, das darüber hinaus dialektal im übertragenen Sinn 'Kegelbahn' bedeuten konnte, da es in den Freudenhäusern zur Unterhaltung auch derartige Spielmöglichkeiten gab, bzw. im Mitteltschechischen bei Václav Jan Rosa auch eine 'Hundehütte' bezeichnete. Für das Slowakische nennen den Ausdruck lediglich Anton Bernolák und Josef Jungmann. Die altschech. Belege lauten *hanbejs* (erstmalig 1374), *hanpejs* (1380), *hanpajs* (1381), *hampajs* (1384), *hampejs* (1419) und *hampýs* (1439). Angesichts dieser Beleglage steht für uns aus lautlicher Sicht mit Vasmer (1933: 401) und Machek (³1971 [1997]: 159) eine Übernahme aus mhd. *han(en)bais*, *han(en)beis* 'Freudenhaus' außer Zweifel.²⁶ Auf den Freudenhäusern war oft das Bild eines Hahnes auf einer Henne angebracht (vgl. mittellat. *gallimordium*). Dem zweiten Wortteil liegt dabei jiddisch *bajis* 'Haus' < hebräisch *בַּיִת* *bajit*, *bēt* 'id.' zugrunde. Diese These stützen zusätzlich die

²⁴ Für die Dialekte der Roma und Sinti interessierte sich bereits intensiv Franz Miklosich (z. B. 1872–1880).

²⁵ Der äußerst gelungene Sammelband von Bláha/Dittmann/Uličná (2013) enthält in diesem Zusammenhang zahlreiche wichtige und erhellende Aufsätze zur Sprachstruktur des Knaanischen und dem betreffenden historischen Hintergrund. Im Kontext mit jüdischen Sprechergruppen sei auch auf das als zweiten Band der Reihe *Sprach- und Kulturkontakte in Europas Mitte. Studien zur Slawistik und Germanistik* erschienene Buch von Kačny/Olszewska/Twardowska (2013) hingewiesen.

²⁶ Holub/Lyer (⁴1992: 175) nennen zwar ebenfalls die richtige Quelle, doch übersetzen sie das mhd. Wort fälschlicherweise als 'Hahnenbiss'. Dies kommt daher, dass früher die Etymologen das tschech. und slowak. Lexem falsch aus mhd. *hanen-peiz* 'Hahnenbiss' herleiteten (z. B. Berneker 1908–1913, I: 376, Janko 1917: 122–123, Mayer 1927: 10, 24, 30, 41, 53, 58, 63, Rudolf 1991: 48, Schneeweis 1912: 17, 24, 29, 39 u. a. m.). Auch lautlich auszuschließen ist angesichts der Beleglage eine Ableitung aus mhd. *han(en)-biz* 'Hahnenbiss', wie sie etwa Rejzek (2001: 194) vorschlug, sie aber neuerdings ebenfalls als weniger wahrscheinlich einschätzt (Rejzek ³2015: 213). Da ließe sich eher – obwohl ebenso unrichtig – an eine Verquickung mit franz. *baiser* 'küssen (Hand, Gesicht)', vulgär auch 'ficken' denken. Eisner (1948 [1996]: 29, 86) geht überhaupt von einer zu jungen Entlehnung < dt. *Hahnpass* aus, das auch semantisch unklar ist.

ehemalige *Hampejzská ulice*, ein sehr alter Straßename der Prager Judenstadt, sowie die mittelschech. Formen *hamous*, *hamouz* ‘Hurenhäusler (als Schimpfwort)’, hannakisch *hamóznik* ‘id.’, die auf alttschech. *hamús* < mhd. *han* ‘Hahn’ + mhd. *hūs* ‘Haus’ zurückgehen, aber auch die bis heute bestehenden vulgären bzw. pejorativen Ausdrücke tschech. *pajzl* und slowak. *pajzel* ‘dürftiges Gasthaus der schlechtesten Art, Kneipe, Beisel’. (Newerkla 2004: 179–180, 410, ²2011: 173, 174, 401).

8. Schluss

Meine kurze, mit Beispielen unterlegte Aufzählung von Kernproblembereichen der Lehnwortforschung hat gezeigt, dass fundierte Lehnwortuntersuchungen letztlich nur auf hieb- und stichfesten philologischen Erkenntnissen aufbauen können, denen solide diachronische und etymologische Forschungen zugrundeliegen, um überhaupt zu allgemeingültigen Aussagen gelangen zu können. Schon der ehemalige Sprachwissenschaftler und Literaturhistoriker an der Prager Karls-Universität Alexandr Stich (1996: 6) warnte davor, das gesamte Erkenntnisinteresse auf Erscheinungen zu konzentrieren, die nicht älter als ein Jahrhundert sind. Das sei zwar vor dem Hintergrund der vorherrschenden gesellschaftlichen Überbetonung des Aktuellen und vermeintlich Zeitgemäßen, selbst wenn sich dieses als noch so ephemere erweist, durchaus zu verstehen. Übersteige jedoch diese Faszination von der Gegenwart oder vom bloß Gestrigen gewisse Proportionen, gehe dabei das allgemeine Bewusstsein für historische Zusammenhänge verloren. Damit münden die Erforschung und die linguistischen Versuche des Verstehens solcher Erscheinungen unweigerlich in ein Produzieren von Ergebnissen, die uns vor allem viel über die sprachwissenschaftlichen Theorien und Methoden der Forscher mitteilen. Ohne ein umfassendes, objektiviertes und nachprüfbares Bild von der Vergangenheit und den Ursprüngen einer Sache erscheint es schwierig, gesichertes neues Wissen zu gegenwärtigen Erscheinungen zu generieren, mögen diese für sich genommen noch so bemerkenswert sein. Das ist jetzt keine Kritik, die sich gegen neue Zugänge in der Lehnwortforschung und innovative Untersuchungsdesigns als solche richtet, sondern vielmehr ein Plädoyer dafür, dabei nicht auf die solide philologische Arbeit zu vergessen, die dazu nötig ist, um überhaupt Lehnwörter identifizieren und in ihrer weiteren Entwicklung richtig einschätzen zu können.

9. Literatur

BELLMANN, GÜNTER (1971): *Slavoteutonica. Lexikalische Untersuchungen zum slawisch-deutschen Sprachkontakt im Ostmitteleutschen* (= Studia Linguistica Germanica 4). Berlin – New York: de Gruyter.

- BELLMANN, GÜNTER (2004): Slawisch/Deutsch (Schwerpunkte). In: Besch, Werner et al. (eds.): *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. 2., vollständig neu bearbeitete und erweiterte Auflage. 4. Teilband. Berlin – New York: de Gruyter, 3229–3260.
- BERGER, TILMAN (2009): Tschechisch-deutsche Sprachbeziehungen zwischen intensivem Kontakt und puristischer Gegenwehr. In: Stolz, Christel (ed.): *Unsere sprachlichen Nachbarn in Europa. Die Kontaktbeziehungen zwischen Deutsch und seinen Grenznachbarn*. Bochum: Universitätsverlag Dr. N. Brockmeyer, 133–156.
- BERNEKER, ERICH (1908–1913, 1914): *Slavisches etymologisches Wörterbuch*. I (A–L), II/1 (М–моръ), Heidelberg: Winter.
- BIRNBAUM, HENRYK (1984): Zu den ältesten Lehnbeziehungen zwischen Slaven und Germanen. In: *Wiener Slawistischer Almanach* 13, 7–19.
- BLÁHA ONDŘEJ; DITTMANN, ROBERT; ULIČNÁ, LENKA (eds.) (2013): *Knaanic Language: Structure and Historical Background. Proceedings of a Conference Held in Prague on October 25–26, 2012* (= Judaica 9). Praha: Academia.
- BOYES BRAEM, PENNY; HAUG, TOBIAS; SHORES, PATTY (2012): Gebärdenspracharbeit in der Schweiz: Rückblick und Ausblick. In: *Das Zeichen. Zeitschrift für Sprache und Kultur Gehörloser* 90, 58–74.
- BRAUNMÜLLER, KURT; HOUSE, JULIANE (eds.) (2009): *Convergence and Divergence in Language Contact Situations*. Amsterdam, Philadelphia: John Benjamins.
- BRÜCKNER, ALEKSANDER (1926–1927, ²1957, ³1974, ⁶1993): *Słownik etymologiczny języka polskiego*. Kraków: Krakowska Spółka Wydawnicza. (Weitere Auflagen: Warszawa: Wiedza Powszechna).
- BRÜCKNER, ALEKSANDER (1929): Die germanischen Elemente im Gemeinslawischen. In: *Archiv für slavische Philologie* 42, 125–146.
- BUSCH, BRIGITTA (2013): *Mehrsprachigkeit*. Wien: Facultas WUV/UTB.
- ČERNÁ-WILLI, RAHEL (2012): Polnisches Deutsch – Deutsches Polnisch. Edition und Analyse einer Sammlung von Paralleltexten des 18. Jahrhunderts aus Teschen/Oberschlesien (= *Slavica Helvetica* 82). Bern: Peter Lang.
- ČJA (1992–2005, 2011): *Český jazykový atlas (ČJA)*. Sv. 1–5. Dodatky. Praha: Academia.
- EISNER, PAVEL (1946 [1992]): *Chrám i tvrz. Kniha o češtině*. Praha: Nákladatelství Lidové noviny.
- EISNER, PAVEL (1948 [1996]): *Čeština poklepem a poslechem*. Praha: Jiří Poláček a B. Just.
- ERSJ (1998, 2003 ff.): *Etimološki rečnik srpskog jezika*. Ogljedna sveska. Sv. 1 ff. Beograd: Srpska akademija nauka i umetnosti.
- ĖSBM (1974 ff.): *Ėtymalahičny sloŭnik belaruskaj movy*. T. 1 ff. Minsk: Belaruskaja navuka.
- ESJS (1989 ff.): *Etymologický slovník jazyka staroslověnského*. Sv. 1 ff. Praha: Academia.
- ESSJ (1973–1980): *Etymologický slovník slovanských jazyků. Slova gramatická a zájmena*. Sv. 1: *Předložky, koncové partikule*. Sv. 2: *Spojky, částice, zájmena a zájmenná adverbia*. Praha: Academia.
- ĖSSJa (1974 ff.): *Ėtimologičeskij slovar' slavjanskich jazykov. Praslavjanskij leksičeskij fond*. T. 1 ff. Moskva: Nauka.
- ESUM (1982 ff.): *Etimolohičnyj slovnyk ukrajins'koji movy*. T. 1 ff. Kyjiv: Naukova dumka.
- GEBAUER, JAN (1903–1916, ²1970): *Slovník staročeský*. 1 (A–J). 2 (K–netbalivost, netbánlivost): Praha: Nákladem České grafické společnosti "Unie". ²Academia.

- HASPELMATH, MARTIN; TADMOR, URI (eds.) (2009): *Loanwords in the world's languages. A comparative handbook*. Berlin – New York: de Gruyter.
- HIRT, HERMANN (1893): Zu den germanischen Lehnwörtern im Slavischen und Baltischen. In: *Beiträge zur Geschichte der deutschen Sprache und Literatur* 23, 330–351.
- HOLUB, JOSEF; KOPEČNÝ, FRANTIŠEK (1952, ³1953): *Etymologický slovník jazyka českého*. Praha: Státní nakladatelství učebnic.
- HOLUB, JOSEF; LYER, STANISLAV (1967, ²1973, ⁴1992): *Stručný etymologický slovník jazyka českého se zvláštním zřetelem k slovům kulturním a cizím*. Praha: Státní pedagogické nakladatelství.
- HOLZER, GEORG (1990): Germanische Lehnwörter im Urslavischen: Methodologisches zu ihrer Identifizierung. In: Holzer, Georg (ed.): *Croatica • Slavica • Indoeuropaea. Wiener Slavistisches Jahrbuch Ergänzungsband VIII*. Wien: Österreichische Akademie der Wissenschaften, 59–67.
- HOLZER, GEORG (2001): Slavisch-deutsche Lautgeschichte im österreichischen Kontaktbereich. In: *Onoma* 36, 91–110.
- HOLZER, GEORG (2013): Urslavische Wortlautungen. In: Grković-Mejdžor, Jasmina; Loma, Aleksandar (eds.): *Miklosichiana bicentennialia. Zbornik u čast dvestote godišnjice rođenja Franca Miklošiča* (= Posebna izdanja SANU knj. DCLXXV / Odeljenje jezika i književnosti 58). Beograd: SANU, Staroslovenski odbor, 65–87.
- HSSJ (1991–2008): *Historický slovník slovenského jazyka*. Zv. I–VII. Bratislava: Veda.
- JANKO, JOSEF (1916): Poznámky a příspěvky k českému slovníku etymologickému. In: *Časopis pro moderní filologii* 5, 97–104, 204–209, 292–296, 407–412.
- JANKO, JOSEF (1917): Poznámky a příspěvky k českému slovníku etymologickému. In: *Časopis pro moderní filologii* 6: 19–24, 118–123, 219–226, 321–329.
- JESPERSEN, OTTO (^o1938 [1952]): *Growth and Structure of the English Language*. Oxford: Oxford University Press.
- JUNGMANN, JOSEF (1835–1839): *Slovník česko-německý*. I–V. W Praze: W kněžecj arcibiskupské knihtiskárně, u Josefy widowy Fetterlowé, řizenjm Wáclawa Špinky.
- KARLÍKOVÁ, HELENA (1999): Die Wanderung des griechischen *κύβαλον* durch die slavischen Sprachen. In: *Münstersches Logbuch zur Linguistik* 6, 19–23.
- KĄTNY, ANDRZEJ; OLSZEWSKA, IZABELA; TWARDOWSKA, ALEKSANDRA (eds.) (2013): *Ashkenazim and Sephardim: A European Perspective*. Frankfurt am Main: Peter Lang.
- KIPARSKY, VALENTIN (1934): *Die gemeinlavischen Lehnwörter aus dem Germanischen. Akademische Abhandlung* (= *Annales Academiae Scientiarum Fennicae / Suomalaisen tiedeakatemia toimituksia*, ser. B 32, 2). Helsinki: Finnische Literaturgesellschaft.
- KLOFEROVÁ, STANISLAVA (1996): Deutsche Sprachreflexe in der tschechischen Lexikographie. In: Bremer, Ernst; Hildebrandt, Reiner (eds.): *Stand und Aufgaben der deutschen Dialektlexikographie. II. Brüder-Grimm-Symposium zur historischen Wortforschung*. Berlin – New York: de Gruyter, 49–56.
- KNUTSSON, KNUT (1929): *Die germanischen Lehnwörter im Slavischen vom Typus buky* (= *Lunds Universitets Årsskrift*, Avd. 1, 21, 9), Lund: Gleerup.
- KNUTSSON, KNUT (1938): Zur Frage der ältesten germanischen Lehnwörter im Slavischen. In: *Zeitschrift für Slavische Philologie* 15, 121–140.
- LAAKSO, JOHANNA (2004): Derivation, morphopragmatics, and language contact – on the role of German influence in Estonian word-formation. In Hyvärinen, Irma; Kallio, Petri; Korhonen, Jarmo (eds.):

- Etymologie, Entlehnungen und Entwicklungen. Festschrift für Jorma Koivulehtu zum 70. Geburtstag* (= Mémoires de la Société Néophilologique de Helsinki LXIII). Helsinki: Société Néophilologique, 171–180.
- KURZOVÁ, HELENA (1996): Mitteleuropa als Sprachareal. In: *Acta Universitatis Carolinae – Philologica* 5, *Germanistica Pragensia* 13, 57–73.
- MACHEK, VÁCLAV (1940): Germano-slavisches Wortstudien. In: *Časopis pro moderní filologii* 26, 161–167.
- MACHEK, VÁCLAV (1957): *Etymologický slovník jazyka českého a slovenského*. Praha: Nakladatelství ČSAV.
- MACHEK, VÁCLAV (²1968, ³1971 [1997]). *Etymologický slovník jazyka českého*. Praha: Nakladatelství ČSAV. [Fotoreprint Nakladatelství Lidové noviny.]
- MARTYNOV, VIKTOR VLADIMIROVIČ (1963): *Slavjano-germanskoe leksičeskoe vzaimodejstvie drevnejšej pory. (K probleme prarodiny slavjan)*. Minsk: Izdatel'stvo Akademii nauk BSSR.
- MARTYNOV, VIKTOR VLADIMIROVIČ (1969): *Tipologija i vzaimodejstvie slavjanskich i germanskich jazykov*. Minsk: Nauka i tehnika.
- MATZENAUER, ANTONÍN (1870): *Cizí slova v slovanských řečech*. Brno: Matice moravská.
- MAYER, ANTON (1927): *Die deutschen Lehnwörter im Tschechischen* (= Forschungen zur Sudetendeutschen Heimatkunde 3). Reichenberg: Gebrüder Stiepel.
- MENZEL, THOMAS; GERD HENTSCHEL (2003): *Wörterbuch der deutschen Lehnwörter im Teschener Dialekt des Polnischen unter Mitarbeit von Pavel Jančák und Jan Balhar* (= *Studia Slavica Oldenburgensia* 10). Oldenburg: BIS – Bibliotheks- und Informationssystem der Universität Oldenburg.
- MIKLOSICH, FRANZ (1867): Die Fremdwörter in den slavischen Sprachen. In: *Denkschriften der kaiserlichen Akademie der Wissenschaften in Wien. Philosophisch-historische Classe* 15, 73–140.
- MIKLOSICH, FRANZ (1872–1880): *Über die Mundarten und die Wanderungen der Zigeuner Europas*. Band 1–12. Wien: K. Gerold's Sohn.
- MIKLOSICH, FRANZ (1886): *Etymologisches Wörterbuch der slavischen Sprachen*. Wien: Wilhelm Braumüller.
- MLADENOV, STEFAN (1909): *Starite germanski elementi v slavjanskite ezici* (= *Sbornik za narodni umotvorenija, nauka i knižnina* 25). Sofija: Bălgarsko knižovno družestvo.
- MUYSKEN, PIETER (2000): *Bilingual Speech. A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- NEWERKLA, STEFAN MICHAEL (2004, ²2011): *Sprachkontakte Deutsch – Tschechisch – Slowakisch. Wörterbuch der deutschen Lehnwörter im Tschechischen und Slowakischen: historische Entwicklung, Beleglage, bisherige und neue Deutungen* (= *Schriften über Sprachen und Texte* 7). Frankfurt am Main: Peter Lang.
- NEWERKLA, STEFAN MICHAEL (2007a): Areály jazykového kontaktu ve střední Evropě a německo-český mikroareál ve východním Rakousku. In: *Slovo a slovesnost* 68, 271–286.
- NEWERKLA, STEFAN MICHAEL (2007b): Kontaktareale in Mitteleuropa. In: Kątny, Andrzej (ed.): *Słowińsko-niesłowińskie kontakty językowe. Slawisch-nichtslawische Sprachkontakte. Materiały z międzynarodowej konferencji naukowej zorganizowanej przez Wydział Filologii Wszechnicy Mazurskiej i Instytut Filologii Germańskiej Uniwersytetu Gdańskiego w dniach 27–28 czerwca 2005 r. Słowińsko-niesłowińskie kontakty językowe w perspektywie dia- i synchronicznej. Slawisch-*

- nichtslawische Sprachkontakte in dia- und synchronischer Sicht*. Olecko: Wydawnictwo Wszechnicy Mazurskiej, 29–48.
- NEWERKLA, STEFAN MICHAEL (2009): *Prečo nemôžeme nehrešiť proti slovenčine...* Zu Norm und Kodifikation des gegenwärtigen Slowakischen in der Außensicht. In: Lehečková, Eva; Táborský, Jan (eds.): *Struktura, variety, funkce. Sborník k 70. narozeninám prof. Oldřicha Uličného* (= Acta Universitatis Carolinae, Philologica 2, 2006. Slavica Pragensia XLI). Praha: Karolinum, 43–55.
- NEWERKLA, STEFAN MICHAEL (2010): Über die Wichtigkeit historiologische Zugänge für die Sprachkontaktforschung. In: *VIEW[Z]. Special issue: Historical Linguistics in the Modern Philologies / Historische Linguistik in den modernen Philologien. Vienna English Working PaperS* 19, 4, 36–51.
- NEWERKLA, STEFAN MICHAEL (2011): Slowakisch und Tschechisch – Distanz trotz Nähe. Norm und Kodifikation des Slowakischen in der Außensicht. In: Kuße, Holger; Woldt, Claudia (eds.): *Tschechisch und Slowakisch: Nähe und Distanz. Beiträge zum 4. Bohemicum Dresdense 13.–14. November 2009* (= Specimina philologiae Slavicae 163). München: Verlag Otto Sagner, 169–184.
- NEWERKLA, STEFAN MICHAEL (2014a): Grundfragen der Lehnwortforschung anhand slawischer Lehnbeziehungen zum Deutschen. In: Bock, Bettina; Kozińska, Maria (eds.): *Schleichers Erben. 200 Jahre Forschung zum Baltischen und Slawischen* (= Studien zur historisch-vergleichenden Sprachwissenschaft 6). Hamburg: Baar, 65–78.
- NEWERKLA, STEFAN MICHAEL (2014b): Mehrsprachigkeit und lexikalische Konvergenz — Gemeinsame Konversationismen in den Sprachen der ehemaligen Habsburgermonarchie. In: Cwanek-Florek, Ewa; Nöbauer, Irmgard (eds.): *Deutsch und die Umgangssprachen der Habsburgermonarchie* (= Symposien und Seminare am Wissenschaftlichen Zentrum der Polnischen Akademie der Wissenschaften in Wien 11). Wien: Polnische Akademie der Wissenschaften, Wissenschaftliches Zentrum in Wien, 11–27.
- PONTIUS, JASON (1997): Language codification and the perception of otherness: The case of Czech and German. In: *Chicago Linguistic Society* 33, *The Panels* (1997), 101–108.
- REJZEK, JIŘÍ (2001, ²2012, ³2015): *Český etymologický slovník*. Voznice: Leda.
- RUDOLF, RAINER (1991): *Die deutschen Lehn- und Fremdwörter in der slowakischen Sprache* (= Beiträge zur Sprachinselforschung 9). Wien: Verlag des Verbandes der wissenschaftlichen Gesellschaften Österreichs.
- SCHNEEWEIS, EDUARD (1912): Lautlehre der deutschen Lehnwörter im Tschechischen. In: *XV. Jahresbericht der Landes-Oberrealschule in Zwittau über das Schuljahr 1911/12*. Zwittau: Verlag der Anstalt – Marcell Morvay, 3–40.
- SKÁLA, EMIL (1992): Deutsch und Tschechisch im mitteleuropäischen Sprachbund. In: *brücken. Germanistisches Jahrbuch. Neue Folge* 1 (1991/1992), 173–179.
- SKÁLA, EMIL (2000): Středoevropský jazykový svaz. In: *Přednášky z XLIII. běhu Letní školy slovanských studií*. Praha: Univerzita Karlova, 77–85.
- SKOK, PETAR (1971–1974): *Etimologijski rječnik hrvatskoga ili srpskoga jezika*. Knj. 1–4. Zagreb: Jugoslavenska akademija znanosti i umjetnosti.
- SNOJ, MARKO (²2003): *Slovenski etimološki slovar*. 2. pregladana in dopolnjena izdaja. Ljubljana: Modrijan.
- ŠRÁMEK, RUDOLF (1998): Zur Wortgeographie der deutschen Lehnwörter in den tschechischen Mundarten. In: Bauer, Walter –Scheuringer, Hermann (eds.): *Beharrsamkeit und Wandel. Festschrift für Herbert Tatzreiter zum 60. Geburtstag*. Wien: Edition Praesens, 295–306.
- SSČ (²1994, ³2003): *Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia.

- SSN (1994, 2006): *Slovník slovenských nářečí*. I. A–K. II. L–P (povzchádzat'). Bratislava: Veda
- STČS (1968): *Staročeský slovník. Úvodní stati, soupis pramenů a zkratk*. Praha: Academia.
- STČS (1968–2008): *Staročeský slovník*. 1.–26. díl. Praha: Academia.
- STENDER-PETERSEN, ADOLF (1927): *Slavisch-Germanische Lehnwortkunde*. Göteborg: Elander
- STICH, ALEXANDR (1996): *Od Karla Havlíčka k Františku Halasovi (lingvoliterární studie)*. Praha: Torst.
- ŠTRELJ, KAREL (1904): *Zur slavischen Lehnwörterkunde* (= Denkschriften der kaiserlichen Akademie der Wissenschaften in Wien. Philosophisch-historische Classe 50, 3), Wien: Gerold.
- THOMAS, GEORGE (1997): The Role of German Loanwords in the Slavic Languages of the Former Habsburg Empire. In: *Canadian Slavonic Papers / Revue canadienne des slavistes* 39, 3–4, 333–359.
- THOMAS, GEORGE (2008): Exploring the Parameters of a Central European Sprachbund. In: *Canadian Slavonic Papers / Revue canadienne des slavistes* 50, 1–2, 123–153.
- THOMASON, SARAH GREY; KAUFMAN, TERRENCE (1988): *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- UHLENBECK, CHRISTIANUS CORNELIUS (1890): *Die lexikalische urverwandtschaft des baltoslavischen und germanischen*. Leiden: Blankenberg & Co. – Leipzig: K. F. Koehler.
- UHLENBECK, CHRISTIANUS CORNELIUS (1893): Die germanische Lehnwortkunde im Altslavischen. In: *Archiv für slavische Philologie* 15, 481–492.
- ULIČNÁ, LENKA (2011): IV. Hlavní proudy středověkého (pre)aškenázského myšlení a tzv. pražská komentátorská škola. Hledání identity v podmínkách izolace a integrace. In: Šedinová, Jiřina; Kieval, Hillel J. (eds.): *Dialog myšlenkových proudů středověkého judaismu. Mezi integrací a izolací*. Praha: Academia, 268–331.
- VASMER, MAX JOHANN FRIEDRICH (1932): Altgermanische Lehnwörter im Slavischen. In: *Zeitschrift für Slavische Philologie* 9, 367–370.
- VASMER, MAX JOHANN FRIEDRICH (1933): Etymologisches. 17. čech. hampejz. In: *Zeitschrift für Slavische Philologie* 10, 401.
- VASMER, MAX JOHANN FRIEDRICH (1934): Ältere germanische Lehnwörter im Slavischen. In: *Zeitschrift für Slavische Philologie* 11, 50–51.
- VINCENZ, ANDRÉ DE (1986): Lehnwörter in Kontakt. Zu einer Klasse deutscher Lehnwörter. In: Olesch, Reinhard; Rothe, Hans (eds.): *Festschrift für Herbert Bräuer zum 65. Geburtstag am 14. April 1986*. Köln – Wien: Böhlau, 713–727.
- WITTMANN, HENRI (1991): Classification linguistique des langues signées non vocalement. In: *Revue québécoise de linguistique théorique et appliquée* 10, 1, 215–288.

Catharina Stamata Krebs-Garić (Wien)

Lehn- und Erbwortschatz des Kroatischen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ im Lichte von Etymologie und Empirie

Abstract

For the following article, the semantic field “religion and belief” in Croatian has been investigated with the purpose of identifying loan and inherited vocabulary. Etymological data helped to decide whether or not a lexeme is to be considered a loanword (or calque, loan meaning extension, or loan creation, respectively). Contact situations, in which most of the loanwords are supposed to be borrowed, were discussed in a separate chapter. Since there occurred more than one counterpart for most of the meanings in the LWT meaning list (part of them being archaic, colloquial, or dialectal), it has proved impossible to investigate the whole amount of loanwords in the mentioned field and to make a comparison with other languages in WOLD. Thus, the determination of the most common Croatian terms remains a desideratum for scientific research.

1. Einleitung

Für den vorliegenden Artikel wird der Lehn- und Erbwortschatz des Kroatischen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ systematisch aufgearbeitet, um so das Ausmaß an Lehn- und Erbwörtern in ebendieser Gruppe ermitteln zu können. Ausgangspunkt der Untersuchung sind die bei Haspelmath/Tadmor (2009a) angeführten 26 Bedeutungen der semantisch-lexikalischen Gruppe „religion and belief“; diese sind wiederum Teil einer 1.460 lexikalische Bedeutungen umfassenden „meaning list“, welche Grundlage für die empirische Untersuchung des Lehn- und Erbwortschatzes in 41 Sprachen der Welt im Zuge des *Loanword Typology (LWT)* Projekts¹ ist. Die Ziele dieses Projekts, die dabei verwendeten Definitionen von lexikalischen und semantischen Entlehnungen, sowie die damit verbundene Problematik werden in Kap. 2 behandelt.

In Kap. 3 werden die kroatischen² (standardsprachlichen wie dialektalen) Entsprechungen zu den Bedeutungen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ der „LWT meaning list“

¹ Das Resultat des LWT-Projekts ist die World Loanword Database (WOLD, erreichbar unter <<http://wold.cldd.org>> [aufgerufen am 28.07.2015].

² Abgekürzt als kr. Weitere verwendete Abkürzungen sind (neben allg. üblichen) arch. (archaisch), pejor. (pejorativ), dial. (dialektal), Bed. (Bedeutung) sowie (nach Sprachzweigen geordnet): idg. (indogermanisch), ursl. (urslawisch), vorursl. (vorur斯拉wisch), nachursl. (nachur斯拉wisch), štok. (štokavisch), nštok. (neuštokavisch), čak. (čakavisch), kajk. (kajkavisch), sr. (serbisch), bosn. (bosnisch), slow. (slowenisch), bulg. (bulgarisch), č. (tschechisch), poln. (polnisch), nsorb. (niedersorbisch), osorb. (obersorbisch), russ. (russisch), aruss. (altrussisch), ksl. (kirchenslawisch), aksl. (altkirchenslawisch), bsl. (baltoslawisch), balt. (baltisch), lett. (lettisch), lit. (litauisch), aprs. (altpreußisch), rom. (romanisch), lat. (lateinisch), klat. (kirchenlateinisch), mlat. (mittellateinisch), vlat. (vulgärlateinisch), alat. (altlateinisch), ital. (italienisch), aital. (altitalienisch), dalm. (dalmatisch), istr. (istriotisch), vegl. (veglitisch), friul. (friulanisch), afriul. (altfriulanisch), venez. (venezianisch), frz. (französisch), port. (portugiesisch), rum. (rumänisch), kelt. (keltisch), germ. (germanisch), got. (gotisch), engl. (englisch), dt. (deutsch), nhd. (neuhochdeutsch), mhd. (mittelhochdeutsch), ahd. (althochdeutsch), abair. (altbairisch), aisl. (altisländisch), gr. (griechisch), mgr. (mittelgriechisch), ngr. (neugriechisch), ved. (vedisch), sanskr. (sanskritisch), avest. (avestisch),

behandelt. Ermittelt werden diese anhand des *Veliki rječnik hrvatskoga jezika* (Anić 2003), durch die bereits vorliegenden Entsprechungen bei Buck (1949/³1971) und Miklosich (1876), sowie mithilfe von Einzelstudien zu diesem Themenbereich (Auty 1976, Katičić 1980, Skok 1953, Meillet 1926); zusätzlich werden der *Rječnik stranih riječi* von Anić/Klaić/Domović (2002) und der *Rječnik sinonima hrvatskoga jezika* von Šarić/Wittschen (2008) hinzugezogen. Die jeweiligen Entsprechungen werden hinsichtlich ihrer Etymologie, ihrer lautgeschichtlichen und semantischen Entwicklung untersucht, wobei auf den *Etimologijski rječnik hrvatskoga ili srpskoga jezika* (ERHSJ) (Skok 1971–1973), sowie auf die erwähnten Einzelstudien zurückgegriffen wird. Zusätzlich wird entsprechendes Material von Derksen (2008), Matasović (2008), Holzer (et al. 2010, 2011, et al. 2012), Pronk-Tiethoff (2012) und dem *Lexikon der indogermanischen Verben* (LIV) (Rix 2001) hinzugezogen.

Die anhand der betreffenden Entsprechungen im Falle von Lehnwörtern ersichtlichen Gebersprachen werden in Kap. 4 unter Berücksichtigung des kulturhistorischen Kontextes der jeweiligen Kontaktsituationen (anhand von Krahwinkler 1992, Katičić 1999, Zagiba 1961, 1964, sowie der bereits angeführten Einzelstudien) behandelt. In Kap. 5 werden die in Kap. 3 ermittelten kr. Entsprechungen für die Bedeutungen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ nochmals in einer Tabelle zusammengefasst, wobei der Lehn- bzw. Erbwortstatus und die Analysierbarkeit der jeweiligen Entsprechung berücksichtigt wird. Weiters werden erste daraus ersichtliche Ergebnisse diskutiert. Auf die sich im Zuge der Bearbeitung ergebenden (v. a. die Synonymie betreffenden) Probleme und hiermit zusammenhängende Forschungsdesiderate wird in Kap. 6 eingegangen.

2. Das Loanword Typology (LWT) Project

Im LWT-Projekt wurde primär das Ziel verfolgt, lexikalische Entlehnungen unter typologischen Gesichtspunkten zu untersuchen und so die Entlehnwahrscheinlichkeit (*borrowability*)³ von Bedeutungen empirisch zu ermitteln und generelle Muster bei der Distribution von Lehnwörtern systematisch zu erforschen (vgl. Haspelmath 2009: 35, Haspelmath/Tadmor 2009b: 1). Durch den Fokus des Projekts auf lexikalische Entlehnungen wurden innerhalb einer Sprache von deren Sprechern analysierbare Einheiten als Erbwörter und somit sowohl Calques als auch mit entlehnten Basen gebildete Wörter nicht als Lehnwörter erachtet, da diese mit – wenn auch entlehntem – Material in der Nehmersprache (und nicht in

pers. (persisch), apers. (altpersisch), npers. (neupersisch), iran. (iranisch), arm. (armenisch), ung. (ungarisch), türk. (türkisch), atürk. (alttürkisch), arab. (arabisch), awar. (awarisch), hebr. (hebräisch), aram. (aramäisch), äthiop. (äthiopisch), syr. (syrisch).

³ Vgl. hierzu die auf dieser Grundlage erstellte *Leipzig-Jakarta-Liste* des entlehnresistenten Basisvokabulars bei Haspelmath/Tadmor (2009b: 69–71, vgl. a. Haspelmath 2009: 36). Die *borrowability* wurde anhand des Lehnwortstatus (*unborrowed score*) einer Bedeutung, deren Vorkommen in den Sprachen des Projekts (*representation score*), der Analysierbarkeit der Entsprechungen (*analyzability/simplicity score*) und dem Alter der jeweiligen Entsprechungen (*age score*) ermittelt (vgl. Tadmor 2009: 65–68).

der Gebersprache) gebildet wurden (vgl. Haspelmath/Tadmor 2009b: 12–15, Haspelmath 2009: 37–39, Tadmor 2009: 66).

Ein Lehnwort (bzw. eine lexikalische Entlehnung, vgl. Haspelmath 2009: 36) wird definiert „[...] as a lexeme that has been transferred from one lect into another and is used as a word (rather than as an affix, for example) in the recipient language“ (Haspelmath/Tadmor 2009b: 13). Bei Lehnwörtern handelt es sich um in der Nehmersprache normalerweise monomorphemische, von Sprechern nicht analysierbare Einheiten (vgl. Haspelmath 2009: 37). Dagegen werden morphosyntaktisch komplexe (analysierbare) Wörter oder Phrasen in (fast⁴) allen Fällen in der Nehmersprache gebildet. „Such words were not considered loanwords, even when they contained borrowed elements“ (Haspelmath/Tadmor 2009b: 12). Lexeme mit entlehnten Basen wurden also nicht als Lehnwörter erachtet – es sei denn, die in der Nehmersprache hinzugefügten Morpheme waren Teil des Entlehnungsprozesses oder Teil der Zitierform des Wortes (d. h. z. B. bei Verben die infinite anstelle einer entlehnten finiten Form) (vgl. Haspelmath/Tadmor 2009b: 15). Morphologische Adaptation (z. B. die Ersetzung eines fremden Suffixes durch ein eigenes oder der Eingliederung ins Deklinationssystem dienende Suffigierung) ist daher nicht als Derivations-, sondern als Teil des Entlehnungsprozesses zu erachten (vgl. Bartels 2009a: 320–321).

Im Gegensatz zu lexikalischen Entlehnungen (*material borrowings*) stehen strukturelle Entlehnungen (*structural borrowings*), von denen Calques (Lehnübersetzungen, *loan translations*) die wichtigste Gruppe ausmachen (vgl. Haspelmath 2009: 38–39). „A calque (or loan translation) is a complex lexical unit (either a single word or a fixed phrasal expression) that was created by an item-by-item translation of the (complex) source unit“ (Haspelmath 2009: 39). Die einzelnen Bestandteile der Lehnübersetzung entsprechen dabei semantisch jenen des Wortes in der Gebersprache (vgl. Haspelmath/Tadmor 2009b: 14).

Zwei weitere hier relevante Arten der strukturellen Entlehnung sind Bedeutungserweiterungen (*loan meaning extensions*), wobei es sich um „[...] an extremely common (and often unnoticed) process whereby a polysemy pattern of a donor language word is copied into the recipient language“ (Haspelmath 2009: 39) handelt, und semantische Entlehnungen (oder *loan creations*), bei denen ein bereits vorhandenes (Erb-)Wort in der Nehmersprache eine zusätzliche Bedeutung nach dem Modell der (Neben-)Bedeutung des entsprechenden Wortes in der Gebersprache erhält (vgl. Haspelmath/Tadmor 2009b: 14, Haspelmath 2009: 39–40). Auch semantische Entlehnungen wurden nicht als Lehnwörter erachtet: „Semantic borrowing (the transfer of meaning without the transfer of words) was not the focus of our

⁴ So wurden z. B. gr. *ἀντί* ‘gegen’ und *χριστός* ‘Christus; der Gesalbte’ in kr. *ἀντικρίστ* ‘Antichrist, Teufel’ (vgl. 3.2.18) auch einzeln ins Kr. entlehnt, die einzelnen Bestandteile sind in der kr. Wortbildung produktiv und Sprechern des Kr. verständlich.

project, which dealt strictly with lexical borrowing (which must involve the transfer of forms together with meanings)“ (Haspelmath/Tadmor 2009b: 14).

Bei entlehnten Wörtern ist zudem relevant, ob es sich um in der Nehmersprache etablierte Wörter (*established borrowings*) oder um sog. *nonce-borrowings* (z. B. Code-Switching) handelt⁵ (vgl. Haspelmath/Tadmor 2009: 12). Zudem müssen Fremdwörter, welche nicht in das System der Nehmersprache integriert sind und von Sprechern leicht als solche erkannt werden, von Lehnwörtern unterschieden werden (vgl. Haspelmath 2009: 43), wenngleich sie dennoch als Lehnwörter gezählt werden.

Da nun aber semantische und strukturelle Entlehnungen (Calques und Bedeutungserweiterungen) sowie mit entlehnten Basen gebildete Wörter in der semantisch-lexikalischen Gruppe „Religion und Glaube“ im Kr. einen nicht unerheblichen Teil ausmachen und die enge Begriffsdefinition eines Lehnworts von Haspelmath/Tadmor (2009b), wie eben dargelegt, problematisch ist, werde ich im Folgenden auch semantische und strukturelle Entlehnungen in die Untersuchung miteinbeziehen (vgl. Kap. 5). Hinzu kommt, dass sich wichtige Kulturkontakte infolge der Christianisierung der Slawen ab dem 7./8. Jh. – und auch im mythologischen Bereich – gerade an (mittlerweile größtenteils ins sprachliche System integrierten) Lehnwörtern und Calques ablesen lassen (vgl. Kap. 4).

3. Die semantisch-lexikalische Gruppe „Religion und Glaube“

Im Folgenden werden zu den Bedeutungen der „LWT meaning list“ die kr. Entsprechungen im Lichte ihrer Etymologie behandelt und deren (ev. auch abweichende) Bedeutungen angegeben. Dabei wird im Falle von Lehnwörtern das Wort in der/den in Frage kommenden Gebersprache(n) nebst Behandlung der Entlehn- und Kontaktumstände und im Falle von Erbwörtern die ursl., bsl.⁶ oder idg. Rekonstruktion angeführt. Ebenso wird festgestellt, ob für eine Bedeutung mehrere kr. Entsprechungen vorliegen.

Die semantisch-lexikalische Gruppe „Religion und Glaube“ wurde gewählt, da diese lt. Tadmor (2009: 64) mit 41,2% den höchsten Gesamtlehnwortanteil aller in der WOLD vertretenen Sprachen aufweist. Auch das Niedersorbische, die einzige in der WOLD vertretene slaw. Sprache, weist in ebendieser Gruppe einen hohen Lehnwortanteil von 52,1% auf (vgl. Bartels 2009a: 314). Da nun das Kr. als besonders puristische Sprache gilt (vgl. u. a. Wingender 1997, Thomas 1988, s. Kap. 6.2), ist es auch Ziel dieser Arbeit zu ermitteln, in welchem Ausmaß Erb- und Lehnwörter hier vertreten sind.

⁵ Dies kann anhand der Vorkommenshäufigkeit eines Lehnworts in der Nehmersprache entschieden werden (vgl. Haspelmath 2009: 40–41). Näheres zum Einbezug der Frequenz eines Wortes findet sich in Kap. 6.3.

⁶ Mit dem Begriff *Baltoslawisch* werden hier sprachliche Entwicklungen bezeichnet, welche sowohl das Balt. als auch das Slaw. betreffen, somit auf gemeinsamen Isoglossen beruhen und auf einen sehr engen Sprachkontakt (oder sogar Sprachbund) zwischen den bereits selbstständigen Sprachzweigen zurückzuführen sind. Keineswegs ist hiermit eine (hypothetische und umstrittene) gemeinsame Zwischengrundsprache der balt. und slaw. Sprachen gemeint.

3.1. „Religion and belief“ bei Haspelmath/Tadmor (2009a) und „religion and superstition“ bei Buck (1949/³1971)

Dass die Zuordnung von Bedeutungen zu semantisch-lexikalischen Gruppen arbiträr ist (vgl. Haspelmath/Tadmor 2009b: 6, Buck 1971: xiii-xiv) und zudem Unterschiede in deren Umfang vorliegen, zeigt ein Vergleich der hier behandelten Gruppe „Religion und Glaube“ bei den bereits vorliegenden Untersuchungen: „religion and belief“ bei Haspelmath/Tadmor (2009a) umfasst 26 Bedeutungen, „religion and superstition“ bei Buck (1971) 32 Bedeutungen. Wie die Benennung der Gruppen schon zeigt, kommen bei Buck (1971) zusätzlich ‘Aberglaube’ und ‘Heide’ hinzu, welche bei Haspelmath/Tadmor (2009a) fehlen. Bei Buck (1971) finden sich zudem die Bedeutungen ‘Schutzgeist’, ‘Teufel’, ‘Engel’ und ‘taufen’, sowie (als Unterkategorien von ‘Priester’) ‘Pfarrer’, ‘Mönch’ und ‘Nonne’. Die bei Haspelmath/Tadmor (2009a) angeführten Bedeutungen ‘Moschee’⁷, ‘Initiationsritus’ und ‘Beschneidung’ fehlen wiederum bei Buck (1971). Diese Tatsache unterstreicht auch, dass es sich bei Bedeutungen aus der religiösen Terminologie, die auf den ersten Blick in beiden Listen fehlen mögen (wie ‘Kreuz’, ‘Christ’, ‘Messe’ oder ‘Gottesdienst’⁸) oder nur in einer der Listen vorhanden sind, eben nicht um universelle religiöse Begriffe handelt, da sie sich nur selten religionsübergreifend finden. Beschränkt man eine Untersuchung der semantisch-lexikalischen Gruppe „Religion und Glaube“ allerdings auf die slaw. Sprachen (wie Miklosich 1876), so müsste diese um eine Reihe von Bedeutungen v.a. aus der christlichen, jüdischen und islamischen Terminologie erweitert werden.⁹ Um den Vergleich mit den anderen Sprachen in der WOLD zu ermöglichen, wird im Folgenden nur auf jene Bedeutungen eingegangen, die sich auch in der „LWT meaning list“ finden. Lediglich ‘Pfarrer’, ‘Messe, Predigt’, ‘Teufel’ und ‘Engel’ werden aufgrund ihrer semantischen Nähe den Bedeutungen ‘Priester’, ‘predigen’, ‘Dämon’ und ‘Geist’ hinzugefügt (vgl. Kap. 3.2.10, 3.2.12, 3.2.18 und 3.2.23).

3.2. Kroatische Entsprechungen in der semantisch-lexikalischen Gruppe „Religion und Glaube“

⁷ Im Prinzip sind aber auch ‘Kirche’ und ‘Moschee’ nur religionspezifische Arten eines ‘Tempels’ – es sei denn, mit ‘Tempel’ wird speziell ein ‘heidn. Tempel’ gemeint, vgl. hierzu (insb. zur Bedeutungsentwicklung) Kap. 3.2.3 und 3.2.4.

⁸ In der vorliegenden Untersuchung wird eine Zuordnung von ‘Messe, Gottesdienst’ zu ‘predigen’ vorgenommen (s. 3.2.12).

⁹ Allein bei den Bezeichnungen für die verschiedenen Angehörigen des höheren und niederen Klerus würden sich mindestens 50 Ausdrücke finden, deren Gebrauch und Bed. je nach Zeit und Region variiert. Vgl. hierzu auch Miklosich (1876: 12–17), der eine Reihe an slaw. Funktionsbezeichnungen für ‘Papst’, ‘Bischof’, ‘Mönch’ usw. anführt, deren Grad an Spezifizierung (man denke nur an die unterschiedlichen Aufgabenbereiche und Befugnisse) sehr hoch ist. Es liegt somit schon eine fachsprachliche Terminologie vor, wobei die spezifischen Bed. wohl auch Sprechern des Kr. nicht immer bekannt sein dürften.

In fast allen Fällen sind für eine Bedeutung mehrere kr. Lexeme als Entsprechungen anzuführen (vgl. hierzu auch Kap. 6), die mehr oder weniger mit der Bedeutung in der „LWT meaning list“ übereinstimmen, čak. oder kajk. kr. Dialekten angehören, kolloquial oder archaisch sein können. Da die meisten kr. Entsprechungen nur marginal von den Bedeutungen der „meaning list“ abweichen, wird hier auf eine Indizierung des Übereinstimmungsgrades als *exact counterpart*, *sub-counterpart*, *super-counterpart* und *para-counterpart* (vgl. Haspelmath/Tadmor 2009b: 9–10) verzichtet, stattdessen werden etwaige abweichende Bedeutungen direkt bei den jeweiligen Entsprechungen angegeben.

3.2.1. ‘religion’

Eine Bezeichnung für ‘Religion’, also „[...] a belief in supernatural powers, the desire, to stand well with them, and practices devoted to this end.” (Buck 1971: 1462), der die umfassende Bed. dieses Begriffs wiedergibt, gibt es lt. Miklosich (1876: 34) in den slaw. Sprachen nicht. Als gelehrter Latinismus findet sich kr. *rèligija* f. ‘Glaube’, ‘Glaubenssystem’ (< lat. *religiō*, ev. zurückgehend auf lat. *lego*, *-ere* ‘sammeln, auswählen’, vgl. Buck 1971: 1463) (vgl. ERHSJ III: 126, Anić 2003: 1326). Kr. *vjerozákon*¹⁰ ‘Religion (als Glaubenssystem)’ ist semant. spezialisiert und schließt den ‘Glauben’ an sich nicht ein (vgl. ERHSJ III: 603), es handelt sich um ein Kompositum aus kr. *vjèra* ‘Glaube’ und *zákon* ‘Gesetz; Religion’ (aus nachursl. **za kona* ‘von Anfang an’, in ursl. Zeit gebildetes Nomen actionis von nachursl. *začēti* ‘anfangen’) vgl. ERHSJ I: 315–316, s.v. *-čēti*).

Gängiger ist hingegen kr. *vjèra* f., aksl. *věra* ‘Glaube, Vertrauen’ (als semant. Entsprechung zu gr. *πίστις*, lat. *fides*, vgl. Miklosich 1876: 43), bei dem es sich um ein Erbwort handelt (vgl. ERHSJ III: 602–603) und welches lt. Anić (2003: 1742) neben ‘Religion’ aber doch auch das ‘Glaubenssystem’ umfasst. Die ursprüngl. Bed. ‘Glaube, Vertrauen’ ist auch ahd. *gilouba* eigen und stellt die gängigste semant. Quelle (neben ‘Verehrung’, ‘Gottesdienst’ und ‘Verständnis’) für ‘Religion’ dar (vgl. Buck 1971: 1462–1463). Ursl. **wērā* (vgl. Matasović 2008: 153) zugrundeliegendes idg. **ueh₁-r-eh₂-* findet sich auch in lat. *vērus* ‘wahr’ und ahd. *wāra* ‘Vertrag, Treue’ (vgl. Derksen 2008: 520).

Als weitere Synonyme für kr. *vjèra* sind bei Šarić/Wittschen (2008: 519) angeführt kr. *pòbožnōst* f. ‘Religiosität, Frömmigkeit’ (Anić 2003: 1056), eine Ableitung von kr. *bôg* ‘Gott’ (vgl. Kap. 3.2.2); kr.

¹⁰ Im Aksl. finden sich zudem (v.a. die ‘religiöse Lehre’ betreffend und somit in semant. Hinsicht spezialisiert) *zakonъ* ‘Gesetz’ und *učeniye* ‘Lehre’, welche in ihrer Bed. ahd. *galauba*, *galaubao* und *lera* entsprechen (vgl. Miklosich 1876: 34). Aksl. *zakonъ* wird auch als Übersetzung von gr. *θρησκεία* ‘Verehrung’ verwendet, welches sich wiederum lt. Buck (1971: 1463) ev. von der idg. Wurzel **dher-* ableitet (davon sanskr. *dhṛ-* ‘Halt, Unterstützung’ und *dharma-* ‘Brauch, Recht, Gesetz’, welches im Sanskr. auch die adäquateste Entsprechung für ‘Religion’ darstellt) (vgl. a. Kap. 3.2.15). Hier wäre m. E. zumindest ein semant. Zusammenhang zu *zakonъ* ‘Gesetz’ gegeben. In diesem Zusammenhang ist auch noch aksl. *zakonъnikъ* ‘Priester’ zu nennen, bei dem es sich um einen Calque aus ahd. *ēuwarto* (ahd. mhd. *ēwart*, ein Kompositum aus *ēwa* ‘Gesetz’ und *wart* ‘Hüter’, eigentlich ‘Hüter des Gesetzes’, vgl. Buck 1971: 1473, Miklosich 1876: 14) handelt (vgl. a. Auty 1976: 173–174). Im heutigen Kr. existiert diese Bed. nicht mehr, kr. *zákonik* m. bedeutet nur ‘Gesetzesbuch’ (vgl. Anić 2003: 1790).

(arch.) *īspovijēd, īspovijēst* f. ‘Bekenntnis’ (vgl. Anić 2003: 468) und kr. *vjeroispovijēst, veroispovijēd* f. ‘Konfession, Glaube, Zugehörigkeit zu einer rel. Gemeinschaft oder Kirche’ (vgl. Anić 2003: 1743), bei denen es sich um Ableitungen (und Komposita mit *vjěra*) vom kr. Abstraktum *vjěst* ‘Kunde, Nachricht’ bzw. direkt von aksl. belegtem *věděti* ‘wissen’ handelt (vgl. ERHSJ III: 574–576; s.v. *věm*, ERHSJ III: 591–592, s.v. *vijěst*) (vgl. a. Kap. 3.2.12); sowie kr. *konfēsija* f. ‘Glaube; Glaubensbekenntnis’ (vgl. Anić 2003: 602), eine Entlehnung von lat. *confessio* ‘Bekenntnis’; lt. Skok sind kr. *īspovijēd, īspovijēst* zudem Calques von lat. *confessio* (vgl. ERHSJ II: 139).

3.2.2. ‘god’

Für ‘Gott’ gibt es neben kr. *Bōg*¹¹ (vgl. Anić 2003: 96; daraus abgeleitet auch kr. *božànstvo* ‘Gottheit’, vgl. Anić 2003: 107, dieses allerdings auch in der Bed. ‘Götze, heidn. Gott’, vgl. a. Šarić/Wittschen 2008: 45), aksl. *bogъ* ‘Gott’ auch noch das bibl. kr. *Gōspōd* m. ‘Herr’ (vgl. Anić 2003: 369) (aksl. *gospodъ*, gr. *κύριος*, lat. *dominus*), welches nie für die Bezeichnung eines Menschen verwendet wird (hierfür nur *gospodънь, κυρίου, domini*, vgl. Miklosich 1876: 35). Idg. **g^host(i)-pot-* (in Urverwandtschaft zu lat. *hospes* ‘Gastgeber’, gr. *δεσπόζω* ‘beherrschen’) wird im Slaw. mit ungeklärtem **-d-* weitergeführt (vgl. Derksen 2008: 180). Bei kr. *bōg* m., ursl. **bagu* ‘Gott’ (vgl. Holzer 2011: 95) handelte es sich ursprüngl. um ein Adj. **bogъ* mit der substantivischen Bed. ‘*Glück, Heil; irdisches Wohl > Gut’ (vgl. a. č. *zboží* ‘Gut’ < **sъbožьje* aus nachursl. **sъbogъ*, vgl. a. sanskr. *subhāgu* ‘glücklich’), welches schon im Ursl. zu einem Agentivum mit der Bed. ‘Glück und Gut verteilende Gottheit’¹² wurde (vgl. ERHSJ II: 178–179, s.v. *bōgat*). Da sich dieselbe semant. Entwicklung auch in sanskr. *bhaga-*, *bhāgas* ‘Spender, Zuteiler, gütiger Herr, Brotherr’ (eigentlich ein Epitheton eines ved. Gottes) neben ‘Wohlstand, Glück’ und avest. *baγa-* ‘Herr’ > ‘Gott’ (apers. *baga-*, npers. nur ‘Gott’) findet, wurde oft angenommen, dass es sich bei *bōg* um eine Entlehnung aus einer iran. Sprache (durch Vermittlung über das Skythische) handelt (vgl. ERHSJ I: 179, Meillet 1926: 168). Lt. Meillet (1926: 168) beweist aber gerade diese Polysemie von ursl. **bagu*

¹¹ Bei Šarić/Wittschen (2008: 45) sind zudem *Stvoritelj* ‘Erschaffer’, *Tvorac* ‘Schöpfer’, *Svevišnji* ‘Allmächtiger’, *Višnji* ‘ib.’, *Svemogući* ‘ib.’, *Svedržeći* bzw. *Svedržitelj* ‘ib.’ (ein Calque von gr. *ὁ παντοκράτωρ* ‘Gott, Allmächtiger’, vgl. ERHSJ I: 448, s.v. *držati*), *Svevideći* ‘Allsehender’, *Sveznajući* ‘Allwissender’, *Ōtac* ‘Vater’ bzw. *Bog Ōtac* ‘Gott Vater’, *Nebeski Ōtac* ‘himmlischer Vater’, *Stvoritelj neba i zemlje* ‘Erschaffer von Himmel und Erde’ und *Dragi Bog* ‘lieber Gott’ als Synonyme aufgeführt, die allesamt als Epitheta erachtet werden können.

¹² In den italischen, kelt., balt. und indo-arischen Sprachen sind die Entsprechungen für ‘Gott’ verbunden mit Wörtern für ‘Firmament (Himmel)’, ‘Tag’ (mit ursprüngl. Bed. ‘scheinen, hell’). Hierzu gehören die auf idg. **deiyō-s* ‘Gott’ (neben **d̥iēu-*, **diu-* ‘Himmel’, ‘Tag’) zurückzuführenden Entsprechungen wie gr. *Ζεύς* (gen. *Διός*) und lat. *Iūpiter* (< gr. *Ζεῦ πάτερ*) mit der allg. Bed. ‘hell, leuchtend’. Nur die slaw. und iran. Entsprechungen für ‘Gott’ gehen auf einen Ausdruck zurück, der ‘gütiger Verteiler’ bedeutet hat (vgl. Buck 1971: 1464, Meillet 1926: 167) (vgl. a. Kap. 4.3).

(‘Gott’ und ‘Wohlstand’) das Vorliegen eines Erbwortes, da auch der Ableitung *bogatъ* (ursl. **bagā`tu*¹³, vgl. Holzer et al. 2012: 8) ‘reich, begütert’ die substantivische Bed. ‘Gut’ zukommt¹⁴ (vgl. a. Miklosich 1876: 35, der zusätzlich *ubogъ* ‘arm’ anführt); Matasović (2008: 47) sieht aufgrund dieser semant. Entwicklung jedoch zumindest einen iran. Einfluss auf das Slaw. (wenn nicht gar eine Entlehnung) als gegeben. Lt. Derksen (2008: 50) ist eine Entlehnung von kr. *bôg* aus einer iran. Sprache wahrscheinlicher: „This hypothesis is supported by the fact, that the etymon does not show the effects of Winter’s Law¹⁵.“ Eine weitere Parallele zu den iran. Sprachen besteht darin, dass mit Ausdrücken wie *bělbog*, *crn(o)bog*, *daždbogъ* > *Dažbog*, *Dabog*, *Zao bog* schon in der vorursl. Mythologie ein Dualismus wie in der zoroastrischen Religion des Awesta bestand (*dobar* und *zao bog*). Im Christentum wurde das ursl. Wort in der Bed. des ‘guten Gottes’ übernommen (vgl. ERHSJ I: 179). Die hier erfolgte semant. Übertragung des christl. Inhalts auf einen Ausdruck, der vormals die heidn. slaw. Götter bezeichnete (vgl. Miklosich 1876: 34), ist eine durchaus übliche Entwicklung: „The old words for a pagan ‘god’ were generally retained for the Christian ‘God’.“ (Buck 1971: 1464)

3.2.3. ‘temple’

Das kr. Wort für ‘Tempel’ *hrâm* m. (im čak. auch ‘Kirche’), aksl. *hramъ*¹⁶ (lt. Anić 2003: 415 sowohl ‘Tempel einer Gottheit’ als auch ‘Ort, an dem der Gottesdienst abgehalten wird’) findet sich in der gesamten Slavia, jedoch gibt es keine balt. Parallelen und auch die idg. Wurzel ist lt. Skok ungesichert. Am ehesten kommt von den bei Skok angeführten Möglichkeiten m. E. aus semant. Gründen wohl idg. **(s)ker-* ‘scheren, kratzen, abschneiden’ (vgl. LIV: 558–559) in Betracht, wodurch sich die ursprüngl. Bed. ‘abgetrennter Teil des Hauses’, später ‘Hof, Speicher (als Teil des Hauses)’ und in weiterer Folge ‘Tempel’, manchmal auch ‘heidn. Tempel’ entwickelte (vgl. ERHSJ I: 683). Diese Bedeutungsentwicklung findet sich lt. Buck (1971: 1465) auch in den Bezeichnungen der meisten idg.

¹³ Anstelle der synthetischen Notation der kr. Akzentuierung, Quantität und Intonation stellt Holzer (vgl. 2011: XII) diese Prosodeme unabhängig voneinander dar: $\grave{a} = \underline{a}$, $\hat{a} = \bar{a}$, $\grave{a} = \underline{a}$, $\acute{a} = \bar{a}$. ($\bar{a} = \bar{a}$), $_ =$ Betonung, $\cdot =$ nštok. steigende Intonation.

¹⁴ Skok (ERHSJ I: 180) führt in diesem Zusammenhang noch die für ‘Schwertlilie, Iris germanica’ verwendeten Bezeichnungen *bôžūr* (als Ableitung von *bôg*) und *perûnika* (ursl. **perawnîkâ/*perawnejkâ*, vgl. Holzer 2011: 139) (als Ableitung von *Perun*, dem Namen der obersten slaw. Gottheit, ursl. **Perawnu* oder **Perawnu*, vgl. Holzer 2011: 138–139) an: „Riječ *bog* je mogla biti slavenska apozicija za *Perun*, a u toj apoziciji nije potrebno baltoslovensko slaganje. Riječ *božur* (i rus., steslav. te bug.) dokazuje nadalje da je *bog* animističkim putem ušao i u botaničku terminologiju.“ (ERHSJ I: 180) Hier muss man allerdings bedenken, dass auch eine Apposition (als Epitheton) entlehnt worden sein kann.

¹⁵ Winters Gesetz ist eine der wichtigsten bsl. Isoglossen und besagt, dass Vokale zwischen stimmhaften Konsonanten (und anscheinend nur in geschlossenen Silben) gelängt werden (vgl. Matasović 2008: 56).

¹⁶ Aksl. *crъky* ‘Kirche’ wird häufig als Übersetzung von gr. *ναός* ‘Tempel, Schrein’ und *ἱερόν* ‘heiliger Ort, Schrein, Tempel’ verwendet, aksl. *hramъ* ‘*Haus’ übernimmt erst später die Bed. ‘Tempel’ (auch ‘Kirche’) (vgl. Buck 1971: 1466).

Sprachen für ‘Tempel’: „Many of the words for ‘temple’ are from ‘dwelling, house’ (with ‘god’ expressed or understood) or ‘holy place’, derived from adjs. for ‘holy’.“

Daneben existieren im Kr. auch noch unüblicheres *sakrarij* m. ‘Heiligtum, Tempel’ (ein gelehrter Latinismus aus lat. *sacrarium*, vgl. ERHSJ III: 191), *světište*¹⁷ n. ‘Heiligtum, Tempel’, eine Ableitung von kr. *svět* ‘heilig’ (vgl. Kap. 3.2.11), und *idòlija* f. ‘heidn. Tempel’¹⁸, eine Entlehnung aus gr. *εἰδωλεῖον* (vgl. ERHSJ I: 709, s.v. *îdō*) (vgl. a. 3.2.19). Šarić/Wittschen (2008: 134) führen (neben *bogòmolja* und *čřkva*, vgl. Kap. 3.2.4) noch kr. *pànteōn* m. als Synonym für ‘Tempel’ an. Es handelt sich um einen Archaismus in der Bed. ‘Tempel aller Götter’ bzw. ‘alle Götter’, der lt. Anić (2003: 992) mittelbar über lat. *panthēum* aus gr. *πάνθειον* entlehnt wurde (aus lautlichen Gründen ist aber eine direkte Entlehnung aus dem Gr. vorzuziehen).

3.2.4. ‘church’

Für die (dem Christentum spezifische) Bed. ‘Kirche’ findet sich die kr. Entsprechung *čřkva* f. ‘Kirche (sowohl Bauwerk als auch Organisation)’ (aksl. *čřky*¹⁹) (vgl. a. Anić 2003: 147). Bei diesem Lehnwort wurde mittels Suffigierung eine Adaptation zur Integration ins morphologische System vorgenommen, die genaue Herkunft von kr. *čřkva*, nachursl. **čřky* f. (*ŭ*-Stamm, vgl. Pronk-Tiethoff 2002: 99) ist jedoch umstritten. Gr. *κυριακόν* (sc. *δῶμα*) ‘Haus des Herrn’ war zwar unüblicher als gr. *ἐκκλησία* ‘*(politische) Versammlung’, es wird aber angenommen, dass es durch die Bed. ‘Kirche (als Bauwerk)’ durch eine unbelegte got. Form in die übrigen germ. und slaw. Sprachen gelangte und die Bedeutungserweiterung auf ‘Kirche (Organisation)’ von lat. *ecclesia* beeinflusst worden ist (vgl. Buck 1971: 1476–1477). Nun ist es allerdings fraglich, ob im Got. jemals ein solches Wort existiert hat bzw. die Notwendigkeit einer Entlehnung aus dem Gr. ins Got. bestanden hat, wenn als Bezeichnungen für ‘Kirche’ ohnehin got. *aikklesji* und *gudhūs* ‘Gotteshaus’ belegt sind (vgl. Pronk-Tiethoff 2012: 99). Miklosich (1876: 17) nimmt eine Entlehnung aus ahd. *chiricha*, *chirihha* an (welches er wiederum entweder auf gr. *κυριακόν*, *κυριακή* oder lat. *circus* zurückführt). Lt. Skok (ERHSJ I: 275) sind als Vermittlersprachen des zugrundeliegenden gr. *κυρι(α)κή* (sc. *στοά*) auch abair. *chirch* oder got. **kyrikō* < gr. *κυρικόν* (volkssprachlich für *κυριακόν*) in Betracht zu ziehen; zudem ist hervorzuheben, dass das germ. Genus mit dem slaw. Genus

¹⁷ Anić (2003: 1515) führt hierfür auch kr. *sanktuarij* als Synonym auf, welches allerdings keinen eigenen Eintrag aufweist.

¹⁸ Speziell den ‘heidn.’ bzw. ‘Götzentempel’ bezeichnen aksl. *kapište* (aksl. *kapъ* m. ‘Götzenbild’, lt. Skok wird angenommen, dass es sich um eine Entlehnung aus dem Awar. handelt; im Kr. hat sich nur das Toponym *Kapište* (Insel bei Lastovo) erhalten, vgl. ERHSJ II: 38), aksl. *trěbište* ‘Ort, an dem geopfert wird’ (vgl. hierzu auch den Namen der Ansiedlung *Trebišće* unter dem Berg *Perūn* bei Mošćenica in Istrien, vgl. Skok 1953: 61, Anm. 68), sowie aksl. *kumirište* ‘Ort der Götzen’ (vgl. Miklosich 1876: 37) (vgl. a. *kumirъ* unter 3.2.7).

¹⁹ Aksl. *hramъ* und *trěbište* wurden zwar auch als Bezeichnungen für ‘Kirche’ verwendet (vgl. Miklosich 1876: 18), in weiterer Folge jedoch verdrängt bzw. nicht mehr in dieser Bed. gebraucht, da vor der Christianisierung existierende Wörter für ‘Tempel’ immer mit dem Heidentum in Verbindung gebracht wurden (vgl. a. Buck 1971: 1477).

übereinstimmt. Den in Frage kommenden germ. Formen (ahd. *kirihha*, *ckīlihha*, bair. dial. *chirhha*) liegt jedenfalls ein westgerm. **kirikō* f. (*ō*-Stamm) ‘Kirche’ zugrunde (vgl. Pronk-Tiethoff 2012: 99). Es kann jedoch auch nicht ausgeschlossen werden, dass die verschiedenen Formen auf mehrfache Entlehnungen zurückzuführen sind (vgl. Auty 1976: 174), denn die unterschiedlichen „[...] forms in the different Slavic languages cannot be derived from a single proto-form. The vocalism of the attested forms points to four different vowels in the initial syllable: **cbr-*, **cir-*, **cer-* and **cěr-*.“ (Pronk-Tiethoff 2012: 99) (vgl. a. die Diskussion dieser Annahme bei Pronk-Tiethoff 2012: 100)

Kr. *bogòmolja* ‘Bethaus in Form einer Kirche, einer Halle, eines Tempels usw.’ deckt nur einen Teil der Bed. von ‘Kirche’ ab (ohne ‘Kirche als Organisation’) bzw. ist als übergeordneter Begriff ‘Ort, an dem gebetet wird’ zu erachten (vgl. Anić 2003: 98); es handelt sich um ein Kompositum mit *bôg* ‘Gott’ und *mòliti* ‘beten’ (vgl. ERHSJ I: 180 s.v. bogat) (vgl. Kap. 3.2.2, 3.2.9). Als weitere Synonyme für ‘Kirche, Bethaus’ führen Šarić/Wittschen (2008: 46)²⁰ kr. *sinagóga* f. (vgl. Anić 2003: 1401), eine mittelbare Entlehnung über lat. *synagōga* aus gr. *συναγωγή* (vgl. ERHSJ III: 238); sowie den Archaismus (und Calque von lat. *synagōga* oder gr. *συναγωγή*) kr. *zbornica* f. ‘*Tempel, Synagoge’ (heute ‘Versammlungsort’) (vgl. Anić 2003: 1828) an, bei dem es sich um ein Denominativum von kr. *zbôr* ‘Versammlung’ handelt, welches wiederum eine Ableitung von kr. *brãti*, nachursl. **b̥rati*, bsl. **ber-* (vom idg. Präsensstamm **b^her-e/o-* der idg. Verbalwurzel **b^her-* ‘tragen, bringen’, vgl. LIV: 76–77) ist (vgl. ERHSJ I: 201–202, 238; Derksen 2008: 73).

3.2.5. ‘mosque’

Bei kr. *džámija*²¹ f. ‘Moschee’ handelt es sich um einen Balkanturzismus (< türk. *cami*, vgl. a. ngr. *τζαμί*) arab. Ursprungs (arab. *ġami* ‘Verband, Verein’) (vgl. ERHSJ I: 470, vgl. a. Anić 2003: 273). Daneben existiert noch unüblicheres kr. *mošéja* f. ‘Moschee’ (vgl. Anić 2003: 778), welches lt. Skok aus arab. *mäsġid* mittelbar zunächst über das Türk. und dann über ital. *moschea* oder frz. *mosquée* und schließlich aus nhd. *Moschee* ins Kr. entlehnt wurde; sowie kr. *mèčet* m. (neben *mečit*) ‘Moschee’, ‘besonderer Grabstein’ bzw. *mèsdžid*, *mesdžet* m. ‘islam. Bethaus ohne Minarett’ (vgl. Anić 2003: 745), bei dem es sich um einen direkt entlehnten Balkanturzismus (türk. *mesced*, *-id*) arab. Ursprungs handelt (arab. *mäsġid* ‘Ort der Anbetung’) (vgl. ERHSJ II: 396).

3.2.6. ‘altar’

²⁰ Neben *Božja kuća* ‘Gotteshaus’, *katedrala* ‘Kathedrale’, *hram* ‘Tempel’, *bazilika* ‘Basilika’, *kapela* ‘Kapelle’, *oratorij* ‘Oratorium’, *kripta* ‘Krypta’ und *džamija*, *mošeja*, *mesdžid* ‘Moschee’ (vgl. 3.2.5).

²¹ Bei *-ija* handelt es sich nicht um ein Ableitungssuffix, sondern um das Anfügen der Endung *-ja* zur Eingliederung von türk., ung. und gr. Lehnwörtern mit der Endung *-i* ins kr. Deklinationssystem (vgl. ERHSJ I: 711–712, s.v. *-ija*).

Als Entsprechung für ‘Altar’ ist kr. *òltār* m. (aksl. *ol̋tarь, oltarь*) anzuführen, das entweder mittelbar über ahd. *altari* (vgl. Miklosich 1876: 18) bzw. über mgr. *ἀλτάριον* oder direkt (wie auch die meisten europ. Wörter für ‘Altar’, so auch č. *oltař*, poln. *oltarz*) aus lat. *altāria, altāre*²² ‘hoher Altar; Brandopfer’ entlehnt wurde (vgl. ERHSJ I: 29, s.v. *àltār*, Buck 1971: 1467; Anić 2003: 935; lt. Matasović 2008: 167, 310 aus dem Vlat. oder frühen Rom. entlehnt). „Kako ne pokazuje metatezu likvida (kao *Albona* > *Labin*), ne potječe iz prvog vremena (7. v.) i nije nastala kontaktom naroda evangelizacijom i nikako ne ide u praslav. doba, premda se nalazi u svim slavinama.“ (ERHSJ I: 29) Die Entlehnung muss noch vor dem Wandel *a* > *o* erfolgt sein, welcher sich im Laufe des 9. Jh. vollzog (vgl. Holzer 2011: 58; Matasović 2008: 102). Nach diesem Wandel wurde selteneres (bei Anić 2003 nicht angeführtes) čak. *òltār*, aksl. *al̋tarь* > russ. *altarь* (vgl. ERHSJ I: 29) erneut entlehnt.

Kr. *žrtvenīk* (aksl. *žr̋tvište, žr̋tvenīkь*, russ. *žertvennik*) ist eine Ableitung aus kr. *žrtva* oder russ. *žertva* ‘Opfergabe, Opfer’ (vgl. a. 3.2.7) und wird (ebenso wie aksl. *tr̋bište, tr̋bьnikь* aus *tr̋ba*, vgl. Miklosich 1876: 18) auch als Bezeichnung eines heidn. Altars gebraucht, hat also schon im Aksl. eine von *ol̋tarь* verschiedene Bed. (vgl. Miklosich 1876: 18, Buck 1971: 1467, ERHSJ III s.v. *žrtva*, Anić 2003: 1870, Šarić/Wittschen 2008: 308).

3.2.7. ‘sacrifice’

Kr. *žrtva*²³ f. ‘dargebrachtes Opfer’ (vgl. a. Buck 1971: 1468; Miklosich 1876: 31, Anić 2003: 1870) dürfte lt. Skok als kirchl. Terminus erst mit Karadžić’ Übersetzung des NT als Russismus in die Volkssprache gelangt sein (vgl. ERHSJ III: 685). Russ. *žertva*, aksl. *žr̋tva*²⁴, slow. *žrtva, žrtav* und seltenes č. *žertva* ‘Opfer, Opfergabe’ sind Ableitungen mit *-tva* von nachursl. **žrti* (aruss. *žereti* ‘opfern’, aksl. *žbr̋, žrti* ‘opfern’) (vgl. Derksen 2008: 566). Das Verb hat sich nur im Russ. erhalten, im Kr. hat es lt. Skok (ERHSJ III: 685) aufgrund der Vermeidung von Homonymie mit aksl. *žbr̋, žr̋ti* (ursl. **ger-*, nachursl. **žer-ti*, kr. *ždrijèti* ‘schlingen’, vgl. ERHSJ III: 673–674) keine Spuren hinterlassen

²² Lat. *āra* (alat. *āsa* < **ās-* mit ungesicherter Bed. ‘brennen, glühen’), *altāria* (wohl von *adolēre* ‘brennen’, im Volksverständnis aber mit lat. *altus* ‘hoch’ verbunden) dürfte lt. Buck (1971: 1467) anfangs einen ‘Altar für Brandopfer’ bezeichnet haben. Lt. Skok (ERHSJ I: 29) ist die lat. Form eine Ableitung von *altus* ‘hoch’ mithilfe des Suffixes *-aris* < *-alis*.

²³ Semant. Entsprechungen sind gr. *θυσία* (*‘Brandopfer’ aus *θύω* ‘opfern < toben, schäumen’), gr. *ἱερείων* ‘Tieropfer’ (*ἱερός* ‘heilig’) (vgl. Buck 1971: 1467), lat. *sacrificium* (*sacer* ‘heilig’ und *facere* ‘machen’); im Lat. finden sich als weitere Entsprechungen für ‘Opfer’ *victima* ‘Tieropfer’ und *hostia* ‘ib.’; letzteres mit ungeklärter Etymologie (vgl. Buck 1971: 1468).

²⁴ Im Aksl. findet sich auch *tr̋ba* (*idol̋skaja, jelin̋skaja*) ‘(heidn.) Opfer’ (vgl., Miklosich 1876: 31). Vgl. a. sr. *treba* ‘heidn. Opfer’, sr. *trebnik* ‘Ritual’, russ. *tr̋ba* ‘Altar’, aksl. *tr̋bište* ‘Tempel’. Aus ursl. **terb-* ‘*lichten, abholzen, roden > reinigen, säubern’ leitet sich *tr̋ba* ‘*Arbeit, Tätigkeit’ ab, dessen Bed. sich lt. Skok über ‘Abholzung’ zu ‘Notwendigkeit’ in kr. *trijeba* f. entwickelte. Da lt. Skok auch religiöse Tätigkeiten im ursl. Heidentum als eine ‘Notwendigkeit’ erachtet wurden, fand das Wort Eingang in die religiöse Terminologie und wurde in der Bed. ‘Opfer’ im Christentum übernommen. Idg. Kognaten sind ungesichert, ev. kommen got. *þorban*, nhd. *bedürfen* in Frage. (Vgl. ERHSJ III: 501–502, s.v. *trijèbiti* (se))

(denominales kr. *žrtovati* ‘opfern’ dürfte somit eine jüngere Ableitung aus entlehntem *žrtva* sein²⁵). Dieselbe Homonymie veranlasste wohl auch Derksen (2008: 566) aksl. *žr̃b̃ti*, *žbr̃q* und aksl. *žr̃ēti*, *žbr̃q* demselben (nachursl.) Etymon **žb̃rti*; **žerti* ‘opfern’ zuzuordnen und beide auf bsl. **gir̃?*-, idg. **g^hrH-* zurückzuführen. Im LIV hingegen werden sie als zwei gesonderte Etyma behandelt: aus der idg. Verbalwurzel **g^herH-* ‘Zustimmung bekunden’ geht schwundstufiges aksl. *žbr̃j̃q* (*žbr̃q* ist sekundär), *žr̃b̃ti* ‘opfern’ hervor (vgl. ved. *á gurá-* zur Bedeutungsentwicklung ‘zustimmen’ über ‘begrüßen’ zu ‘preisen’) (vgl. LIV: 210–211), hingegen leitet sich aus dem Präsensstamm **g^hrh̃₃-é-* der idg. Verbalwurzel **g^herh̃₃-* ‘verschlingen’ schwundstufiges ursl. **žer-* und aksl. *žbr̃q*, *žr̃ēti* ab (vgl. LIV: 211–212).

Als weitere Entsprechungen für ‘Opfer’ (von denen einige jedoch selten oder nur dial. gebräuchlich sind) finden sich: kajk. *aldōv* ‘Opfer’ (und *aldovati* ‘opfern’), eine Entlehnung aus dem ung. Partizip *aldó* der Verbalwurzel *ald* ‘opfern; segnen’ (vgl. ERHSJ I: 26, s.v. *aldomāš*, Miklosich 1876: 32); kr. *kūrbān* m. ‘Opfer’, aber auch die ‘Feierlichkeiten rund um den Festtag, an dem dieses Opfer dargebracht wird’ (vgl. Anić 2003: 657) ist ein Turzismus aus türk. *kurban* < arab. *kurban*, auch in ngr. *κορβάκι* (vgl. ERHSJ II: 242); čak. *oblānta* f. ‘Oblate, Hostie’ ist wohl ein Überbleibsel aus dem Dalm., aus substantiviertem lat. Partizip Perfekt *oblata*²⁶ von *offero*, mit *n*-Epenthese im Gegensatz zu slow. *ōblat* m. (vgl. ERHSJ II: 537), bei Anić (2003: 897) findet sich nur (bei Šarić/Wittschen 2008: 291 als kolloquial markiertes) lt. Anić aus dem Ital. entlehntes *oblānda* ‘Oblate, Hostie’, welches allerdings m. E. eher auf das Friul. als Gebersprache verweist, da es die Ergebnisse der westromanischen Sonorisierung²⁷ aufweist (vgl. a. ERHSJ II: 537); čak. *ofar* m. ‘Opfer, Gabe’ ist ein Germanismus (protestantischen) lat. Ursprungs aus ahd. *offar*, *opfar*, mhd. *opfer*, *opfern* < lat. *offerre* ‘darbringen; opfern’ (die ahd. Form mit *pf* ist lt. Buck (1971: 1468) und Skok (ERHSJ II: 545) wohl aus einer Überkreuzung mit klat. *operāri* im Sinne von ‘religiöse Riten durchführen’ > ‘opfern’ entstanden) (vgl. Miklosich 1876: 32, Buck 1971:

²⁵ Ein weiteres Indiz für eine Entlehnung aus dem Russ. könnte m. E. das Ausbleiben des Wandels *žr* > *ždr* (vgl. Holzer 2011: 78) sein, der im 15. Jh. nach der zweiten Liquidametathese und nach Verlust des schwachen *ɔ* (als Auslöser für diesen Wandel) stattgefunden hat. „Do umetanja *-d-* nije došlo u riječi *žrtva* < *žb̃rtva* zbog izbjegavanja teške skupine **ždrt-*.“ (Matasović 2008: 162) Diese Umgehung eines Lautgesetzes zur Vermeidung von Konsonantenclustern müsste dann nicht notwendig gewesen sein, wenn das Wort aufgrund der oben erwähnten Homonymie erst später aus dem Russ. (wo dieser Wandel nicht durchgeführt wurde, da es anstelle der Liquidametathese zu einem Vokaleinschub vor der Liquida kam, vgl. a. kr. *ždrijèbe* vs. russ. *žerebéc* < ursl. **džerben* ‘Fohlen’, vgl. Matasović 2008: 150; 162; Holzer 2011: 165–166) entlehnt und in weiterer Folge „kroatisiert“ (vgl. hierzu die Ausführungen von Thomas 1988: 73, 101, vgl. a. Kap. 6.2) worden ist.

²⁶ Bei lat. *oblata* dürfte es sich lt. Skok um einen Calque aus gr. *προσφορά* handeln (vgl. ERHSJ III: 51, s.v. *proför*).

²⁷ Bei der westromanischen Sonorisierung (oder Lenierung) werden die intervokalischen Tenues *p*, *t*, *k* (auch in Kombination mit Resonanten) zu den Mediae *b*, *d*, *g*; sie tritt z. B. beim kr. Exonym *Čabdad* aus rom. **Čiβ̃ēdāde* < lat. *Cīvitātem* auf, jedoch nicht bei den kr. Endonymen *Cāvtat* und *Cāptat*, welche Entlehnungen desselben Toponyms, allerdings aus dem autochthonen Rom. auf heutigem kr. Gebiet sind. Die westrom. Sonorisierung fand spätestens im 6. Jh., in nördl. ital. Dialekten wohl schon im 5. Jh. statt, und betrifft alle westrom. Sprachen, somit auch das Friul., nicht aber das Ital. und balkanroman. Sprachen wie Rum. und Dalm. (Vgl. Holzer 2011: 19, 97–99; Katičić 1980: 29, Skok 1953: 53, Anm. 13; 54, Anm. 25).

1469); kr. *ostija* f., ‘Hostie’ ist als gelehrter Latinismus aus lat. *hostia* ‘Tieropfer’ bzw. über ital. *ostia* ‘Symbol des Opfers Christi’ vermittelt worden (vgl. ERHSJ II: 572), daneben findet sich noch der gelehrte Latinismus *hòstija* als kirchl. Terminus für ‘Hostie’ (vgl. Anić 2003: 415); kr. *prosfor* m., *pròsfora* f. ‘Hostie’ ist ein kirchl. Gräzismus aus *προσφορά* ‘Opfer, Opfergabe’ (Präfix *πρός* ‘hin zu’ und *φέρω* ‘tragen’) (vgl. ERHSJ III: 51, Anić 2003: 1237); als Calque des gr. Wortes könnte m. E. aksl. *prinosa* (vgl. a. Miklosich 1876: 31) angesehen werden; kr. *proskura* ‘Hostie’ (vgl. Anić 2003: 1237), auch *proskomidija* f. (vgl. ERHSJ III: 52 s.v. *prosfora*) ‘(orthodoxer) Altar, Opferstätte; dargebrachtes Opfer (als Teil der kath. Eucharistie)’ ist ein kirchl. vermittelter Balkangräzismus (< ngr. *προσκομιδή* ‘Dargebrachtes’); kr. *svetba* ‘Sakrament, Opfer, Hostie’ ist ein Abstraktum von kr. *svēt* ‘heilig’ (vgl. ERHSJ III: 370, s.v. *svēt*) (vgl. Kap. 3.2.11). Als weitere Synonyme für *žrtva* finden sich bei Šarić/Wittschen (2008: 559) (neben *žrtvena životinja* ‘geopfertes Tier’ und *prinošenje žrtve* ‘dargebrachtes Opfer’) kr. *pâljenica* f. ‘religiöses (Brand-)Opfer’ (vgl. Anić 2003: 988), wohl eine Ableitung von kr. *pâliti* ‘brennen’, ursl. **pol-* (vgl. ERHSJ II: 593), idg. **pōlh₁-* (vgl. Derksen 2008: 390) bzw. (ungesichert, da nur im Slaw. weitergeführt) **pel-* ‘aufflammen’ (vgl. LIV: 469); und kr. *posvećenje* n. ‘dargebrachtes Opfer’ (vgl. Anić 2003: 1119), eine Ableitung von kr. *svēt* ‘heilig’ (s. Kap. 3.2.11).

3.2.8. ‘to worship’

Als kr. Entsprechung für ‘als Gottheit verehren’ findet sich *posvétiti*, *posvećivati* ‘heiligen, ehren’, bei dem es sich um eine Ableitung von kr. *svēt* ‘heilig’ handelt (vgl. ERHSJ III: 370, s.v. *svēt*, Anić 2003: 1119) (vgl. a. 3.2.11). Bei kr. *obožavati* ‘als Gottheit verehren’ (vgl. Anić 2003: 900) handelt es sich um eine Ableitung von kr. *bôg* ‘Gott’ (vgl. 3.2.2). Kr. *(po)štòvati* aus nachursl. *čtovati* ist ein Denominativum von kr. *čâst* ‘Ehre’ (Faktitivum *čtiti* aus nachursl. **čtb*) (ERHSJ I: 298–299, s.v. *čâst*, Anić 2003: 1554). Aksl. *čbtq*, *čisti* ‘zählen, lesen (Schrift); ehren’ lässt sich aus auf Grundlage des Wurzelaorists der idg. Verbalwurzel **k^heġt-* ‘bemerken, erkennen’ gebildetem **k^hit-é-* ableiten (vgl. LIV: 382–383).

Kr. *släviti* ‘feiern, preisen’ ist ein Denominativum des gesamt- und ursl. Nomen actionis *slàva* ‘Ruhm, Ehre’, bei dem es sich um die Dehnstufe zur *o*-Stufe **slov* ‘Wort’ (vgl. Vollstufe *sluti* ‘heißen, berühmt sein’) handelt (vgl. ERHSJ III: 288–289, s.v. *sluti*, Anić 2003: 1425). Aus idg. **k^hleu-os-* ‘Ruhm, Ehre’ (vgl. Matasović 2008: 85), eine Erweiterung der idg. Verbalwurzel **k^hleu-* ‘hören’ (vgl. LIV: 334), leitet sich bsl. **šlōu²-* (vgl. Derksen 2008: 453) und schließlich ursl. **slāwā*, kr. *slava* (= *slàva*) ab (vgl. Holzer 2011: 149) (vgl. a. 3.2.13). Bei kr. *poklòniti se* ‘huldigen’ (vgl. Anić 2003: 1081) fand eine semant. Entwicklung bei *poklòniti* von ‘Ehrerbietung darbringen’ zu ‘geben, schenken’ und ‘den religiösen Akt des Gebets vollziehen’ statt (vgl. ERHSJ II: 102–103, s.v. *klòniti se*), vgl. a. kr. *klànjati se* ‘verbeugen’; ksl. *klaněti se*, *pokloniti se* (ursprüngl. ‘beugen’) wurde regelmäßig als Übersetzung für gr.

προσκυνέω, wörtl. ‘verbeugen’ verwendet (vgl. Buck 1971: 1470). Es könnte in Verbindung mit kr. *sloniti* ‘(an)lehnen’ (vgl. kr. *záslon* und *záklon* ‘Zufluchtsort’) stehen (vgl. ERHSJ III: 287) und somit auf die idg. Verbalwurzel **k̑leǵ-* ‘sich anlehnen’ (vgl. LIV: 332) zurückzuführen sein.

Als weitere Synonyme (vgl. Šarić/Wittschen 2008: 292, 356, 357, 432) finden sich kr. *divinizírati* ‘als Gottheit verehren’ (vgl. Anić 2003: 224), eine Ableitung von kr. *divin* ‘göttlich’, welches wiederum eine Entlehnung aus ital. *divino* m. aus lat. *dīvīnus* ‘göttlich’ ist (vgl. ERHSJ I: 410); kr. *díviti se* ‘anbeten; bewundern’ (vgl. Anić 2003: 225) ist eine Ableitung vom kr. Adj. *dīvan* ‘herrlich’ (vgl. ERHSJ I: 409–410), nachursl. **divьnъ* ‘erstaunlich, toll’, eine Ableitung mit *-ьнъ* von nachursl. **divь* m. ‘Erstaunen, Verwunderung’, die idg. Wurzel ist ungesichert, aber lt. Derksen wohl **d^hHi-uo-* oder **d^heiH-uo-* (vgl. Derksen 2008: 108–109); bei kr. *adorírati*, *adòrati*, *adorávati* ‘verehren’ (vgl. Anić 2003: 3) handelt es sich um eine Entlehnung aus ital. oder lat. *adōrāre* ‘verehren, anbeten’ (vgl. ERHSJ I: 9); kr. *uvážávati* ‘(ver-)ehren’ (vgl. Anić 2003: 1699) ist eine Ableitung von kr. *vága* ‘Waage’ (vgl. *vágati* ‘abwägen’, *vážiti* ‘gelten’) und dieses wiederum eine noch aus (nach)ursl. Zeit stammende Entlehnung aus ahd. *vaga*, die sich lt. Skok in allen slaw. Sprachen außer dem Bulg. findet (vgl. ERHSJ III: 558–59); kr. *uznòsiti*, *ùznijèti* ‘anpreisen, verherrlichen’ (vgl. Anić 2003: 1707) ist eine Ableitung von kr. *nòsiti* ‘tragen’ (vgl. ERHSJ II: 512–513, s.v. *nèsti*), nachursl. **nositi*, bsl. **nós-* (vgl. lit. *našinti* ‘verkünden’), idg. **h₁nok-* (vgl. Derksen 2008: 356), Iterativum von idg. **h₁nek-* ‘erhalten, nehmen’ (vgl. LIV: 250); kr. *glorificírati* ‘verehren’ ist eine Entlehnung aus lat. *glōrificāre* ‘verherrlichen, preisen’ (vgl. Anić 2003: 357, Anić/Klaić/Domović 2002: 502); kr. *veličati* ‘lobpreisen’ (vgl. Anić 2003: 1722) stellt eine Ableitung des kr. Adj. *velik* ‘groß’, nachursl. **velikъ* dar (vgl. ERHSJ III: 573), lt. Derksen (2008: 514) erfolgte die Ableitung mittels *-ik* von idg. **uelh₁-* ‘drängen, pressen’ bereits in bsl. Zeit; kr. *hváliti* ‘loben, preisen’ (vgl. Anić 2003: 421) ist ein Denominativum von kr. *hval* f. ‘Lob, Ruhm, Verherrlichung; Dank’, ursl. **xwālīns* (vgl. Matasović 2008: 260), die idg. Wurzel ist ungesichert (vgl. ERHSJ I: 696–697).

3.2.9. ‘to pray’

Kr. *mòliti se*²⁸ ‘beten’ (gegenüber *mòliti* ‘bitten, beten’), aksl. *moliti se* ‘beten’ (vgl. ERHSJ II: 451, Anić 2003: 772; vgl. a. Miklosich 1876: 28, Buck 1971: 1471) liegt die bsl. Wurzel **mold-* zugrunde, welche mit *-la* zu *moldlā* erweitert wurde. Da sich hiervon č. *modla* ‘Götzenbild’ (bsl. **molda?*, vgl. Derksen 2008: 320) erhalten hat, dürfte *moliti se* schon im slaw. Heidentum die Anbetung einer Gottheit bedeuten und durch die Christianisierung die heutige Bed. ‘beten’ übernommen haben. Eine andere (und lt. Skok wahrscheinlichere) Erklärung liefert bulg. *molitva* ‘blutiges Opfer’, welches den Gebrauch des Reflexivums *se* plausibel macht und somit den Schluss nahelegt, dass in ursl. Zeit *molitva* ‘Opfer’ und

²⁸ Bei Šarić/Wittschen (2008: 238) finden sich für *moliti se (Bogu)* als Synonyme nur *upućivati molitve* ‘Gebete aussenden’ und *izgovarati/govoriti molitve* ‘Gebete aussprechen’.

moliti ‘opfern’ bedeutet haben muss (vgl. ERHSJ II: 452). Die zugrundeliegende idg. Verbalwurzel ist **meld^h*- ‘feierlich sprechen, verkünden’, von deren Iterativum **mold^h-éje-* sich aksl. *moljǫ*, *moliti* ‘bitten, flehen’ ableiten lässt. Ursrl. **modliti* setzt eine singuläre Metathese **ld > dl* voraus, welche noch vor der Liquidametathese anzusetzen ist (vgl. LIV: 432, vgl. a. Derksen 2008: 320).

3.2.10. ‘priest’

Bei kr. *svěćenik* m. ‘Priester’ (aksl. *svěštenikъ*²⁹, lt. Miklosich (1876: 13) ein Calque von gr. *ἱερέυς*, ‘Priester’, vgl. gr. *ἱερός* ‘heilig’) handelt es sich um ein mittels des Suffixes *-ik* substantiviertes Partizip Perfekt Passiv von kr. *svétiti* ‘heiligen’, eine Ableitung von *svēt* ‘heilig’ (vgl. ERHSJ III: 371–372, Anić 2003: 1513) (vgl. a. 3.2.11); kr. *jěrej* m. ‘orthodoxer Priester’ (vgl. Anić 2003: 523) ist eine Entlehnung von gr. *ἱερέυς* (vgl. ERHSJ I: 774). Der Archaismus kr. *glagòljāš* m. bezeichnet einen ‘kath. Priester, der die Messe in aksl. Sprache und mit Büchern in glagolitischer Schrift hält’ (vgl. Anić 2003: 351) und leitet sich von kr. *glagol* m. ‘*Wort’ (heute nur ‘Verb’) (vgl. ERHSJ I: 564–565) ab, nachursl. **golgolati* ‘sprechen’ ist ein Denominativum von nachursl. **golgolъ* ‘Rede’, lt. Derksen eine Reduplizierung der Schallwurzel **gol-* aus idg. **gal-* (wie auch in nachursl. **gòlsъ* m. ‘Stimme’, bsl. **golsos*, lt. Derksen wohl idg. **gal-so-*) (vgl. Derksen 2008: 174, 176). Kr. *latināš* m. bezeichnet hingegen einen ‘Priester, der die Messe in lat. Sprache hält’ (vgl. Anić 2003: 673) und ist eine Ableitung von kr. *Làtin* in der Bed. ‘Katholik’, welches wiederum einen gelehrten Latinismus aus lat. *latīnus* (ein mittels *-īnus* abgeleitetes Ethnikon des Choronyms *Latium*) darstellt (vgl. ERHSJ II: 275).

Kr. *dùhovnĭk* (neben *dùhōvnĭk*) ‘Geistlicher; Ordensmann, Mönch’ ist (wie auch aksl. *duchovъnikъ*) ein Calque aus dem Ahd. (vgl. nhd. *Geistlicher*) oder aus gr. *πνευματικός* (vgl. Miklosich 1876: 14, ERHSJ I: 372, s.v. *dāh*, Anić 2003: 264); ein weiterer Calque ist kr. *dušòbrižnĭk* ‘Seelsorger’ aus nhd. *Seelsorger* (vgl. ERHSJ I: 373, Anić 2003: 267) (vgl. a. 3.2.23 zu kr. *dùh* ‘Seele, Geist’).

Kr. *pǒp* ‘Priester der röm. u. gr. Kirche’ (im Kr. in pejor. Bed., vgl. Anić 2003: 1100, lt. Šarić/Wittschen 2008: 455 kolloquial) findet sich in der gesamten Slavia. Es ist ungeklärt, ob es sich um ein Lehnwort aus ahd. *phapho*, *phaffo* oder gr. *παπᾶς* ‘Geistlicher’ handelt (vgl. Miklosich 1876: 13; ERHSJ III: 8–9). Eine direkte Übernahme aus dem Gr. dürfte aber aus phonologischen (wie auch semant. Gründen, vgl. ERHSJ III: 9) auszuschließen und eine Entlehnung aus dem Westgerm. (ahd. *phaffo*) oder Got. (**papa*) anzunehmen sein (vgl. Pronk-Tiethoff 2012: 134). Lt. Auty (1976: 171) liegt ahd. *phaffo* zugrunde. Ebenfalls in pejor. Verwendung ist (kolloquiales) kr. *mantĭjāš* m. ‘Priester, Pfarrer, jmd. der ein Priesterkleid trägt’ (vgl. Anić 2003: 726), eine Ableitung von kr. *màntija* f. ‘Kutte, Priesterkleid’, einem Balkangräzismus des gr. Deminutivs *μαντίον* (vgl. ERHSJ II: 372).

²⁹ Bei sr. *svěštenik* handelt es sich hingegen um eine Entlehnung aus dem Aksl., da es *-št-* anstelle lautgesetzlich zu erwartendem *-ć-* < ursl. **tj* aufweist (vgl. a. Matasović 2008: 314) (vgl. a. Kap. 4.4).

Čak. *pr̥vad* ‘(höchster Rang eines) Diakon, Priester’ muss aufgrund seiner heutigen Lautung³⁰ relativ früh aus friul. **preved(e)* (< vlat. *praebiter* < lat. *presbyter* < gr. *πρεσβύτερος*) entlehnt worden sein (vgl. Skok 1953: 54–55, ERHSJ III: 61, Katičić 1980: 32). Daneben existiert jüngerer kr. *pr̥zbiter* m. (vgl. Anić 2003: 1182), eine Entlehnung aus klat. *presbyter* ‘Ältester, Priester’ < gr. *πρεσβύτερος* und direkt aus dem Gr. entlehntes sr. *pr̥zviter* (vgl. ERHSJ III: 61, s.v. *prvad*). Kr. (arch.) *podreka* ‘Oberhaupt der Kirche’ ist eine Entlehnung aus friul. *padriarca* (< lat. *patriarcha*), kommt jedoch heute nur noch in Anthroponymen vor. Das Vorhandensein der westrom. Sonorisierung deutet auf Herkunft aus Aquileia hin (vgl. Skok 1953: 54, Katičić 1980: 32). Kr. (dial.) *m̥sn̥ik* und kajk. *mašnik* ‘Priester’ sind Ableitungen aus dem kr. Latinismus *m̥sa* f. ‘kath. Messe, wichtigster Teil des Gottesdienstes’ bzw. dem Friulanismus kajk. *m̥ša* f. (vgl. ERHSJ II: 430–431) (vgl. 3.2.12).

Bei kr. *r̥ab̥m* m. ‘Rabbi’ bzw. ‘Lehrer; hebr. Priester; Herr’ handelt es sich um eine Entlehnung von hebr. *rabb* mit Postposition *-i* für die 1.P. *rabbi* ‘mein Lehrer, mein Herr’ (vgl. ERHSJ III: 91, s.v. *raban*¹, Anić 2003: 1272); kr. *h̥d̥ža* m. ‘Priester; Glaubenslehrer’ bzw. ‘muslim. Priester’ ist ein Balkanturzismus pers. Herkunft (pers. *haḡe* < *hāwaḡa*, vermittelt über türk. *koca* ‘Ältester, Lehrer’ bzw. Überkreuzung mit türk. *koç* ‘stark, mächtig, groß’) (vgl. ERHSJ I: 677–678, Anić 2003: 411); kr. *š̥am̥an* m. ‘Priester bei Tataren, Mongolen usw., Schamane’ ist eine über nhd. *Schamane* vermittelte Entlehnung aus dem Russ., welche ursprüngl. auf sanskr. *sehama* zurückgeht (vgl. Anić 2003: 1525, Anić/Klaić/Domović 2002: 1364).

Kr. *ž̥akan* (ERHSJ III: 670) bzw. *ž̥akan* (Anić 2003: 1857) m. ‘Diakon, Kleriker, niedrigster Priesterrang’ wurde aus dem Istr. oder Dalm. entlehnt³¹, wohin es über lat. *diāconus* < gr. *διάκονος* ‘Kirchendiener’ vermittelt wurde (vgl. ERHSJ III: 670, Holzer 2011: 165, Skok 1953: 56). Im später entlehnten gelehrten Latinismus kr. *d̥akon* ‘Diakon, Hilfspriester’ ist *d̥* erhalten und unterliegt den zu erwartenden lautgesetzlichen Entwicklungen; kr. *d̥ijakon* m. ‘Diakon, Hilfspriester’ hingegen wurde (wie kr. *d̥ijāk* ‘Schüler’, vgl. Anić 2003: 213) mittelbar über das Ung. (vgl. a. Kap. 4.4) entlehnt (vgl. ERHSJ III: 670, Anić 2003: 276, vgl. a. Miklosich 1876: 15). Kajk. *pr̥pošt* (*pr̥pošt*) m. ‘kath. Aufseher, Vorsteher’ ist lt. Anić (2003: 1169) ein über das Venez. oder Ital. vermittelter Latinismus, lt. Skok

³⁰ Lt. Skok (1953: 55, Anm. 26) geht der Sonant *r* auf eine friul. Form mit Metathese **pervede* zurück (zur im Friul. stattfindenden *r*-Metathese vgl. Marchetti 1985: 131, Skok 1953: 60). Zugrundeliegendes friul. **pervede* würde im Kr. (nach Durchlaufen der Zweiten Liquidametathese) allerdings (nach Anwendung der bei Holzer 2011 angeführten chronologischen Abfolge der slaw. bzw. kr. Lautgesetze) dasselbe lautliche Output liefern wie friul. **prevede* (auslautendes *-e* muss bei der Entlehnung jedenfalls noch erhalten gewesen sein, um die heutige Lautung *pr̥vad* zu erhalten). Die Entlehnung muss zudem nach dem slaw. Wandel *w > β* (da ansonsten rom. *β* mit slaw. *b* substituiert worden wäre, vgl. Holzer 2011: 34) und vor dem Wandel *i > ɔ* erfolgt sein, das Wort muss also unmittelbar nach der Entlehnung im Slaw. **prividi* gelautet haben (bei slaw. *i* für geschlossenes rom. *e* handelt es sich ebenfalls um eine Lautsubstitution, welche wohl auch hier durchgeführt wurde, vgl. hierzu a. Holzer 2011: 35).

³¹ Anić (2003: 1857) führt *žakan* als eine Entlehnung aus dem Gr. auf, doch lt. Skok (1953: 56, Anm. 33) ist der Wandel *d̥ > ž̥* wohl alpenromanischen Ursprungs; Holzer (2011: 165) zeichnet folgende Entwicklung nach: gr. *διάκονος* > lat. *diāconus* > vlat. **diākonu* > [...] rom. **dzākonu* > slaw. **džā kunu* > [...] kr. *žakan* (= *žakan*).

hingegen weist das Vorhandensein von *š* auf einen Magyarismus (< lat. *praepos(it)us* ‘Aufseher, Vorsteher’) hin (vgl. ERHSJ II: 42, s.v. kapitul). Dagegen ist kr. *prepòzīt* m. ‘Aufseher, Vorsteher, Prior’ (vgl. Anić 2003: 1169) eine direkte Entlehnung aus lat. *praepositus* (vgl. ERHSJ III: 7, s.v. -pònjat).

Kr. *vikār* m. ‘Stellvertreter eines Bischofs’ ist eine Entlehnung aus lat. *vicārius* ‘Stellvertreter’ (vgl. ERHSJ III: 592, Anić 2003: 1733); bei kr. *kapèlān* m. ‘(arch.) Priester, der sich um die Pfarre kümmert’, bzw. ‘Heeres-, Spitalspriester usw.; Hilfspriester’ (vgl. Anić 2003: 550) handelt es sich um eine Entlehnung aus klat. *capellanus* (vgl. klat. *cappella* ‘kleine Kirche’), im Kajk. existiert zudem über nhd. *Kaplan* entlehntes *kaplan* (vgl. ERHSJ II: 41, s.v. kàpela); kr. *kànonīk* m. ‘Domherr, Chorherr, Kanoniker’ ist lt. Anić (2003: 548) ein über das Dt. vermitteltler Latinismus (< lat. *canōn*), lt. Skok hingegen eine Ableitung aus kr. *kànon* m. ‘Kirchengesetz’, welches aus gr. *κανόν* entlehnt wurde (vgl. ERHSJ II: 34–35); kr. *kùstōd* m. ‘Küster’ ist eine Entlehnung aus lat. *custōdis*; kajk. *kustoš* m. ‘einer der Chorherren, Kanoniker’ hingegen geht (aufgrund des *š*, wie auch bei *plèbānuš* ‘Pfarrer’) auf über das Ung. vermittelte lat. *custōs* ‘Hüter, Wächter’ zurück (vgl. ERHSJ II: 246).

Bei kr. *pàstīr* m. ‘Hirte; Priester’ (vgl. Anić 2003: 1005) handelt es sich um ein Erbwort, welches von kr. *pàsti* ‘hüten, weiden’ (nachursl. **pasti* ‘Weide, Herde’, idg. Desiderativum **peh₂s-* der idg. Verbalwurzel **peh₂(i)-* ‘schützen, hüten, weiden’, vgl. Derksen 2008: 392, LIV: 460) mithilfe des Suffixes *-ir* zur Bildung von Nomina agentis abgeleitet wurde; in Urverwandtschaft steht lat. *pāscere* ‘hüten’, wovon mithilfe des lat. und panrom. Suffixes *-or* lat. *pāstor* ‘Hirte; Pfarrer’ abgeleitet und ins kr. als *pāstor* m. ‘protestantischer Priester, Pfarrer’ (vgl. Anić 2003: 1005) entlehnt wurde (vgl. ERHSJ I: 728, s.v. -ir, ERHSJ II: 614–615, s.v. pàsti¹).

Schließlich sind als weitere Synonyme³² für ‘Priester’ noch zu nennen: kr. *ispovjèdnik* m. ‘Beichtvater’ (vgl. Anić 2003: 468); eine Ableitung aus kr. *ìspovijed* f. ‘Beichte, Bekenntnis’ (vgl. ERHSJ III: 575, s.v. vēm) (vgl. a. 3.2.1); kr. *nadarbènik* m. ‘Priester’ (vgl. Anić 2003: 798), eine Ableitung von kr. *nadarbina* ‘Dargereichtes’, welches wiederum aus dem kr. Abstraktum *darba* von kr. *dàti* ‘geben’ abgeleitet wurde (nachursl. **dàti*, bsl. **do?*-, **do?d-mi*, vgl. Derksen 2008: 96; idg. **deh₃-* ‘geben’; reduplizierter Präsensstamm³³ **dé-doh₃-*/**dh₃-*, vgl. LIV: 105–106), genaugenommen handelt es sich um einen Calque von lat. *praebēnda* ‘Nahrungs-, Beköstigungsgeld’; kr. *prebèndār* m. ‘kath. Priester, der vom Ertrag des Landes oder einer Schenkung lebt’ stellt eine Entlehnung aus lat. *praebendarius* dar (vgl. ERHSJ I: 383 s.v. dāti, Anić 2003: 1148).

³² Šarić/Wittschen (2008: 455) führen außerdem *sluga Božji* ‘Diener Gottes’, *duhovni pastir* ‘geistiger Hirte’, *òtac* ‘Vater’, *velečasnī* ‘Hochwürden’, *kaluđer* ‘Mönch, Ordensmann’, *redòvnik* ‘Ordensangehöriger, Mönch’, aber auch *žrec* und *vrač* ‘Magier’, welche ursprüngl. auch ‘Priester’ bedeutet haben (vgl. 3.2.21), an.

³³ Wobei in bsl. Zeit **ded-* < **dé-dh₃-* zur Vermeidung von Homonymie mit **ded-* ‘setzen’ aus idg. **d^hé-d^hh₁-* analogisch zu **dód-* umgebildet wurde (vgl. LIV: 106).

Für ‘Pfarrer’ (vgl. für weitere slaw. Entsprechungen Miklosich 1876: 15) finden sich kr. *pàrōh* (ERHSJ III: 687) bzw. *pàroh* (Anić 2003: 1000) m. ‘orthodoxer Priester’ und kr. *parohijān* ‘ib.’ (vgl. Anić 2003: 1000) (< gr. *παροικία*, wohl mit Entwicklung von ‘Aufenthalt in einem fremden Land’ zur bei Skok angegebenen Bed. ‘Diözese’ bzw. ‘Diözese in einem fremden Land’) neben seltenerem *parok* m. (< klat. *parochia*) (vgl. ERHSJ II: 610); kr. *žùpnīk* (ERHSJ III: 687) bzw. *žûpnīk* m. (Anić 2003: 1871) ist eine Ableitung aus kr. *žúpa*³⁴ f. ‘Verwaltungseinheit, Gau’, auch ‘kirchl. Verwaltungsbezirk, Diözese’ (vgl. ERHSJ III: 687–688); kr. *plòvan* (vereinzelt auch *plovàn*, vgl. Anić 2003: 1051, ERHSJ II: 689) m. ‘Pfarrer, Priester’ ist eine Entlehnung aus friul. *plevan* < lat. *plebānus* (vgl. a. Katičić 1980: 32), daneben finden sich auch der Latinismus kr. *plèbān* m. und (über das Ung. vermittelte) kajk. *plèbānuš*³⁵ (*plebānuš*) m. (vgl. Anić 2003: 1046, ERHSJ II: 246, s.v. kustoš, ERHSJ II: 42, s.v. kapitul). Das Fehlen des toskanischen Wandels *pl-* > *pi-* in *plòvan* weist zwar auf das Friul. als Gebersprache (vgl. Skok 1953: 55), es kann sich aber auch um ein dalm. Wort handeln (vgl. ERHSJ II: 689); eine (lt. Miklosich 1876: 15) Entlehnung aus ital. *piovano* ist aus lautlichen Gründen auszuschließen.

3.2.11. ‘holy’

Das kr. Adj. *svēt*, aksl. *svěť* ‘heilig’ (vgl. Anić 2003: 1514) stellt lt. Skok eine Ableitung mithilfe des Suffixes *-to* aus der idg. Wurzel **kuen-* ‘feiern, heiligen’ dar (bsl. **šwentos-*, idg. **kuen-to-*, vgl. Derksen 2008: 476; ursl. **swentu*, vgl. Holzer 2011: 154) und dürfte mit idg. **keu-* ‘scheinen’ in Verbindung stehen (vgl. ERHSJ III: 371–372, vgl. a. Buck 1971: 1476). Demzufolge handelt es sich bei bsl. **šwentos-* wohl um den Präsensstamm **k̑ui-né/n-t-* der idg. Verbalwurzel **k̑uei-* ‘hell aufleuchten’ und bei aksl. *svěť*, kr. *svjjet* ‘Licht, Welt’, aksl. *světiti* ‘leuchten, erleuchten’ um das dazugehörige Kausativum **k̑uit-éh₁/h₁-* (vgl. LIV: 340, Derksen 2008: 476, ERHSJ III: 368–369, s.v. svànuti). Daneben existiert noch der Internationalismus kr. *sàkrālan* ‘sakral, heilig’, der lt. Anić (2003: 1364) aus mlat. *sacer* ‘heilig’ entlehnt wurde (vgl. a. Anić/Klaić/Domović 2002: 1263).

3.2.12. ‘to preach’

Bei kr. *propovijèdati*, aksl. *propovědati* ‘predigen, verkünden’ (vgl. a. aksl. *povědati* ‘verkünden, berichten’, ursprüngr. ‘bekannt machen’, vgl. Buck 1971: 1478, Miklosich 1876: 27, Anić 2003: 1235) dürfte es sich um einen Calque aus lat. *praedicāre* ‘verkünden’ (klat. ‘predigen’) handeln. Im Aksl. belegtes *věděti* ‘wissen’ ist im heutigen Kr. verschwunden und nur noch in Ableitungen wie *povijest*

³⁴ Es ist lt. Skok unklar, ob hier eine (vorursl.) Entlehnung aus dem Awar. (vgl. *kagan*, *bajan*, č. *pán* – dieses ist lt. Matasović 2008: 48 jedoch eine nur auf das Č. und Poln. begrenzte Entlehnung aus iran. **gu-pāna-* ‘Beaufsichtiger des Viehs’) oder aus gr. *δίοπος* ‘Aufseher, Chef’ oder gar ein idg. Erbwort vorliegt (vgl. ERHSJ III: 687–88, vgl. a. Matasović 2008: 48). Lt. Holzer (2011: 168–169) ist ursl. *džawpānu* > kr. *žûpān* ‘Dorfältester’ eine vorursl. Entlehnung aus dem Awar.

³⁵ Šarić/Wittschen (2008: 455, 560) weisen *plovān* und *pleban(uš)* als kolloquial aus, ersteres zudem als arch.

‘Geschichte’ oder *svjèdok* ‘Zeuge’ zu finden (vgl. ERHSJ III: 574–576; s.v. vēm). Aksl. *vědě* ‘ich weiß’ liegt idg. **uóǵd-h₂-ei* zugrunde, das auf den Perfektstamm **uóǵd-/*ǵid-* (bsl. **woi?d-*) der idg. Wurzel **ǵeid-* ‘erblicken’ zurückgeführt werden kann; von derselben Wurzel lässt sich kr. *vǵjeti* ‘sehen’ ableiten (vgl. LIV: 665–667, Derksen 2008: 518). Etymologisch hierher gehört auch vom kr. Abstraktum *vǵjest* ‘Kunde’ abgeleitetes kr. *navijéstiti*, *naviještati* ‘predigen, verkünden’ (vgl. Anić 2003: 834, Šarić/Wittschen 2008: 388).

Daneben ist kr. *predikati*, *predikovati*, *prodikovati* ‘predigen’ anzuführen (vgl. a. ERHSJ III: 40 s.v. pridika, Anić 2003: 1151), welches lt. Miklosich (1876: 28) wie nsorb. *prátkovaš* eine mittelbare Entlehnung (über ahd. *bredigōn*, *predigon* ‘predigen’, *prediga* ‘Predigt’) aus lat. *praedicāre* ‘predigen’ darstellt (lt. Skok allerdings direkt aus lat. *praedicāre*). Pejor. kr. *moralizirati* ‘predigen, moralisieren’ (vgl. Anić 2003: 776, Šarić/Wittschen 2008: 388) ist eine Ableitung von kr. *mòrāl* m. ‘Moral’, dieses wiederum ein Latinismus aus lat. *morālis* ‘moralisch’ (vgl. ERHSJ II: 455).

Für die ‘Predigt’ finden sich kr. *pròdika* f. (Anić 2003: 1217), *prèdika* f. (Anić 2003: 1151), wobei ersteres lt. Anić eine direkte Entlehnung aus lat. *praedicāre* ‘predigen’, letzteres jedoch eine mittelbare Entlehnung über das Ital. darstellt, und kr. *prǵdika* f. (vgl. ERHSJ III: 40, Anić 2003: 1186); aus kr. *propovijèdati* abgeleitetes kr. *pròpovijèd* f. (vgl. Anić 2003: 1235, ERHSJ III: 575 s.v. vēm) sowie (als Teil der Predigt) mittelbar über lat. *homīlia* ‘Rede; Predigt’ aus gr. *ὁμλία* ‘Rede’ entlehntes kr. *hòmīlija* f. (vgl. Anić 2003: 412, Šarić/Wittschen 2008: 388, Anić/Klaić/Domović 2002: 562).

Als Entsprechung für ‘Messe’ findet sich kr. *mīsa* f. ‘kath. Messe, wichtigster Teil des Gottesdienstes’ (vgl. Anić 2003: 758), lt. Skok ein später Latinismus (< substantiviertes Partizip Perfekt lat. *missa* (sc. *prēx*) < *missus* < *mittere* ‘schicken, senden’) aus dem 15. Jh.; bei sr. *mīsa* kann es sich lt. Skok auch um einen Gräzismus (< mgr. *μίσσα*) handeln. Daneben existieren kajk. *māša* f. und čak. *měša* f., welche noch die Entwicklung von (klassischem) lat. *ī* > vlat. *ę*, welches nachursl. als *i* adaptiert wurde > *ь* (wie in lt. Auty 1976: 172 nur in den Kiever Blättern bezeugtem *mьša*) > kr. (dial.) *a*, *e* (vgl. hierzu a. Holzer 2011: 23, 35) aufzeigen, so auch slow. *māša*, č. *mše*. Lt. Miklosich (1876: 27) handelt es sich bei ksl. *mьša*, kajk. *maša* und sr. *misa* um direkte Entlehnungen aus lat. *missa* oder mittelbare Entlehnungen über ahd. *mëssa*, *missa*; lt. Auty (1976: 172–173) eher um Entlehnungen aus dem Ahd., auch wenn in diesem Fall die Herkunft schwer zu entscheiden ist, denn „[...] die Wiedergabe eines orthographischen *s* als Zischlaut *š* würde der Phonologie des Althochdeutschen ebensogut entsprechen wie dem von deutschen (bairischen oder fränkischen) Priestern gesprochenen Latein.“ (Auty 1976: 172) Lt. Skok ist die Annahme einer Vermittlung ins West- und Südslaw. über das Ahd. überflüssig, wenn das Wort auch (zusammen mit anderen kirchl. Termini) aus dem Friul. Aquileias entlehnt worden sein kann (vgl. ERHSJ II: 430–431). Bei kr. *slūžba bǵžjā* neben *bogoslūžja* f. ‘Gottesdienst’ (‘kath. Messe’, vgl. Anić 2003: 1432, s.v. služba) könnte es sich m. E. um einen Calque des gr. (*θεία*) *λειτουργία* ‘öffentlicher Dienst,

Gottesdienst' (vgl. aber a. nhd. *Gottesdienst*) handeln. Kr. *bōžjā* stellt eine Ableitung aus kr. *bōg* 'Gott' (vgl. 3.2.2.) dar, *služba* f. 'Dienst' ist eine Ableitung mittels dem Suffix *-ba* aus kr. *sluga* m. 'Gefolgschaft, Dienerschaft' (vgl. Anić 2003: 1431), die Etymologie des bsl. und ursl. Abstraktums ist allerdings ungeklärt, lt. Skok besteht ev. ein Zusammenhang zu idg. **k̑lou-* mit seltener Suffixerweiterung *-ga* (vgl. ERHSJ III: 287–288), somit läge die idg. Verbalwurzel **k̑leu-* 'hören' (vgl. LIV: 334–335) zugrunde. Ein weiteres Synonym³⁶ für 'Gottesdienst, Messe' ist kr. *òbred* m. 'Gesamtheit der religiösen Tätigkeiten bzw. der Sakramente' bzw. 'Zeremonie als solche' (vgl. Anić 2003: 902, Šarić/Wittschen 2008: 46), eine Ableitung von kr. *réd* 'Reihe, Abfolge'; nachursl. **rêdъ* m. 'Reihe, Linie', bsl. **rind-* (vgl. Derksen 2008: 436), die idg. Wurzel ist ungesichert, lt. Skok kommt ev. idg. **ar-* und somit eine Urverwandtschaft zu lat. *òrdo* 'Reihe; Ordnung' in Frage (vgl. ERHSJ III: 122–123). Als Syntagma gehört auch noch kr. *obredni govor* 'Messrede' (vgl. Šarić/Wittschen 2008: 388) hierher.

Daneben finden sich noch kr. *litùrgija*, *lejturgija*, *letùrgija* f. 'Messe, Gottesdienst' (vgl. Anić 2003: 691, Šarić/Wittschen 2008: 46, 233), wobei *litùrgija* entweder mittelbar über klat. *liturgia* oder direkt (nach der ngr. Aussprache von *ει*) aus gr. *λειτουργία* 'öffentlicher Dienst, Gottesdienst' entlehnt wurde (vgl. ERHSJ II: 285); sowie čak. *ofcij* m., *ofcije* n. 'Gottesdienst' als Entlehnungen aus lat. *officium* 'Dienst', neben dem Italianismus kr. (dial.) *ofčije* n. und dem Venezianismus kr. (dial.) *ofciji* (vgl. ERHSJ II: 545).

3.2.13. 'to bless'

Bei kr. *blagoslòviti*³⁷, aksl. *blagosloviti* 'segnen' (vgl. Anić 2003: 89) handelt es sich um einen Calque von gr. *εὐλογέω* oder lat. *benedicere* 'Gutes über jmd. sprechen, loben, preisen' (gr. *εὐ*, lat. *bene*, aksl. *blagъ* 'gut' und *λόγος* 'Rede, Wort', *dīcere* 'sagen', *slovo* 'Wort', vgl. a. ERHSJ I: 167–168). Der erste Teil des Kompositums ist das kr. Adj. *blāg* 'süß, gut', nachursl. **bōlgъ* 'gut' (aksl. *blagъ* 'gut, gütig'),

³⁶ Deutlich spezialisiert sind die bei Šarić/Wittschen (2008: 46, 223) angeführten Synonyme für 'Gottesdienst, Messe' (neben kr. *večernjica*, *povečerje* 'Abendmesse', *po(l)nočka* 'Mitternachtsmesse', *bdjénje* 'Nachtwachen', *vigilija* 'ib.' und *crna misa* 'schwarze Messe') kr. *misa zadušnica* bzw. *zádušnica* f. 'Totenmesse' (vgl. Anić 2003: 1778) und kr. *pòdūšje* n. 'orthodoxe Totenmesse' (vgl. Anić 2003: 1070), beides Ableitungen von kr. *dúša* 'Seele' (vgl. ERHSJ I: 373, s.v. *dāh*), kr. *rèkvijem* m. 'Totenmesse', eine Entlehnung von lat. *requiēm* 'Totenmesse' aus lat. *requiēs* 'Ruhe, Rast, Erholung' (vgl. ERHSJ I: 324, s.v. *čil*, Anić 2003: 1325); kr. *paràstos* m. (mit *a > o*) 'orthodoxe Totenmesse' (vgl. Anić 2003: 997) und sr. *parastas* m. 'Requiem' sind Balkangräzismen aus gr. *παράστασις* in der Bedeutung 'Requiem' (vgl. ERHSJ II: 607); kr. *opijélo* n. 'orthodoxe Messe für verstorbene Heilige' (vgl. Anić 2003: 944) ist eine Ableitung von kr. *pjèvati* 'singen', nachursl. **pê-* aus ablautendem **poi-* der idg. Wurzel **peiH-* 'singen' (vgl. ERHSJ II: 671–672, LIV: 465). Kr. *euharistijsko slavlje* bzw. *euharistija* f. bedeutet 'Teil der Messe' bzw. 'Opfer von Wein und Brot an sich' (vgl. Anić 2003: 301) und ist ein Latinismus aus lat. *euharistija* 'Kommunion' < gr. *εὐχαριστία* (Abstraktum von *εὐχαριστέω* 'danken', ein Kompositum von *εὐ* 'gut' und *χάρις* 'Liebe') (vgl. ERHSJ I: 498, s.v. *evkaristija*). Kr. *pontifikāl* m. 'von einem Bischof durchgeführte Riten' ist ein Latinismus (< lat. *pontificālis* 'bischöflich, oberpriesterlich') (vgl. Anić 2003: 1100).

³⁷ Šarić/Wittschen (2008: 43) führen zudem (neben *davati blagoslov* 'Segen erteilen' und *prizvati Božju milost/zaštitu* 'Gottes Gnade erbitten') noch bereits in 3.2.8 angeführte kr. *hvaliti*, *slāviti*, *veličati*, *(po)štòvati*, *posvećivati* und *posvétiti* an.

ursl. **balgu* (vgl. Holzer et al. 2012: 7, Derksen 2008: 51–52), dessen Etymologie (wie auch jene von kr. *blâgo* ‘Wohlstand, Geld, Vieh, Gut’, nachursl. **bôlgo* n., ursl. **balga*, vgl. Holzer et al. 2012: 8) aber lt. Derksen ungeklärt ist. Lt. Skok (ERHSJ I: 166–167) ist aufgrund der ältesten Bed. ‘gut’ wahrscheinlich, dass die Wurzel **bol-* zugrundeliegt, welche sich auch im kr. Komparativ *bôljī* ‘besser’ und im Superlativ *nâjboljī* ‘am besten’ erhalten hat. Das Suffix bzw. der Formant *-gъ* in nachursl. **bolgъ* ist lt. Skok sehr selten, allerdings findet sich hier lt. Skok eine Parallele im avest. *bərəjaya* ‘jmd. willkommen heißen’ (aus idg. **bhelgh-*; lt. Rix jedoch zu idg. **b^herg^h-* ‘beachten, bewahren’, vgl. LIV: 79–78). Der Vergleich mit sanskr. *bhārgāh* ‘Heiligkeit’ (aus idg. **bhelg-*) legt lt. Skok nahe, dass auch kr. *blâg* die idg. Wurzel **bhel-* ‘scheinen’ zugrunde liegt, nachursl. **bolgъ* wäre somit ablautend (vgl. ERHSJ I: 167). Eine andere sowohl lautlich als auch semant. wahrscheinliche Möglichkeit stellt m. E. die Zuordnung zu einer schwundstufigen Form der idg. Verbalwurzel **b^hleg-/b^hleg-* ‘glänzen’ dar (vgl. LIV: 86–87). Beim zweiten Glied handelt es sich um ein denominales Verb von nachursl. **slōvo* n. ‘Wort’ (Deverbativum von *sluti* ‘heißen, berühmt sein’, vgl. aksl. *slovō, sluti*) (vgl. ERHSJ III: 288–289, s.v. *sluti*, Buck 1971: 1479–1480) (vgl. a. 3.2.8). Aus der idg. Wurzel **k^hley-* ‘hören’ (vgl. LIV: 334) (bzw. der Erweiterung **k^hley-os-*) leitet sich *ursl. **slawa* ‘Wort’, nachursl. **slōvo*, kr. *slōvo* (= *slōvo*) ‘Buchstabe’ ab (vgl. Holzer 2011: 149, Derksen 2008: 454, Anić 2003: 1430).

3.2.14. ‘to curse’

Kr. *pròklēti, pròklinjati* (vgl. Anić 2003: 1225), aksl. (*pro*)*klēti* ‘verfluchen, verwünschen’ und kr. *ùklēti* ‘verfluchen’ (vgl. Anić 2003: 1657) sind deverbale Ableitungen von kr. *klēti* (*se*), welches ursprüngl. (als Rechtsausdruck) ‘einen Eid ablegen, schwören’ bedeutet hat (vgl. aksl. *klēti se*), heute aber neben ‘beteuern’ primär ‘fluchen’ bedeutet und lt. Skok in Verbindung mit kr. *klōniti se* (vgl. 3.2.8) steht (vgl. ERHSJ II: 97–98, s.v. *klēti* (*se*), Buck 1971: 1481). Vom Präsensstamm derselben slaw. Wurzel (nachursl. **klēti, klbnō*) leitet sich lt. Skok auch kr. *kunuti, kunući* ‘fluchen’ ab (vgl. ERHSJ II: 97–98, s.v. *klēti* (*se*), Anić 2003: 1225 s.v. *prokleti*).

Als Synonym für kr. *proklētstvo* ‘Fluch’ (vgl. Anić 2003: 1225) findet sich aus gr. *ἀνάθεμα* ‘Weihgabe’ entlehntes kr. *anàtema, ànatema* f. ‘Fluch’, auch ‘Kirchenbann, Exkommunikation’ sowie denominales kr. *anatemizirati, anatemati, anatemisati* ‘verfluchen’ (vgl. Anić 2003: 20; ERHSJ I: 40). Šarić/Wittschen (2008: 385) führen zudem als Synonym³⁸ kr. *ùrok* m. ‘Verwünschung, Verzauberung’ (vgl. Anić 2003: 1681) an (vgl. a. 3.2.20), nachursl. **rekti* ‘sagen’, idg. **rek-* ‘ordnen, festlegen, bestimmen’ (vgl. Derksen 2008: 433, LIV: 506, vgl. a. ERHSJ III: 121, s.v. *rěci*).

³⁸ Neben *baciti anatemu/prokletstvo/urok (na koga)* ‘jmd. verfluchen’, *udariti prokletstvom* ‘jmd. mit einem Fluch belegen’; *zazvati zlo (na koga)* ‘Übel hervorrufen’ und *zazeljeti/poželjeti (komu) zlo* ‘jmd. Schlechtes wünschen’.

3.2.15. 'to fast'

Beim kr. Substantiv *pōst* m., aksl. **postъ* 'Fasten' bzw. 'Fastenzeit' (vgl. Anić 2003: 1115) und dem denominalen Verb *pōstiti* (vgl. Anić 2003: 1116), aksl. *postiti* 'fasten' handelt es sich um Ausdrücke, die in der gesamten Slavia vertreten sind. Es wird allgemein angenommen, dass es sich um Lehnwörter aus ahd. *fasta*, *fastēn*³⁹ (so auch Auty 1976: 171) aus der Zeit der Salzburger Mission (vgl. Kap. 4.2) handelt, welche selbst ev. aus ungesichertem lat. *fastus* entlehnt wurden (vgl. ERHSJ III: 14). Lt. Pronk-Tiethoff (2012: 135) können sowohl nachursl. **postъ* 'Fasten' als auch nachursl. **postiti se* 'fasten' aus dem Westgerm. oder dem Got. (*fastubni* n., (*ski*)*fastan*) entlehnt worden sein.

Čak. *žěžn* m. 'Fasten', *žěžināt* 'fasten' ist eine über das Dalm. oder Istr. (**dzędzūnu*) vermittelte Entlehnung aus lat. *iēiūnāre* bzw. *iēiūnus*⁴⁰ 'fasten' (vgl. Holzer 2011: 166, ERHSJ III: 679, Anić/Klaić/Domović 2002: 1503). Kr. *gōvjeti* 'abstimmen, einstellen' bedeutete ursprüngl. '*fasten' (wie heute lt. Skok auch noch im Russ. und Bulg.). Es findet sich in allen slaw. Sprachen außer dem Poln.; ursl. **gawē tēj* (vgl. Holzer et al. 2012: 24) liegt lt. Skok die idg. Wurzel **ghou-* wie in lat. *favere*⁴¹ 'gewillt sein' und ahd. *goumōn* 'erquicken, schmausen' zugrunde (vgl. ERHSJ I: 597). Lt. Derksen (2008: 181) liegt nachursl. **gověti* idg. **g^{uh}ou-* zugrunde, welches m. E. vielleicht in Verbindung zu idg. **g^{uh}ed^h-* 'bitten, wünschen' (vgl. aksl. *žędati* 'begehren, dürsten') stehen könnte (vgl. LIV: 217).

Mit kr. *kōrizma* neben (dial.) *korisma* f. (Anić 2003: 617 führt nur *korizma* an) wird die gesamte Fastenzeit, d. h. '40-tägiges Fasten' bezeichnet. Es handelt sich um eine mittelbare Entlehnung des lat. substantivierten Zahladjektivs *Quadragesima* f. (sc. *dies*) (> vlat., ital. *quaresima*; zu lat. *qu-* > rom. *k-* vgl. Holzer 2011: 22) '40 Tage vor Ostern', welches wiederum selbst einen Calque aus gr. *τεσσαρακοστή* (sc. *ἡμέρα*) '40. Tag (vor Ostern, als Beginn der Fastenzeit)' darstellt und durch Synekdoche auf die gesamte Dauer des Fastens übertragen wurde. Anhand der Lautung (Vorhandensein der westrom. Sonorisierung) ist ersichtlich, dass *korisma* dalm. Herkunft (Split; vgl. vgl. *koresma*, alb. *kreshmë*) ist, *korizma* hingegen aus dem nördl. Istr. oder Friul. (Aquileia) entlehnt wurde (vgl. ERHSJ II: 154–155, Katičić 1980: 31). Ein weiteres Synonym für '40-tägiges Fasten' und ein Calque des lat. *quadragesima* stellt kr. *četrdesètnica* f. (vgl. Anić 2003: 165) dar, es setzt sich aus kr. *čètiri* 'vier' (idg. **kstétur-*) und kr. *dēsēt* 'zehn' (idg. **dek^m-t-*) zusammen (vgl. Derksen 2008: 87, 100; ERHSJ I: 318, s.v. *četiri*, ERHSJ I: 394).

³⁹ Vgl. das ahd. Adj. *fasti* mit ursprüngl. Bed. 'fest, standhaft', mit Bedeutungsentwicklung über 'beibehalten, bewahren' wie got. *fastan* 'fasten; beibehalten'; die Bed. 'auf Essen verzichten' ist im Germ. und Slaw. sekundär (vgl. Buck 1971: 1483).

⁴⁰ Die Etymologie von lat. *iēiūnus*, ist ungeklärt, die Bed. dürfte aber ebenso 'fasten, hungrig sein', nur ohne religiöse Konnotation, gewesen sein; vgl. a. gr. *νηστεία*, *νηστέω* (von *νηστικός* 'nicht essen, fasten') (vgl. Buck 1971: 1483).

⁴¹ Lt. Rix geht lat. *fauēre* 'gewogen sein' jedoch auf das Kausativum-Iterativum **d^hou-éje-* der idg. Verbalwurzel **d^hey-* 'laufen, eilen' zurück (vgl. LIV: 147–148).

Šarić/Wittschen (2008: 354) führen zudem als weitere Synonyme an: kr. *němřs* m. ‘Tag, an dem bzw. Zeit, in der man kein Fleisch isst (im Gegensatz zum Fasten)’ (vgl. Anić 2003: 850), eine Ableitung von kr. *mřsiti* (*se*) ‘Fleisch essen = Fastenbrechen (als Gegenteil von ‘fasten’)’, deverbale *mřs* m. ‘Fleischspeise; Zeit, in der man Fleisch essen darf’; idg. Parallelen sind ungesichert, aksl. *mřsiti se* ‘entweihen’ weist lt. Skok auf eine Bedeutungsentwicklung von ursprüngl. ‘sündigen’ zu ‘Fastenbrechen’ hin (vgl. ERHSJ II: 470); kr. *umjerenōst* f. ‘Enthaltbarkeit’ (vgl. Anić 2003: 1666), eine Ableitung von kr. *mjěra* f. ‘Maß’ (vgl. ERHSJ II: 436–437), nachursl. **měra* f. ‘ib.’, eine *r*-Ableitung der idg. Verbalwurzel **meh₁-* ‘(ab)messen’ (vgl. Derksen 2008: 312, LIV: 424–425); sowie (kolloquiales) kr. *suzdržavanje* ‘Entziehung, Mäßigung’ bzw. *suzdržati* ‘zurückhalten’ (vgl. Anić 2003: 1510), eine Ableitung von kr. *držati* ‘halten’ (vgl. ERHSJ I: 448), nachursl. **dbržati* ‘ib.’, lt. Derksen ev. eine Erweiterung der idg. Verbalwurzel **d^her-* ‘befestigen, fixieren’ (vgl. Derksen 2008: 137–138, LIV: 145–146).

3.2.16. ‘heaven’

Kr. *něbo* n. (pl. *neběsa*⁴²) bedeutete ursprüngl. ‘Wolke, Firmament’ und erst als christl. Terminus ‘Himmelreich’ (vgl. ERHSJ II: 509; Anić 2003: 839). Idg. **neb^h-es-* (**neb^hos-* ‘Feuchtigkeit, Nebel, Wolke’ > ‘Himmel’) geht auf die idg. Verbalwurzel **neb^h-* ‘feucht werden, bewölkt werden’ (vgl. LIV 448) zurück und liegt bsl. **nebo, nebes*⁴³, ursl. **neba* ‘Himmel’, kr. *něbo* (= *něbo*); ursl. NAPI. **nebesā*, kr. *nebe.sa* (*neběsa*) zugrunde (vgl. Matasović 2008: 100, Holzer 2011: 132, Derksen 2008: 347–378). Lt. Miklosich (1876: 48) wird ‘Himmel, Himmelreich’ in allen slaw. Sprachen als *nebesa* (pl.) bezeichnet, der Sg. *nebo* sei in dieser Bed. seltener bzw. erst in jüngerer Zeit in Verwendung und bezeichne eher ‘Firmament’. Wenn dasselbe Wort für ‘Himmel, Firmament’ und ‘Himmelreich, Paradies’ verwendet wird (was oft der Fall ist, da fast alle europ. Wörter für ‘Himmel’ ursprüngl. das ‘Firmament’ bezeichnet haben), so findet oft eine Differenzierung mit anderen sprachlichen Mitteln statt: engl. *heaven* ‘Himmel’ vs. pl. *heavens* ‘Firmament’, umgekehrt lett. *debess* ‘Firmament’ vs. pl. *debesis* ‘Himmel’. Für ‘Himmelreich’ ist der Pl. im gr. NT (*οὐρανοί* pl. von *οὐρανός*) gebräuchlicher, so auch in der Vulgata (*caeli* pl. von *caelum*), im Ahd. (Tatian, *himila* pl. von *himil*) und Ksl. (*nebesa*) (vgl. Buck 1971: 1484, Miklosich 1876: 48).

⁴² Der arch. kr. NPI. *neběsa* ist ein Relikt der ursprüngl. *s*-Deklination (der Übergang zur *o*-Deklination hat im Kr. schon im 13. Jh. stattgefunden), daneben existiert auch ein NPI. nach der *o*-Dekl. kr. *něbā* (vgl. Matasović 2008: 205–206).

⁴³ Von der idg. Wurzel lassen sich sanskr. *nāphas-* ‘Nebel, Wolkenwand, Firmament’ und gr. *νέφος* ‘Wolke, Wolkenwand’ ableiten, von der Bsl. Wurzel lit. *debesis* ‘Wolke’ (vgl. Derksen 2008: 347–348).

Ein anderes Wort für ‘Himmelreich’ ist kr. *râj* m. ‘Paradies, Aufenthaltsort der Seligen’, aber auch ‘Garten Eden’⁴⁴, aksl. *rajъ* (vgl. Miklosich 1876: 48, Anić 2003: 1278). Als ursl. Erbwort (**rā̄ju* ‘Paradies’, vgl. Holzer et al. 2010: 53) findet es sich in der gesamten Slavia, idg. Entsprechungen sind (wie auch für nhd. *Himmel*) ungesichert und auch die ursprüngl. Bed. ist ungewiss; es dürfte sich aber um eine vorursl. Entlehnung aus einer iran. Sprache handeln (vgl. u. a. Pronk-Tiethoff 2012: 236). Ursl. **rā̄ju* muss allerdings schon im Ursl. die (Neben-)Bed. ‘angenehmer Aufenthaltsort, Insel der Seligen’ gehabt haben, da für aksl. *rajъ* in Übersetzungen die Bed. von gr. *παράδεισος* (in der Septuaginta ‘Paradies’) adaptiert wurde (wie dies auch bei slaw. *pakao* ‘*Pech’ für gr. *ἄδης* geschehen ist, vgl. 3.2.17) (vgl. ERHSJ II: 100). Daneben finden sich noch kr. *džènet* m. ‘Paradies, Himmelreich’ (ein Balkanturzismus aus türk. *cennet* < arab. *ġennet* ‘Garten, Paradies’; vgl. ERHSJ I: 472, Anić 2003: 274) und der Internationalismus kr. *paradis* m. (< gr. *παράδεισος*) (vgl. ERHSJ II: 100, Anić/Klaić/Domović 2002: 1033).

3.2.17. ‘hell’

Kr. *pàkao* m. (vgl. Anić 2003: 985), aksl. **pьklъ* (belegt als aksl. Adj. *pьklьnъ* ‘zur Unterwelt gehörig’) bedeutete ursprüngl. ‘Pech, Teer’ (heute in der Standardsprache wie in den Dialekten durch kr. *smòla* ‘Pech’ und *kàtran* ‘Teer’ ersetzt) und bekam erst durch das Christentum die Bed. ‘Hölle, Inferno’. Als idg. Wurzel führt Skok **poi-* ‘von Feuchtigkeit, Saft, Fett, Harz usw. strotzen’ an (vgl. ERHSJ II: 588). Es handelt sich somit wohl um eine *k*-Ableitung der idg. Verbalwurzel **pejH-* ‘anschwellen’ (vgl. LIV: 464–465), die bsl. Wurzel ist **pikil-/pikul-*, unverwandt sind lat. *pix*, gr. *πίσσα* ‘Pech’ (vgl. Derksen 2008: 426). Lt. Buck (1971: 1486) sind kr. *pakao*, č. *peklo*, poln. *piekło* und russ. (dial.) *peklo* ‘Hölle’ Entlehnungen aus mhd. *pech* ‘Pech’ (auch auf die ‘Hölle’ selbst übertragen). Es ist hier m. E. aufgrund der Tatsache, dass es sich bei kr. *pàkao* um ein slaw. Erbwort mit derselben (partiellen bzw. Grund-)Bed. wie mhd. *pech* handelt, allerdings fraglich, ob hier nicht in Anlehnung an das Dt. eine Bedeutungsübertragung bzw. -erweiterung und somit eine semant. Entlehnung stattgefunden hat und eine zusätzliche formale Entlehnung aus dem Mhd. nicht als überflüssige Annahme anzusehen ist. Lt. Miklosich (1876: 49) ist das Wort an sich ererbt, die Bed. ‘Hölle’ gehe aber vielleicht vom Ahd. aus.

Als Synonym⁴⁵ für *pakao* findet sich auch der kr. Gräzismus *pandèmonĭj* (neben *pandemônĭj*) ‘Sitz der Dämonen, Hölle’ < gr. *πανδαιμόνιον* ‘Gesamtheit aller Dämonen bzw. deren Aufenthaltsort’

⁴⁴ Für diese Teilbedeutung existiert auch das aus dem Hebr. kommende Synonym kr. *éden* ‘Garten Eden’, ‘Paradies’ (vgl. Anić 2003: 279), vgl. a. *edenski vrt* ‘Garten Eden’, *rajski vrt* ‘ib.’, *Elizej* ‘Ort der Seligen’, *Elizejske poljane* ‘Elysische Gefilde’ und *kraljevstvo nebesko* ‘himmlisches Königreich’ als weitere bei Šarić/Wittschen (2008: 399) angeführte Synonyme.

⁴⁵ Šarić/Wittschen (2008: 324) führen noch kr. *vječno prokletstvo* ‘ewige Verdammnis’ und *paklena vatra* ‘Höllenforte’ an.

(vgl. Anić 2003: 990). Daneben existiert noch aksl. *geena* < gr. *γέεννα*⁴⁶ (in der Vulgata als *gehenna*), welches im Kr. erneut als Balkanturzismus arab. Ursprungs (arab. *gehenem* < hebr. Kompositum *gē ben Hinnom* ‘Garten des Sohnes Hinnoms’, hebr. *ge-hinnom* ist der Name einer Schlucht, in der dem Moloch geopfert wurde) als kr. *dže(h)ènem* bzw. *džehènnem*, *džénem* m. ‘Hölle’ entlehnt wurde (vgl. ERHSJ: I: 471, Anić 2003: 273).

3.2.18. ‘demon’ (inkl. ‘devil’)

Die kr. Entsprechung für ‘Dämon, böser Geist’ ist am ehesten *zlòduh* m. (vgl. Anić 2003: 1841), ein Kompositum aus kr. *zlò* (aksl. *zъlъ*) ‘schlecht’ und *dùh* ‘Geist’ (vgl. 3.2.23) (vgl. ERHSJ III: 642, s.v. *zào*, vgl. a. Buck 1971: 1489). Aufgrund der bei Skok (ERHSJ III: 667) angeführten balt. und idg. Parallelen dürfte m. E. die idg. Verbalwurzel **ǵʰuer-* ‘krumm gehen’ (von der sich auch **ǵʰuér-* ‘wildes Tier’ ableiten lässt, vgl. a. kr. *zvìjer* ‘ib.’), mit einem *l*-Fortsatz (vgl. LIV: 182) zugrundeliegen, die bsl. Wurzel ist **zulʔ-* (vgl. Derksen 2008: 551).

Daneben findet sich kr. *démon* (ERHSJ I: 391) bzw. *dèmōn* m. (Anić 2003: 201) (aksl. *demonъ*, vgl. Miklosich 1876: 42) ‘Dämon, böser Geist; guter oder böser Genius; gefallener Engel, Teufel’, welches aus (k)lat. *daemon* < gr. *δαίμων* (‘göttliche Macht’ bei Homer, in späterer kirchl. Bed. ‘böser Geist, Dämon’, vgl. Buck 1971: 1488) entlehnt wurde und einen Internationalismus darstellt (vgl. ERHSJ I: 391). Als sowohl ‘guter’ als auch ‘böser Genius; Schutzgeist’ ist noch kr. *gènj* anzuführen, ein Latinismus aus lat. *genius* (vgl. Anić 2003: 345, ERHSJ I: 727 s.v. *inžinj*). Daneben findet sich auch mittelbar über das Türk. (*cin*) aus dem Arab. (*ǧinn* ‘Geist, Dämon’) entlehntes kr. *džīn* m. ‘Dämon; guter oder böser Geist’ (vgl. Anić 2003: 274–275, ERHSJ I: 474–475). Šarić/Wittschen (2008: 75, 551) führen zudem als weitere Synonyme für ‘Dämon, böser Geist’ (neben den untenstehenden Bezeichnungen für ‘Teufel’ sowie *zvijêr* f. ‘wildes Tier, Bestie’ und *vukòdlak* m. ‘Werwolf’) noch an: kr. *Levijàtan* m. ‘(bibl.) böser Geist’, eine bibl. übermittelte Entlehnung aus hebr. *liwjatan* ‘der sich Windende’ (vgl. Anić 2003: 680, Anić/Klaić/Domović 2002: 808); kr. *vàmpīr*, *upīr* m. ‘Vampir, Blutsauger’ (vgl. Anić 2003: 1714, Anić/Klaić/Domović 2002: 1458), ein ursl. folkloristischer Terminus, bei dem es sich lt. Skok ev. um eine Entlehnung aus nordtürk. *ubyr*, *ubyrly* ‘Hexe’ oder um ein Erbwort mit privativem *u* und der ungesicherten Wurzelform **pir* (wie in *lèpir*, *lèptīr* ‘Schmetterling’) von *per-* ‘fliegen’ handeln könnte (vgl. ERHSJ III: 564, ERHSJ II: 289, s.v. *lèpir*), vgl. idg. **(s)per-* ‘fliegen’ (vgl. LIV: 579); sowie kr. *nèmān* f. ‘großes, furchterregendes Phantasietier aus Märchen’ bzw. ‘Geist, Gespenst, Erscheinung;

⁴⁶ Ein weiteres aus dem Gr. entlehntes Synonym für ‘Hölle’ stellt sr. *ad*, aksl., russ. *adъ* < gr. *ἄδης* (‘Ort der verstorbenen Seelen’, mit ungeklärter Etymologie, ev. von **ἀ-φιδ-* ‘unsichtbar’ oder von **αἰφιδ-*) (vgl. Miklosich 1876: 49–50, Buck 1971: 1485–1486) dar, welches sich lt. Skok (ERHSJ I: 8) allerdings nur bei orthodoxen Slawen findet, Teil des ksl. Wortschatzes ist und nicht in die Volkssprache Eingang gefunden hat (es ist auch nicht bei Anić 2003 aufgeführt).

Teufel’, eine Ableitung mit Dissimilierung $m > n$ von kr. *mámiti* ‘anlocken, verlocken’, welches lt. Skok auf idg. **mā-* ‘Handbewegung, Gestikulation’ zurückgeht (vgl. ERHSJ II: 366, s.v. *mámiti*, vgl. Anić 2003: 848) und m. E. in Verbindung zur idg. Verbalwurzel **meh₂-* ‘Zeichen geben’ (vgl. LIV: 425) stehen könnte.

An dieser Stelle werden auch die kr. Entsprechungen für ‘Teufel’ angeführt, da Wörter für ‘Teufel’ häufig auch zur Bezeichnung der niederen Teufel oder Dämonen⁴⁷ verwendet werden.

Kr. *vrâg* m. ‘Teufel’ neben *vrâžica* f. ‘Teufelin’ (vgl. Anić 2003: 1755) bedeutete ursprüngl. wie im aksl. *vragъ* und russ. *vrag* ‘Feind’ (vgl. Miklosich 1876: 42, Buck 1971: 1487), die Bedeutungsübertragung war durch das Christentum bedingt. Durch Vgl. mit dem Balt. (aprs. *wargš* ‘böse’, lit. *var̃gas* ‘Elend, Not’) ist lt. Skok ersichtlich, dass es sich ursprüngl. um ein Adj. mit der Bed. ‘böse’ gehandelt haben muss (vgl. a. poln. *wrog* ‘böse’ neben ‘Teufel’). (Vgl. ERHSJ III: 617) Aus idg. **(h₁)-uorg-o-* lässt sich bsl. **worǰgós* (vgl. Derksen 2008: 527–528; Matasović 2008: 119; ev. besteht ein Zusammenhang mit idg. **h₂uerg-* ‘sich umdrehen, sich wenden’, vgl. LIV 290–291) und schließlich ursl. **wargu-*, kr. *vrâg* ableiten (vgl. Matasović 2008: 119). Kr. *ďāvao*, *ďāvō* m. ‘Teufel, Satan’ (vgl. Anić 2003: 276) ist eine Entlehnung aus gr. *διάβολος*⁴⁸ ‘Verleumder’ oder lat. *diabolus* (die zugrundeliegende Form mit $b > v$ kann entweder aus dem Vlat. oder Mgr. stammen) (vgl. ERHSJ I: 401–402, s.v. *dijavol*, Buck 1971: 1487). Lt. Miklosich (1876: 41) kann aksl. *dijavolъ*, *ďjajavolъ* aber auch mittelbar über ahd. *tiufal* entlehnt worden sein.

Bei kr. *satan* m., *sātana* ‘Teufel’ handelt es sich lt. Skok um eine erneute (gelehrte) Entlehnung von hebr. *sātān* ‘Feind’ über gr. *Σατᾶν*, *Σατανᾶς* (vgl. ERHSJ: 206–207, vgl. a. Miklosich 1876: 41). Daneben existieren kr. *sotōna* (*soto.na*), eine direkte Entlehnung aus gr. *σατανᾶς*, und kr. (dial.) *sòtona* (*so.tona*), eine mittelbare Entlehnung über lat. *satanas*. Bei beiden handelt es sich um frühe Entlehnungen, da sie den Wandel $a > o$ noch mitmachen (vgl. Holzer 2011: 150, Matasović 2008: 102). Für Anić (2003: 1377, 1449) liegen jedoch bei *satāna* f. und *sotōna* ‘böser Geist, Teufel’ direkte Entlehnungen aus dem Hebr. vor. Durch Vermittlung über bosn. Muslime findet sich im Bosn. zusätzlich der Balkanturzismus *šejtan* (ERHSJ III: 207) bzw. *šéjtān* ‘böser Geist, Dschinn; Teufel’ (Anić 2003: 1529) (< türk. *şeytan* < arab. *šaitan* < äthiop. *sajeṭān*, vgl. ERHSJ III: 207).

⁴⁷ „In the NT Grk. *διάβολος*, always sg. ‘the devil, Satan’, and *δαίμων* or *δαίμόνιον* ‘demon’ are kept apart, and this distinction was followed in the Vulgate. But in the English and German versions, from the earliest to the present, the words for ‘devil’ have been used also in the latter sense [...]“ (Buck 1971: 1488) So bedeutet got. *unhulpa* ‘Teufel’, aber *unhulpō* ‘Dämon’, ahd. *tiufal*, *diufal* (Diphthong wohl durch Assoziierung mit ahd. *tiuf* ‘tief’) sowohl ‘Teufel’ als auch ‘Dämon’. Ungebräuchlicher ist ahd. *unholdo* ‘Dämon, feindselig’ (wie got. *unhulpō*, von got. *hulps*, ahd. *hold* ‘gnädig, gütig’ mit Negativpräfix) (vgl. Buck 1971: 1487–1488), von welchem aksl. *neprijaznъ* ein Calque ist (vgl. Miklosich 1876: 42, Auty 1976: 173), vgl. a. kr. *nèprijatelj* m. ‘Feind, Gegner’ (Anić 2003: 858), arch. auch ‘Teufel’ (vgl. Šarić/Wittschen 2008: 523).

⁴⁸ Ableitung von *διαβάλλω* ‘auseinanderwerfen, herumwerfen, verleumden’ (von *βάλλω* ‘werfen’), welches selbst ein Calque von hebr. *sātān*, *sātān* ‘Widersacher’ ist (vgl. ERHSJ I: 402, Buck 1971: 1487).

Der čak. Archaismus **hūd** f., m. mit den Ableitungen **hudoba** (vgl. Anić 2003: 419), **hūdīc** (slow. *hudič*) ‘Teufel’ f. bedeutete ursprüngl. ‘dürr; ärmlich; schlecht’, die Bedeutungsübertragung ist durch die Christianisierung erfolgt (vgl. ERHSJ I: 692, Miklosich 1876: 42). Lt. Derksen (2008: 205–206) ist die ursprüngl. Bed. von nachursl. **hūdъ* ‘klein, dünn’ (ursl. **xaw`du* ‘schmächtig’, vgl. Holzer et al. 2012: 30) und es leitet sich von idg. **ksoud-ó-* ab (vgl. sanskr. *ksudrá-* ‘klein’). Lt. Rix (LIV: 372) leitet sich aksl. *hudъ* ‘klein, gering’ allerdings vom Präsensstamm **k^(u)seyud-e-* (nicht vom Kausativum) der zugrundeliegenden idg. Verbalwurzel **k^(u)seyud-* ‘zerstieben, in kleine Teile zerfallen’ ab.

Bei kr. **zlōca** f. ‘Teufel’ (vgl. Anić 2003: 1841) handelt es sich um eine Ableitung aus dem kr. Adj. *zlō* ‘schlecht’ (s.o.), vgl. aksl. *zъloděj* (vgl. ERHSJ II: 642, s.v. *zào*, Miklosich 1876: 42); kr. **paklènica** f. ‘Teufel, Teufelin’ und kr. **pàklenjāk** m. ‘Teufel; Verurteilter in der Hölle’ sind Ableitungen von kr. *pàkao* ‘Hölle’ (vgl. 3.2.17) (vgl. ERHSJ II: 588; vgl. a. Anić 2003: 985); sr. **kūsī** m. ‘Teufel’ (vgl. Šarić/Wittschen 2008: 523, bei Anić 2003 nicht vorhanden) ist eine Substantivierung des kr. Adj. *kūs* ‘Stück, Bissen’ (vgl. ERHSJ II: 245–246), nachursl. **kōsъ* ‘Stücken, Bissen’, bsl. **kanъd-*, idg. **k^(u)ond-* oder **k^(u)onHd^(h)-* mit Suffix *-*so-* (vgl. Derksen 2008: 244, s.v. *koṣati*).

Kr. **àntikrist** m. ‘Antichrist, Teufel’ (vgl. Anić 2003: 25) stellt einen wohl über lat. *antichristus* vermittelten Internationalismus (gr. *àvri* ‘gegen’ und gr. *χριστός* ‘Christus; der Gesalbte’) dar (vgl. ERHSJ I: 46, s.v. *anti-*, ERHSJ II: 197 s.v. *krist*, Anić/Klaić/Domović 2002: 91). Kr. **Màmon** m. ‘(bibl.) Satan, Widersacher, böser Gott (dualistisch)’ bzw. ‘einer der Teufel’, aber auch ‘bibl. Personifizierung Gottes’, ist eine aus dem Lat. durch das Gr. und Hebr. (wohl bibl.) vermittelte Entlehnung aus aram. *mammōn* oder *māmōna* ‘Gut, Reichtum, Geld, Ertrag’ (vgl. Anić 2003: 723, ERHSJ II: 366). Kr. **bělāj** m. bedeutet neben ‘Unglück, Leid’ auch ‘Teufel’ (vgl. Anić 2003: 67), es handelt sich um einen Balkanturzismus (*bela*) arab. Ursprungs (*belā* ‘Unglück, Übel, Elend’) (vgl. ERHSJ I: 135). Kr. **belzèbub** m. ‘Teufel’ ist eine bibl. vermittelte Entlehnung (über klat. *Beelzebūb* < gr. *Βεελζεβούλ*) aus hebr. *Ba`al zewūl* ‘Herr des Himmels’ oder *Ba`al zewūw*, wörtl. ‘Herr der Fliegen’⁴⁹ (vgl. Anić 2003: 67, Anić/Klaić/Domović 2002: 170) und dürfte sich lt. Skok auch im kr. Syntagma *Vrziño kòlo* ‘Tanz der Teufel’ (verbunden mit dem Gebirgszug Velebit), sowie im Toponym *Vrzići* als aksl. vermitteltes *Verziul*, *Verzil* < *Belzebub*, *Belzevul* finden (vgl. ERHSJ III: 633–634, s.v. *Vřzin*). Weitere Synonyme (vgl. a. Šarić/Wittschen 2008: 135, 523) bzw. Appositionen oder eigentlich Euphemismen aufgrund des Tabus der Bezeichnungen für ‘Teufel’ (vgl. u. a. ERHSJ I: 135, s.v. *bělāj*) sind kr. **nàpasnik** m., **nàpast** f. ‘Versucher, Teufel’ (vgl. Anić 2003: 818), letzteres ist ein schon ursl. gebildetes Abstraktum von kr. *pàsti* ‘fallen’, nachursl. **pàsti* ‘ib.’, idg. **pōd-* (*ō*-Stufe zu idg. Verbalwurzel **ped-* ‘treten; fallen, sinken’, vgl.

⁴⁹ Vgl. zu den Deutungsversuchen die Ausführungen im Digitalen Wörterbuch der deutschen Sprache (DWDS) der Berlin-Brandenburgischen Akademie der Wissenschaften unter <<http://www.dwds.de/?qu=Beelzebub>> [aufgerufen am 27.7.2015].

LIV: 458) (vgl. ERHSJ II: 615–616, s.v. pàsti, Derksen 2008: 392); etymologisch gehört auch kr. *pali andeo* m. ‘gefallener Engel’ hierher (vgl. Anić 2003: 1755, s.v. vrag) (vgl. 3.2.23 zu *andeo*). Kr. *nečastivī* m. ‘Teufel’, eigentlich ‘Ehrloser’ (vgl. Anić 2003: 840) ist eine Ableitung von kr. *čâst* f. ‘Ehre’, dieses wiederum ein Abstraktum des nachursl. Präsensstamms *čbtq, čisti* (vgl. a. kr. *štòvati* in 3.2.8); bei kr. *nečastivnik* m. ‘Teufel’ handelt es sich ebenfalls um einen Euphemismus für *vrâg* (vgl. ERHSJ I: 298). Beim kr. Adj. *rèpat* ‘mit einem langen Schwanz oder Schweif versehen’ (vgl. Anić 2003: 1329) handelt es sich um eine Ableitung von kr. *rêp* ‘Schwanz, Schweif’, nachursl. *rëpb*, die idg. Wurzel und idg. Kognaten sind ungesichert (vgl. ERHSJ III: 128–129, s.v. rêp). Bei kr. *čfnī* m. ‘Teufel’, eigentlich ‘Schwarzer’ (vgl. Anić 2003: 148), handelt es sich ebenfalls um einen Zusatz zu Personennamen (vgl. ERHSJ I: 277–278), es ist lt. Derksen (2008: 92) auf nachursl. *čbrnъ ‘schwarz’, bsl. *kirsnos, idg. *krs-no- zurückzuführen.

3.2.19. ‘idol’

Kr. *kùmīr* m. ‘Götzenstatue, Idol’, aksl. *kumirъ, kumirъ* ‘Statue’ (vgl. Miklosich 1876: 36) war lange Zeit auf das Ksl. beschränkt, wurde jedoch im 19. Jh. „[...] od nekih purista i ljubitelja neologizama [...]“ (ERHSJ II: 233) wieder aufgegriffen und so auch in der Volkssprache etabliert. Kajk. *kumer* ‘schmutzig’ (semant. Überkreuzung mit *pogan* ‘heidn.; schmutzig’) und slow. *kúmer* ‘Magerkeit, dürrer Mensch’ zeigen lt. Skok allerdings, dass das Wort auch davor in der Volkssprache bestanden haben muss. Lt. Anić (2003: 654) handelt es sich um einen ursprüngl. aus dem Aksl. stammenden Russismus, wie das Wort ins Aksl. gekommen ist, bleibt allerdings weiterhin unklar; ein Zusammenhang mit syr. *kumrā* ‘Priester’ und arm. *k’urm*, pl. *k’rmac* ‘heidn. Priester’ ist lt. Skok nicht auszuschließen (vgl. ERHSJ II: 233, vgl. a. Buck 1971: 1492).

Bei kr. *kīp* m. (vgl. a. russ.-ksl. *kypъ*) ‘Gestalt, Form, Bild, Idol, Götze’ handelt es sich um eine Entlehnung aus ung. *kép* ‘Bild’, welches wiederum aus dem Atürk. entlehnt wurde. Parallelen zu *kīp* finden sich im jakutischen und uigurischen *kep*, im mongolischen und südtürk. *kap* ‘Form, Bild’ (vgl. ERHSJ II: 38, s.v. kapъ, Miklosich 1876: 36, Anić 2003: 570).

Weitere Synonyme für ‘Götze, Idol’ (vgl. Šarić/Wittschen 2008: 138, 180) sind (neben *lažni bog* ‘falscher Gott, Abgott’) kr. *fētīš* m. ‘verehrtes heiliges Objekt mit übernatürlichen Kräften’, eine mittelbare Entlehnung über frz. *fétiche* aus port. *feitico* ‘Zauberei, Hexerei’ (< lat. *facitius* ‘künstlich’) (vgl. Anić 2003: 313, Anić/Klaić/Domović 2002: 437); kr. *simulâkr(um)*, *simulâkrum* m. ‘Götze, Statue’, ‘Erscheinung, Geist’, eine Entlehnung aus lat. *simulacrum* (vgl. Anić 2003: 1400), kr. *poklónac* m. ‘(verehrte) Statue’ (vgl. Anić 2003: 1081), wohl eine Ableitung von kr. *poklòniti* ‘darbringen’ (vgl. ERHSJ II: 102–103, s.v. klòniti se), kr. *tòtēm* m. ‘verehrtes Totemtier bzw. dessen Symbol’, lt. Anić (2003: 1607) eine mittelbare Entlehnung über amerikanisches Engl. aus einer „exotischen“ Sprache

(womit wohl eine indigene amerikanische Sprache gemeint ist) (vgl. Anić 2003: 1607, Anić/Klaić/Domović 2002: 1418). Kr. *īdol* m. (ERHSJ I: 709) bzw. *idōl* (Anić 2003: 426), aksl. *idolъ* ‘als Gottheit verehrte Statue, Götzenbild’ ist ein gelehrter Latinismus (lat. *īdolum* < gr. *εἶδωλον*⁵⁰ ‘Bild’) (vgl. ERHSJ I: 709).

3.2.20. ‘magic’

Als Entsprechungen⁵¹ für ‘Magie, Zauberei’ sind das kr. Verb *bājati* (ERHSJ I: 93) bzw. *bájati* (Anić 2003: 51) ‘zaubern, Magie ausüben, Geister beschwören’ und das daraus abgeleitete kr. Substantiv *bajanje* n. ‘Zauberei’ anzuführen (heute allerdings auch oft pejor. als ‘Falschmeldung, Gerücht’, vgl. ERHSJ I: 93). Die ursprüngl. Bed. der ursl. mythologischen Termini entspricht jener von urverwandtem gr. *φημί* ‘sprechen’ (vgl. a. russ.-ksl. *bajati* ‘Märchen erzählen’). Eine Verbindung zu ung. *báj* ‘Magie’ < türk. **bagy*⁵² ‘Bindung, Bund’ ist fraglich. Ursl. **bā`jā`tēj* ‘rezitieren u. dgl.’ (vgl. Holzer et al. 2010: 40) liegt die idg. Verbalwurzel **b^heh₂-* zugrunde, welche sich auch noch in lat. *fārī* ‘sagen’ findet (vgl. ERHSJ I: 93, Derksen 2008: 33–34, LIV: 69). Kr. *bāhoriti* ‘zaubern’ (vgl. Anić 2003: 51) und *bahorenje* n. ‘Zauberei’ stehen in etymologischer Verbindung zu kr. *bājati* und sind wie dieses Termini aus der ursl. Mythologie (vgl. ERHSJ I: 92).

Kr. *vráčati* ‘Zaubersprüche hersagen’ und *vračanje* n., *vrādzbina*, *vračarija* f., *vrāštvo* n. ‘Zauberei’ sind Ableitungen von kr. *vráč* ‘Wahrsager, Magier’ (vgl. 3.2.21); lt. Skok könnte die zugrundeliegende idg. Wurzel **uer-* ‘sprechen, sagen’ in Verbindung zu idg. **uorgo-*, aus der sich auch kr. *vrâg* ‘Teufel’ ableitet (vgl. 3.2.18), stehen (vgl. ERHSJ III: 616, 617, Anić 2003: 1755, 1757).

Bei kr. *čâr* f. (ERHSJ I: 295; lt. Anić 2003: 156 aber m.), aksl. *čarъ* (pl.) ‘Zauber’, kr. *čarati* ‘zaubern, Zaubersprüche hersagen’ handelt es sich um den lt. Skok ältesten slaw. Terminus aus dem Bereich der Magie mit ursprüngl. Bed. ‘fertigen, tun’. Relevante, heute noch verwendete Ableitungen sind kr. *čaròlija*, *čaròbija*, *čârnost*, *čârnošt* f. ‘Zauberei, Hexerei’ und das Deminutivum *čârka* f. ‘Zauber’ (vgl. Anić 2003: 157). Hierbei fand eine Bedeutungsspezialisierung wie in sanskr. *krtyā-* von ‘machen, tun’ zu ‘Handlung’, speziell ‘Magie, Hexerei’ statt (vgl. ERHSJ I: 295, Buck 1971: 1496, Anić 2003: 157). Aus der idg. Verbalwurzel **k^her-/k^hēr-o-* ‘(ab)schneiden, schnitzen’ > ‘herstellen, machen’

⁵⁰ Von *εἶδος* ‘Figur’ abgeleitet; vgl. a. den seltenen Calque *obličenje* im Aksl. (vgl. Miklosich 1876: 36, Buck 1971: 1492).

⁵¹ Als weitere Synonyme für ‘Zauberei’ führen Šarić/Wittschen (2008: 62, 113, 222, 523) noch die Spezifizierungen *crna magija* ‘schwarze Magie’ und *bijela magija* ‘weiße Magie’, sowie *proricati budućnost/sudbinu* ‘Zukunft/Schicksal voraussagen’, *čitanje/gledanje sudbine/budućnosti* ‘Zukunft/Schicksal lesen/sehen’, *čitati/gledati sudbinu* ‘ib.’, *bacanje uroka* ‘verwünschen’, *gledati u kavu/dlan* ‘aus dem Kaffeesatz oder der Hand lesen’, *gledati u grah* ‘Bohnenlegung’, *gledati u kavenu šalicu* ‘aus dem Kaffeesatz lesen’, *bacati karte* ‘Karten legen’, *okrenuti karte* ‘ib.’, *gledati u karte* ‘Karten lesen’ sowie *magična formula* ‘magische Formel’ und *vudu* ‘Voodoo’ an.

⁵² Ein Verbindungsglied könnte m. E. das arm. Erbwort *bay* ‘sagt’ (vgl. a. LIV: 69) liefern, sofern das türk. Wort wiederum ein Lehnwort aus dem Arm. ist.

(vgl. a. sanskr. *κρῆνóτι* ‘machen’) leitet sich bsl. **kēr-/ker-* und in weiterer Folge ursl. **čērá-*, kr. *čâr* ab (vgl. LIV: 391, Matasović 2008: 135, Derksen 2008: 78–79). Aufgrund der Tabuisierung von kr. *čâr* wurde anstelle dessen *copràti*, *coprija*, *còparnjica* (s.u.) entlehnt bzw. der Ausdruck mit Euphemismen ersetzt (vgl. ERHSJ I: 295, s.v. *čâr*), so auch mit kr. *čîni* f. pl. bzw. m. ‘Zauberei, Verwünschung, Zaubersprüche hersagen’ (vgl. Anić 2003: 168), welches im Sg. *čîn* auch ‘Handlung’ neben ‘Form, Art’ und ‘Reihe, Rang’ bedeutet (vgl. ERHSJ I: 325–326); nachursl. **čînъ* m. liegt idg. **k^hei-no-* (vgl. Derksen 2008: 89) bzw. **k^hei-nu-* zugrunde (vom Präsensstamm der idg. Verbalwurzel **k^hej-* ‘sammeln, schichten’, vgl. LIV: 378–379). Weitere kr. Ableitungen sind *opčinjavanje* n. ‘Zauberei’, *opčinjavati* ‘bezaubern’ und *opčiniti* ‘verzaubern’ (vgl. Anić 2003: 942, ERHSJ I: 326).

Kr. *gátati* ‘zaubern; vorhersagen’ kommt ebenfalls aus dem Bereich der ursl. Mythologie. Das Wort wird in den übrigen slaw. Sprachen entweder basierend auf *gat-* (bulg. *gátam se*, rus. *gatatъ*) oder *gad-* (slow. *gâdati*, č. *hâdati*, rus. *gadatъ*, bulg. *gâdkam* ‘sprechen’) weitergeführt, wobei ungeklärt ist, ob es sich um zwei separate oder um ein und dieselbe Wurzel handelt (vgl. ERHSJ I: 555, vgl. a. Anić 2003: 342). Es besteht lt. Derksen (2008: 160) auch die Möglichkeit, dass bei nachursl. **gâdàti* ‘raten’ die idg. Wurzeln **g^het-* (nachursl. **gatàti*, kr. *gátati* ‘wahrsagen, raten’) und **g^hōd-/g^hed-* ‘fassen’ überkreuzt wurden oder dass eine Verbindung zu idg. **g^hed^h-* ‘(zusammen)treffen, sich vereinigen’ (Iterativ **g^hod^h-éje-*, bsl. **gad-*, nachursl. **goditi*, kr. *gòditi* ‘gefallen’) vorliegt (vgl. Derksen 2008: 160–161, 172, vgl. a. LIV: 194, 195).

Bei kr. *màđija* f., *màgija* ‘Magie, Zauberei’ handelt es sich lt. Skok um einen Internationalismus aus gr. *μαγεία*⁵³ oder lat. *magia* (vgl. ERHSJ II: 350, Anić 2003: 713, Buck 1971: 1496); lt. Anić (2003: 712; s.v. *mađioničar*) wurde das Wort mittelbar über das Dt. entlehnt (nhd. *Magie* < frz. < lat. < gr.). Kr. *còprati* ‘zaubern’ und *còprija* (*coprija*) f. ‘Zauberei (oft pejor.)’ (vgl. Anić 2003: 146) sind alte Germanismen aus ahd. *zoubar* (nhd. *Zauber*) (vgl. ERHSJ I: 272, s.v. *còparica*).

Kr. *okultizam* m. ‘Okkultismus’ (vgl. Anić 2003: 933) ist eine Entlehnung aus lat. *occultus* ‘geheim, verborgen’ (vgl. Anić/Klaić/Domović 2002: 981). Kr. *negromàntija*, *negromàncija* (auch *nekromantija*) f. ‘Zauberei, Wahrsagerei’ (vgl. Anić 2003: 844, Anić/Klaić/Domović 2002: 943) ist eine Ableitung von kr. *negromanat* ‘Zauberer’, dieses wiederum ein über mlat. *necromantes* vermittelter Gräzismus (vgl. ERHSJ I: 599, s.v. *grabancija*). Bei Kr. *spiritizam* m. ‘Spiritismus; Glaube an Kommunikation mit Geistern oder übernatürlichen Wesen’ (vgl. Anić 2003: 1453, Anić/Klaić/Domović 2002: 1331) handelt es sich wohl um den Latinismus *spiritismus* (< lat. *spīritus* ‘Geist’, vgl. ERHSJ III: 410, s.v. *špirit*). In eher pejor. Verwendung sind kr. *opsjenárstvo* n. ‘Scharlatanerie’ und kr. *òpsjena* f.

⁵³ Gr. *μαγεία*, *μαγευτική* (sc. *τέχνη*), *μαγική τέχνη* sind Derivate von gr. *Μάγοι*, ursprüngl. der Name eines medizinischen Stammes, später als Bezeichnung der pers. ‘Priester’, apers. NSg. *maguš*, welche im gr. Volksglauben mit magischen, weil unbekanntem Riten assoziiert wurden. Von gr. *μαγικός* stammen lat. *magicus*, *magica ars* und *magicē* (aus gr. *μαγική τέχνη*) und späteres *magia* (aus gr. *μαγεία*) (vgl. Buck 1971: 1494–1496).

‘Täuschung’ (vgl. Anić 2003: 949), Ableitungen aus kr. *sjèn* ‘Schatten’ (vgl. ERHSJ III: 250–251), nachursl. **sěnb* f. ‘Schatten’, idg. **skoHi-n(-i)-* (vgl. Derksen 2008: 447), hier liegt die idg. Verbalwurzel **skēH(i)-* ‘schimmern, scheinen’ (vgl. LIV: 546) zugrunde.

Kr. *ùrok* m. (*ùroci* pl.) ‘Zauberei, Verwünschung’ (vgl. Anić 2003: 1681), aksl. *urokъ* ‘Absicht’ ist eine Ableitung von kr. *rěci* ‘sagen’, nachursl. **rekti* ‘ib.’, idg. **rek-* ‘(an)ordnen, festlegen, befehlen’ (vgl. ERHSJ III: 120–122, Derksen 2008: 433, LIV: 506). Hierher gehören auch die kr. Ableitungen *pròreći*, *proricati* ‘weissagen, orakeln’ und *proricanje* n. ‘Wahrsagerei’ (vgl. Anić 2003: 1236, 1237). Bei kr. *zazivanje* n. ‘Beschwörung’ und *zazívati* ‘beschwören’ (vgl. Anić 2003: 1826) handelt es sich um Ableitungen von kr. *zvàti* ‘rufen’ (vgl. ERHSJ III: 665–666), nachursl. **zъvati* ‘beschwören, anrufen’, idg. Präsensstamm **ǵhuH-e-/ *ǵ^houH-e-* von **ǵ^hueH-* ‘rufen’ (vgl. Derksen 2008: 551, LIV: 180–181). Als Entsprechung für ‘Magie’ im weitesten Sinne bzw. für ‘durch übernatürliche Kräfte Hervorgerufenes’ kann hier noch kr. *čùdo* n. ‘Wunder’ (vgl. Anić 2003: 174) angeführt werden, eine Ableitung von kr. *čùti* ‘hören, wahrnehmen’ (vgl. ERHSJ I: 339–340), welches auf die idg. Verbalwurzel *(s)*keuh-* ‘wahrnehmen, schauen’ zurückgeht (vgl. LIV: 561), nachursl. **čùdo* n. ‘Wunder’ liegt lt. Derksen (2008: 91) idg. **keud-os-* zugrunde.

3.2.21. ‘sorcerer or witch’

Kr. Entsprechungen für ‘Zauberer’ und ‘Zauberin’ (vgl. a. 3.2.20) sind *màdionĭk*, *mađioničar*, *magičar*, *mâg* (letzteres auch mit der eigentlichen Bed. ‘Weiser’) m., *mađionica*, *mađioničarka* f. (Ableitungen aus kr. *màđija* < gr. *μαγεία* oder lat. *magia*, vgl. ERHSJ II: 350, Anić 2003: 712), mit kr. *vračati* in Verbindung stehendes kr. *vrâč* m. ‘Wahrsager, Magier’ – auch mit den Nebenbed. ‘Arzt’ (wie im Slow. und Russ., diese Nebenbed. hat sich auch im kajk. Verb *vrāčiti* ‘heilen’ erhalten, vgl. ERHSJ I: 295, Anić 2003: 1755) und ‘Priester, jmd. der sich um kultische Riten kümmert’ – neben *vračar* m. ‘Wahrsager, Magier’ und *vrâčara*, *vračàrica* f. ‘Wahrsagerin, Magierin’ (vgl. ERHSJ III: 616, Anić 2003: 1755), sowie kr. *gàtara*, *gàtalica* f. ‘Wahrsagerin’, *gàtalac* m. ‘Wahrsager’ aus kr. *gátati* ‘zaubern, wahrsagen’ (vgl. ERHSJ I: 555, Anić 2003: 342). Ebenso existieren die aus kr. *čâr* ‘Zauber’ abgeleiteten Substantive kr. *čaròbnĭk*, *čàrobnĭk* m., *čàrobnica* f. (vgl. ERHSJ I: 295, Anić 2003: 157, Buck 1971: 1499); bei kr. *čaratār* m. ‘Zauberer’, *čaratàrica* f. ‘Zauberin’ handelt es sich um eine Entlehnung von ital. *cerretano* ‘Schwindler, Betrüger’, wobei lt. Skok ev. eine Überkreuzung mit ital. *ciarlatano* > frz. *charlatan* > kr. *šarlàtan* m. ‘Scharlatan, Betrüger’ stattgefunden hat (vgl. ERHSJ I: 253, Anić 2003: 1527).

Zudem findet sich im Kr. der (allerdings nicht wie *žrtva* ‘Opfer’ durch das Christentum vermittelte) Russismus *žréc* m. ‘Schamane, Zauberer’ (vgl. Anić 2003: 1870, ERHSJ III: 685, s.v. *žrtva*). Das nachursl. Agentivum **žbrъcb* m. ‘Priester’, aksl. *žrъcb*, stellt eine Ableitung aus nachursl. **žrti* ‘opfern’ (vgl. 3.2.7) dar und findet sich sonst nur im russ. und bulg. *žrec*, hier allerdings in der Bed.

‘(heidn.) Priester’ (vgl. Derksen 2008: 567). Bei kr. **vilènjāk** m. ‘Hexer, Zauberer’ (vgl. Anić 2003: 1734) handelt es sich um eine Ableitung von kr. *vīla* ‘Fee’ (vgl. 3.2.22) (vgl. ERHSJ III: 593).

Kr. **iluziònist** m. ‘Magier’, **iluziònistkinja** f. ‘Magierin’ sind Ableitungen aus dem kr. Latinismus *ilūzija* f. ‘Erscheinung, Illusion’ (< lat. *illūsiō*) (vgl. Anić 2003: 430). Kr. **šāmān** m. bedeutet neben ‘Schamane (Priester mancher Völker wie Tataren und Mongolen)’ auch im pejor. Sinne ‘jmd. der Macht über andere besitzt; Spiritist’, und wurde lt. Anić (2003: 1525) über nhd. *Schamane* aus dem Russ. entlehnt und geht ursprüngl. auf sanskr. *sehama* zurück (vgl. a. Anić/Klaić/Domović 2002: 1363). Kr. **opsjènār** m. ‘Zauberer’ (vgl. Anić 2003: 949) ist eine Ableitung aus kr. *sjèn* ‘Schatten’ (vgl. 3.2.20); bei kr. **negròmant**, **nekròmant** m. ‘Zauberer’ handelt es sich um einen über mlat. *necromantes* vermittelten Gräzismus (ev. auch über ital. *negromante*) mit zugehörigen Ableitungen **negròmantkinja**, **negròmantica** f. ‘Zauberin’ (vgl. Anić 2003: 844, ERHSJ I: 599, s.v. grabancija, Anić/Klaić/Domović 2002: 943); kr. **okùltist** (*okultist*) m. ‘jmd. der Okkultismus betreibt’ ist ein Internationalismus (< lat. *occultus*) (vgl. Anić 2003: 933, Anić/Klaić/Domović 2002: 981); bei kr. **spřitist** (*spiritist*) m., **spřitistkinja** (*spiritistica*) f. ‘jmd. der sich mit Spiritismus beschäftigt’ handelt es sich ebenfalls um Internationalismen (< lat. *spiritismus*) (vgl. Anić 2003: 1453, Anić/Klaić/Domović 2002: 1331). Als weiteres Synonym für ‘Zauberer’ führen Šarić/Wittschen (2008: 63, 311) noch (kolloquiales) kr. **švarckinstler** an, welches sich nicht bei Anić 2003, Anić/Klaić/Domović 2002: 1331 und auch nicht im ERHSJ findet, bei dem es sich aber offensichtlich um eine Entlehnung aus nhd. *Schwarzkünstler* handelt.

Kr. **vjěštica** f. (ERHSJ III: 603) bzw. *vjěštica* (Anić 2003: 1744) ‘Hexe’, **vjěštac** m. ‘Hexer, Zauberer’ sind Ableitungen von kr. *vjěšt* (vgl. aksl. *věštъ*), welches ursprüngl. ‘erfahren, bewandert, kundig’ bedeutete und mit *-jo* von nachursl. **věd-to* (vgl. aksl. *věděti* ‘wissen’, vgl. a. 3.2.12) gebildet wurde (vgl. ERHSJ III: 603 s.v. *vjěšt*, Anić 2003: 1744, Buck 1971: 1499). Die zugrundeliegende idg. Verbalwurzel ist **uejd-* ‘erblicken’, von deren Perfektstamm **uójd-/*uid-* bsl. **woi?d-*, aksl. *vědě* (aus idg. **uójd-h₂eǵ*) abgeleitet wurde (vgl. LIV: 665–667, Derksen 2008: 518). Dieselbe Bed. ‘Hexe’ weisen (kolloquiale) kr. **còprnica**, **còprnjica**, **còparica**, **còparnjica** f. (vgl. Anić 2003: 146, Šarić/Wittschen 2008: 520), Entlehnungen aus nhd. *Zauberin* bzw. ahd. *zoubar*, und kr. **štrìga** auf, welches aus ital. *strega*, friul. *stria* oder dem Dalm. entlehnt worden sein kann (< lat. *strìga*) (vgl. ERHSJ I: 272, ERHSJ III: 417 s.v. *štrìga*; vgl. a. ERHSJ III: 603 s.v. *vjěšt*; Anić 2003: 1555 kennt nur kr. *štrìga* ‘Ohrwurm’). Kr. **babaróga** f. ‘Hexe; erfundenes und meist im Dunkeln befindliches Wesen, vor dem sich Kinder fürchten’ (vgl. Anić 2003: 47, Šarić/Wittschen 2008: 520) liegt nachursl. **bàba* f. ‘alte Frau’, bsl. **bá?ba?* zugrunde (vgl. Derksen 2008: 32); der zweite Teil des Kompositums geht auf kr. *róg* ‘Horn’, aksl. *rogъ* und somit auf die idg. Verbalwurzel **(h₂)reg^h-* ‘sich aufrichten’ zurück (vgl. LIV: 498, ERHSJ III: 153–155).

3.2.22. ‘fairy or elf’

Als Entsprechung für ‘Fee’ findet sich das kr. Erbwort *vīla* f. (vgl. Anić 2003: 1733), welches allerdings eine ungesicherte Etymologie aufweist; es könnte sich lt. Skok um eine Ableitung von kr. *vīti* ‘winden’ (aus idg. **u̯ieh₁-* ‘umwickeln, umhüllen’, vgl. ERHSJ III: 600–601, LIV: 695) oder aufgrund der Zugehörigkeit zum mythologischen Bereich (vgl. a. slow. *vīlac* ‘Zauberer’ und kr. *vilènjāk* in 3.2.21) ev. um eine Ableitung (**vidla*) von kr. *vidjeti* ‘sehen’ (vgl. a. 3.2.12) handeln und somit in Zusammenhang mit kr. *vještica* ‘Hexe’ (vgl. 3.2.21) stehen (vgl. ERHSJ III: 593). Hierher gehört als Ableitung von *vīla* auch kr. *vilènica* f. ‘Fee’ bzw. ‘Ort, an dem Feen wohnen’ (vgl. Anić 2003: 1734).

Bei dem bei Buck (1971: 1500) angeführten kr. *rūsālka* f. ‘Fee; Luftgeist’ handelt es sich lt. Anić (2003: 1357) um einen Russismus in der Bed. ‘Wald- oder Wasserfee der alten Slawen’; daneben existiert kr. *rūsālka* ‘Fee’ bzw. ‘jmd. der bei religiösen Riten in Trance fällt’ (vgl. Anić 2003: 1357). Es handelt sich lt. Skok bei beiden um Überbleibsel aus dem Balkanrom. (*rosalis*), welche über das Dalm. vermittelt wurden (vgl. ERHSJ III: 175; s.v. rusalje).

Bei Šarić/Wittschen (2008: 517) finden sich zudem als weitere Synonyme für ‘Fee’ (kolloquiales) kr. *gôrkinja* f. ‘Bergfee’ (Anić 2003: 366), wohl eine Ableitung aus kr. *gòra* ‘Berg’, nachursl. **gorà*, bsl. **gor-/gir-*, idg. Verbalwurzel **g^hrH-* (vgl. Derksen 2008: 177–178), bzw. liegt das idg. Wurzelnomen **g^hōrHs* zugrunde (vgl. Matasović 2008: 179); kr. *jèžerkinja* f. ‘Fee (in Volkssagen)’ (vgl. Anić 2003: 525), eine Ableitung aus kr. *jèzero* n. ‘See’ (vgl. ERHSJ I: 781), nachursl. **èzero* n.; **èzerb* m. ‘Teich’, bsl. **ézero*, idg. **h₁éǵ^h-(e)r-o-* (vgl. Derksen 2008: 148), die idg. Verbalwurzel ist wohl **h₁eg^{uh}-* ‘trinken’ (vgl. LIV: 231); sowie kr. *vodàrica* f. ‘Wasserfee’ (vgl. Anić 2003: 1749), eine Ableitung von kr. *vòda* ‘Wasser’ (vgl. ERHSJ III: 611), ursl. **wadā* (vgl. Matasović 2008: 166), bsl. **wondōr*, idg. **uod-r/n-*⁵⁴ (vgl. Derksen 2008: 523).

3.2.23. ‘ghost’ (inkl. ‘angel’)

Als Entsprechungen für ‘Geist, Phantom’ ist kr. *ävēt* (ERHSJ I: 77) bzw. *ávet* f. und denominales *avètinja* f. (Anić 2003: 44) ‘Geist, Gespenst, Phantom’ anzuführen, das sich außer im Kr. und Sr. in keiner anderen slaw. Sprache findet. Eine Entlehnung aus arab. *hajālet* ‘ib.’ lehnt Skok aus phonologischen Gründen ab und schlägt stattdessen eine Ableitung mit dem Kollektivsuffix *-et* aus der nicht belegten Verbalwurzel **avjō*, *aviti* vor (vgl. kr. *javiti* ‘(an)zeigen’, nachursl. **aviti* ‘zeigen’ bsl. **āw-* (?), bei dem es sich lt. Derksen (2008: 30–31) ev. um eine Entlehnung aus einer iran. Sprache mit zugrundeliegendem idg. **h₂ēy-ēi-* handeln könnte, wohl zur idg. Verbalwurzel **h₁ey-* ‘sehen, erblicken’, vgl. LIV: 243), womit die ursprüngl. Bed. von *ävēt* ‘dasjenige, das sich zeigt’ wäre (vgl. ERHSJ I: 77–78). Es könnte lt.

⁵⁴ Die zugrundeliegende idg. Verbalwurzel **ued-* ‘quellen’ hat sich nur im Indo-iran. erhalten, das uridg. Substantiv **uod-r* n. ‘Wasser’ ist hingegen weit verbreitet (vgl. LIV: 658–659).

Buck (1971: 1502) allerdings auch mit türk. oder arab. *afét* ‘Übel, Schaden’ in Verbindung stehen (so auch Anić/Klaić/Domović 2002: 147).

Daneben findet sich noch kr. *ùtvara* f. ‘Gespenst, Erscheinung’ (vgl. Anić 2003: 1698) (vgl. kr. *utvoriti se* ‘erscheinen’), eine Ableitung aus der Dehnstufe **tvar-* (zu *tvor-* ‘Aussehen, äußere Erscheinung’; aksl. *tvorjǫ, tvoriti* ‘schaffen, machen, tun’, dem das Kausativum-Iterativum **tvorH-éje-* der idg. Verbalwurzel **twerH-* ‘fassen’ zugrundeliegt, vgl. LIV: 656) (vgl. ERHSJ III: 530–531, s.v. *tvòriti*; vgl. a. Buck 1971: 1502). Bei kr. *bàük* m. ‘Gespenst (so auch im Kommunistischen Manifest)’ bzw. ‘furchteinflößende Erscheinung; erfundenes Schreckgespenst für Kinder’ handelt es sich um ein Deverbativum der kr. Interjektion idg. Herkunft *bàu bàu* ‘Laute zur Beruhigung weinender oder unruhiger Kinder’ (vgl. ERHSJ I: 123, Anić 2003: 63, Šarić/Wittschen 2008: 420). Kr. *vodènjāk* m. ‘Wassergeist; Wassermann’ (vgl. Anić 2003: 1749, Šarić/Wittschen 2008: 98) ist eine Ableitung aus kr. *vòda* ‘Wasser’ (vgl. hierzu auch 3.2.22).

Weitere Synonyme sind kr. *prividénje* n. ‘Erscheinung, Geist’ (vgl. Anić 2003: 1208), neben kirchl. verwendetem kr. *vedenie* n. ‘Vision’, bei dem es lt. Skok zu einer Überkreuzung von kr. *vìdjeti* ‘sehen’ mit lat. *vedea* (aus lat. *vidēre* ‘wissen, sehen’) gekommen ist (vgl. ERHSJ III: 586, s.v. *vìdjeti* (se)); kr. *sàblast* f. ‘Erscheinung’ (vgl. Anić 2003: 1360), auch ‘Ärgernis’, ein Deverbativum aus kr. *sablázni* ‘skandalisieren’ (wiederum ein Denominativum aus kr. *blázan* ‘Skandal, Ärgernis’), das sich lt. Skok auch in anderen slaw. Sprachen, jedoch nicht im Balt. findet und eine ungesicherte Etymologie aufweist (vgl. ERHSJ I: 170–171, s.v. *blázan*); (seltenes) kr. *pričina* f. ‘Gespenst, Erscheinung’ (vgl. Anić 2003: 1185), eine Ableitung aus kr. *čìni* ‘Zauberei, Verwünschung’ (vgl. 3.2.20); kr. *prìkaza* f., *prìkazánje* n. ‘Erscheinung, Geist’ (vgl. Anić 2003: 1192), eine Ableitung aus kr. *prikázati se* ‘erscheinen’ (vgl. ERHSJ II: 69, s.v. *kázati*), nachursl. **kázàti* ‘zeigen’, idg. **k^heġ-* ‘sehen, erblicken’ (vgl. Derksen 2008: 222, LIV: 383–385); der Gräzismus (*φάντομα*) (lt. Anić 2003: 307) oder (eher) Gallizismus (*fantôm*) (vgl. ERHSJ I: 505, s.v. *fantàzija*, Anić/Klaić/Domović 2002: 426) kr. *fàntòm* m.; kr. *fantàzma* f., *fantàzam* m. ‘Erscheinung, Geist’, ein gelehrter Gräzismus (gr. *φάντομα*) (vgl. Anić 2003: 307, ERHSJ I: 505, s.v. *fantàzija*, Šarić/Wittschen 2008: 107); kr. *simulàkr(um)*, *simulàkrum* m. ‘Erscheinung, Geist’, aber auch ‘Statue, Götze’, eine Entlehnung aus lat. *simulacrum* (vgl. Anić 2003: 1400); und kr. *snòvid* m., *snòvidénje* n. ‘Erscheinung im Traum, Gespenst’ (vgl. Anić 2003: 1442), Ableitungen von kr. *vìdjeti* ‘sehen’ (vgl. 3.2.12) und Komposita mit kr. *sàn* ‘Traum, Schlaf’ (vgl. ERHSJ III: 309–310, s.v. *spàti*), nachursl. **sànb* ‘Schlaf, Traum’, bsl. **su(o)pnum*, idg. **sup-n-o* (vgl. Derksen 2008: 481) der Verbalwurzel **suep-* ‘einschlafen’ (vgl. LIV: 612–613).

Kr. *dùh* (Anić 2003: 264, Akzent nach Karadžić lt. ERHSJ I: 372) bzw. *dùh* (Holzer et al. 2012: 19, Akzent in der Aussprache von Dubrovnik lt. ERHSJ I: 372) m. bedeutet eher ‘Seele, Geist’, wurde aber in der Bed. ‘Geist’ auch in die christl. Terminologie aufgenommen (vgl. kr. *Sveti Duh* ‘heiliger

Geist') (vgl. ERHSJ I: 372, s.v. *dāh*, Anić 2003: 264, vgl. a. *dūhovnīk* 'Geistlicher' in Kap. 3.2.10), als Ableitung mit *-ia* bedeutet kr. *dúša* f. 'Seele' (aus ursl. **dōsjā* lt. Matasović (2008: 148) bzw. ursl. *dawsjā* lt. Holzer et al. (2012: 19), bsl. **dousja?*, idg. **d^hous-i-eh₂-*, vgl. Derksen 2008: 127, ERHSJ I: 372–374, s.v. *dāh*). Kr. *dūh* liegt die ablautende Dehnstufe **d^hous*⁵⁵ der idg. Verbalwurzel **d^hues-* '(ein-, aus-)atmen' (vgl. LIV: 160) zugrunde, welche bsl. **dousos* (vgl. Derksen 2008: 124) und lt. Holzer et al. (2012: 19) ursl. **dawxu* bzw. lt. Matasović (2008: 144) ursl. **dōxu* ergibt.

Als 'Schutzgeist' im weitesten Sinne kann kr. *āndeo* m., aksl. *an̄gelъ* 'Engel, Gottesbote' (vgl. Anić 2003: 21) betrachtet werden. Der christl. Terminus wurde direkt aus gr. *ἄγγελος*⁵⁶ 'Bote' (später 'Engel') entlehnt (vgl. ERHSJ I: 42–43, Buck 1971: 1468). Als weitere Synonyme für 'Engel' führen Šarić/Wittschen (2008: 29) kr. *serafim*, *serafin* an, bei dem es sich um eine über das Lat. oder Gr. vermittelte (bibl.) Entlehnung des hebr. *serāfīm* 'sechsflügeliger Engel' handelt (vgl. ERHSJ III: 223, Anić/Klaić/Domović 2002: 1286). Zudem ist hier noch kr. *heruvim* f. 'dreiflügeliger Engel' < klat. *cherūbīm* < hebr. *kerūbhīm* pl. (*kerūbh* sg. 'übernatürliches Wesen') neben *kerūbīn* < ital. *cherubino* < mlat. *cherubinus* (vgl. ERHSJ I: 665, Anić/Klaić/Domović 2002: 687) anzuführen.

3.2.24. 'omen'

Der Bed. 'Omen' entspricht kr. *znāmēnje* n. (vgl. Anić 2003: 1844), aksl. *znamēnje* '(Vor-)Zeichen; Omen', eine Ableitung aus kr. *znāmēn* 'Zeichen' (aksl. *znamę*), welches mit dem heute nicht mehr produktiven Suffix *-men* (idg. **ġneh₃-men-*, vgl. Derksen 2008: 546, Matasović 2008: 203–204) aus ursl. **znātej* (kr. *znāti*) 'kennen' (idg. **ġneh₃-* 'erkennen', vgl. Matasović 2008: 85, LIV: 168–169) abgeleitet wurde (vgl. ERHSJ III: 658, s.v. *znāti*, Buck 1971: 1504). Ebenfalls in ursl. Zeit von **znātej*, allerdings mithilfe des Suffixes *-ko* abgeleitet wurde kr. *znāk* m. 'Zeichen', von welchem wiederum kr. *prēdznāk* m. 'Vorzeichen' eine Ableitung darstellt (vgl. Anić 2003: 1155, 1844; ERHSJ III: 658). Weiters weisen auch kr. *pretkazanje* n., eine Ableitung des kr. Verbs *kázati* 'zeigen' (vgl. ERHSJ II: 69), welches auf die idg. Verbalwurzel **k^hek-* 'sehen, erblicken' (vgl. LIV: 383–385) zurückgeht (vgl. a. kr. *príkaza* in 3.2.23); und kr. *pòjava* f. (vgl. Anić 2003: 1079), eine deverbale Ableitung von kr. *jáviti* 'anzeigen, bekannt machen' (vgl. ERHSJ I: 762–763) (vgl. a. 3.2.23), die (Neben-)Bed. 'Vorzeichen, Omen' auf. Daneben ist noch

⁵⁵ „The root shape **d^hous-* is peculiar to Balto-Slavic. Elsewhere in Indo-European – and in Balto-Slavic as well – we find full grades of the type **d^hue/ous-* [...]“ (Derksen 2008: 124). Kr. *dāh* m. 'Hauch, Atem' (ursl. **duxu* > **dъxъ* > kr. *dāx*, vgl. Holzer 2011: 100, 2012: 14) hingegen ist auf die Schwundstufe der idg. Verbalwurzel **d^hues-* '(ein-, aus-)atmen' zurückzuführen, wie im zugehörigen Aoriststamm **d^hues-/d^hus-* (> aksl. *iz-dъše* 'hauchte aus, gab den Geist auf'), Präsensstamm **d^hus-jé-* (> aksl. *dušę*, *duhati* 'atmen, wehen') und Essiv **d^hus-hjé-* (> aksl. *dyšę*, *dyšati* 'atmen, wehen') (vgl. LIV: 160). Im Bsl. und Germ. existiert zudem die sekundäre Vollstufe **d^heus-* > got. *dius* n. 'wildes Tier', vgl. a. germ. **deuza-* 'Tier' aus **d^heus-ó-* '*mit Atem versehen' (vgl. LIV: 160, Derksen 2008: 124).

⁵⁶ Im Aksl. selten mit *сѣлъ* übersetzt (auch 'Apostel', vgl. Derksen 2008: 480), vgl. a. slow. *sel*, sr. *vjesnik* 'Bote' oder durch eine Umschreibung (slow. *krilatec* 'Geflügelter', aksl. *šestokrilatъcb* ist ein Calque) ausgedrückt (vgl. Miklosich 1876: 41).

čak. *pégula* neben *pégola* (vgl. Anić 2003: 1014) zu nennen, welches neben ‘Teer, Pech’ auch im übertragenen Sinn ‘Pech, schlechtes Omen’⁵⁷ bedeutet. Hierbei handelt es sich lt. Skok und Anić um eine Entlehnung aus ital. *pégola* < lat. *pīcula* (Deminutiv von *pix* ‘Pech, Teer’, welches in Urverwandschaft mit aksl. *рѣкъль*, kr. *pàkao* ‘Hölle’ steht) (vgl. ERHSJ II: 631). Der Internationalismus kr. *omen* ist eine Entlehnung aus lat. *ōmen* (dessen Etymologie selbst allerdings ungeklärt ist, vgl. Buck 1971: 1503).

Als weitere Entsprechung⁵⁸ für ‘Omen’ ist kr. *kôb* f., *koba* f. ‘Omen’ (kr. *kòbiti* ‘Unglück vorhersagen’) anzuführen, welches als gesamt- und ursl. mythologischer Terminus das ‘Wahrsagen aus dem Vogelflug’ bezeichnet und somit als ‘Schicksal’ generell sowohl ‘gutes Omen’ (und ‘Glück’) als auch ‘schlechtes Omen’ bedeutet. Auf die ursprüngl. mythologische Bed. bezieht sich auch die Ableitung kr. *kôbac* m. (ursl. **kabiiku*, vgl. Holzer et al. 2012: 37) ‘Sperber, Accipiter nisus L., Merlin’, welche in Urverwandschaft mit nhd. *Habicht* und aisl. *haukr* ‘Habicht’ steht (vgl. ERHSJ II: 116; Anić 2003: 583). „The suggestion that this etymon is cognate with **koby* is based on augural rites.“ (Derksen 2008: 227). Mit der Bedeutungsentwicklung von kr. *kôb* vergleichbar ist gr. *οἰωνός* ‘Greifvogel; der Ornithomantie dienender Vogel’ (mit ungeklärter Etymologie) neben *ὄρνις* ‘Vogel, Omen’, sowie lat. *auspicium* ‘Wahrsagerei mithilfe von Vögeln, Omen’ (von *auspex* als Kompositum mit *avis* ‘Vogel’ und der Wurzel von *specere* ‘beobachten’) (vgl. Buck 1971: 1503–1504). Lt. Derksen (2008: 227) könnte es sich bei nachursl. **koby* (ursl. **kabi* ‘(vorhergesagtes) Schicksal’, vgl. Holzer et al. 2012: 37) um eine Entlehnung aus einem europ. Substrat handeln, das vielleicht mit air. *cob* ‘Sieg’, sicher mit aisl. *happ* ‘Erfolg’ in Urverwandschaft steht. Auch Skok (ERHSJ II: 116) sieht aufgrund der möglichen Parallele der Vogelnamen *kobac* und *Habicht* zu klat. *capus*, *capys* etruskischen Ursprungs die Möglichkeit gegeben, dass es sich (auch aufgrund fehlender balt. Parallelen) um ein voridg. Substratwort handelt.

3.2.25. ‘circumcision’

Der Bed. ‘Beschneidung’ entspricht kr. *sūnet* (ERHSJ III: 362) m. bzw. *sūnet* (Anić 2003: 1502) und daraus abgeleitetes *sūnečēnje* n. und *sūnetiti* ‘beschneiden’ (vgl. Anić 2003: 1502). Es handelt sich um einen Balkanturzimus aus türk. *sūnnet* < arab. *sūnnet* (vgl. ERHSJ III: 362). Daneben existiert als kr. Erbwort *òbrezati* ‘beschneiden’ bzw. *òbrezivanje* n. ‘Beschneidung’ (vgl. Anić 2003: 902, ERHSJ III: 134 s.v. rezati), das auf idg. **ureh₁ǵ-* ‘brechen, reißen’ zurückgeht (vgl. LIV: 698); als Latinismus

⁵⁷ Nur die Bed. ‘Pech’ aber nicht ‘Omen’ weisen der kr. Germanismus *pěh* m. < nhd. *Pech* (vgl. Anić 2003: 1014, ERHSJ II: 631 s.v. *pégula*) und das kr. Erbwort *smòla* f. auf (vgl. Anić 2003: 1438, lt. Skok (ERHSJ III: 294) ev. auf idg. **smel-*, **(s)meld-* ‘weich werden’ mit *s*-mobile wie in ahd. *smelzen* ‘schmelzen’ (vgl. LIV: 431) oder idg. **suel-* ‘schwelen, brennen’ wie lit. *svelù* ‘glimmen, schwelen’ (vgl. LIV: 609) zurückzuführen), bei denen eine Bedeutungsübertragung von ‘Pech, Teer’ zu ‘Unglück’ stattfand.

⁵⁸ Šarić/Wittschen (2008: 183, 343) führen als weitere kr. Synonyme für ‘Omen’ *ùsud* ‘Schicksal’, *sùdbina* ‘ib.’, *ùdes* ‘ib.’, *zla sùdbina* ‘böses Schicksal’, *nèscréca* ‘Unglück’, *zla srěca* ‘ib.’; sowie für ‘Erscheinung’ (*neobično*) *bíce* ‘übernatürliches Wesen’, *entitēt* ‘Entität’, *fenòmēn* ‘Phänomen’, *dogadžaj* ‘Vorkommnis’, *slučaj* ‘Vorfall’, *očitovanje* ‘Manifestation’ an.

(*circumcīso, circumcidere*) findet sich noch kr. *cirkumcīzija* f. ‘Beschneidung (sowohl religiös als auch medizinisch)’ (vgl. Anić 2003: 142, Anić/Klaić/Domović 2002: 228, Šarić/Wittschen 2008: 293).

3.2.26. ‘initiation ceremony’

Für ‘Initiationsritus’ wird hier (in Anlehnung an Bartels (2009b), welcher hierfür die nsorb. Entsprechung *konfirmacija* wählt) die Bedeutungsspezialisierung ‘Konfirmation’ gewählt. Neben dem (seltenerem) Latinismus *konfirmacija* f. (< lat. *confirmatio*, vgl. a. Anić/Klaić/Domović 2002: 723) wird im Kr. *krīzma* f. (vgl. Anić 2003: 639) neben *krīsma* (dial.) f. zur Bezeichnung der ‘Firmung’ verwendet. Es handelt sich um eine mittelbare Entlehnung des lat. kirchl. Terminus *chrisma* für *consignatio* (> ital. *crēsima*) < gr. Abstraktum *χρῖσμα* (von *χρίω* ‘Salbe’), welche sich lt. Skok (ERHSJ II: 197) in allen slaw. Sprachen außer dem Russ. und Bulg. findet. Kr. *krīzma* dürfte aufgrund der durchgeführten westrom. Sonorisierung über das Friul., *krīsma* hingegen über das Dalm. vermittelt worden sein (vgl. ERHSJ II: 196–197; vgl. a. ERHSJ II: 155, s.v. korisma). Daneben existieren noch kr. *sveta pōtvrda* f. bzw. *sakrament potvrde* ‘Konfirmation’ (vgl. Anić 2003: 1129), eine Ableitung des kr. Adjektivs *tvŕd* ‘fest’ (vgl. ERHSJ III: 531), nachursl. **tvŕdъ*, idg. **tuorH-o-* (vgl. Derksen 2008: 501), idg. Verbalwurzel **tuŕH-* ‘fassen’ (vgl. LIV: 656) (vgl. a. kr. *ūtvara* in 3.2.23); sowie mittelbar über ital. *firma, firmare* aus lat. *firmus* entlehntes (jedoch nur regional verwendetes) kr. *firma* f. ‘Firmung’; neben čak. *fērma* und *bērma* (über aital. *ferma, fermare*) und čak. *bīrma* f. (vgl. slow. *birma*) (vgl. Anić 2003: 318, ERHSJ I: 512 s.v. ferma; vgl. a. Miklosich 1876: 30, Šarić/Wittschen 2008: 199).

4. Kontaktsituationen

Die in der semantisch-lexikalischen Gruppe „Religion und Glaube“ vorgefundenen Lehnwörter im Kr. stammen vorrangig aus dem Romanischen, Germanischen und Griechischen. Hinzu kommt ein nicht unerheblicher Anteil aus dem Türkischen, Ungarischen und (im Falle vorursl. entlehnter mythologischer Termini) iran. Sprachen. Ein Großteil wurde im Zuge der Christianisierung entlehnt, es handelt sich somit um *cultural borrowings* (vgl. Haspelmath 2009: 46–47), da die betreffenden Bezeichnungen im Zuge der Vermittlung eines neuen Konzepts (dem Christentum) aus der Sprache der jeweiligen Missionare übernommen wurden (vgl. a. Miklosich 1876: 2). Da einige Lehnwörter aber nicht nur in den bereits vorhanden Wortschatz eingefügt wurden, sondern auch mit bereits vorhandener (ev. auch umgedeuteter) religiöser Terminologie koexistieren, handelt es sich in diesen Fällen um *core borrowings* (vgl. Haspelmath 2009: 48–49).

4.1. Romanisch

Bei die kirchl. Terminologie betreffenden Lehnwörtern aus dem Romanischen kommen mehrere Gebersprachen in Frage (vgl. Skok 1953: 56), zum einen das mittlerweile ausgestorbene autochthone Romanische auf heutigem kr. Gebiet wie das Istriotische und Dalmatische – allerdings finden sich hieraus entlehnte Wörter vorrangig in kr. Dialekten (vgl. Holzer 2011: 15, Matasović 2008: 309) –; zum anderen allochthones Venezianisch, Istrorumänisch und Toskanisch (vgl. Holzer 2011: 15), wobei die größte Anzahl an romanischen Lehnwörtern im Kr. aus dem Venezianischen kommt (vgl. Matasović 2008: 311), in diesen Fällen handelt es sich jedoch um jüngere Entlehnungen. (Kirchen-)Latein wurde in vielen Fällen über das Althochdeutsche oder Griechische vermittelt (vgl. Auty 1976), durch die im 8. und 9. Jh. stattfindende fränkische Missionierung⁵⁹ der Slawen von Aquileia aus (vgl. Katičić 1999: 188, 316, Holzer 2011: 11, Skok 1953: 51, 61) hat das Friulanische (bzw. „aquilejisches Kirchenlatein“, vgl. Zagiba 1964: 291) einen bedeutenden Anteil an Lehnwörtern im Bereich der kr. religiösen Terminologie hinterlassen (vgl. Katičić 1980: 31). Allerdings wurden die Südslawen wohl auch schon vor der fränkischen Eroberung vom Patriarchat Aquileia ausgehend missioniert (vgl. Krahwinkler 1992: 118, 161, Matasović 2008: 310); „welsche“ Missionare waren zudem ab Ende des 8. Jh. auch in Pannonien (zu Beginn im Zuge der Missionierung sowohl der Awaren als auch der Slawen) und Mähren tätig (vgl. Zagiba 1964: 280, 288). Allerdings ist auch eine Christianisierung der Kroaten durch die erneuerten dalmatinischen Kirchensitze zwischen dem 7. und 9. Jh. nicht auszuschließen (mit dem Diözesanzentrum Salona), wenngleich die Quellenlage über die kirchl. Verhältnisse dieser Zeit mangelhaft ist (vgl. Katičić 1999: 189, Zagiba 1964: 276). Lehnwörter aus dem Friulanischen zeichnen sich durch die im 5. Jh. stattfindende Westromanische Sonorisierung aus (vgl. Skok 1953: 53 Anm. 13; 54, Anm. 25, vgl. a. Matasović 2008: 310, vgl. a. Anm. 27 in Kap. 3.2.7), welche im autochthonen Romanischen des heutigen Kroatien nicht stattfand (vgl. Holzer 2011: 19). Dass viele der betreffenden Wörter dem im Laufe des frühen 9. Jh. stattfindenden Wandel $a > o$ unterliegen (vgl. Holzer 2011: 58) zeigt, dass es sich um frühe Entlehnungen handeln muss (vgl. Skok 1953: 54).

4.2. Germanisch

Vorursl. Kontakte zum Germanischen sind zeitlich und räumlich schwer zu bestimmen (vgl. Matasović 2008: 49, vgl. a. Pronk-Tiethoff 2012). Die meisten Lehnwörter, welche die religiöse Terminologie betreffen (darunter auch viele Latinismen und einige Calques), kommen allerdings aus dem

⁵⁹ Nach der fränkischen Eroberung Friauls 774/776 (vgl. Krahwinkler 1992: 158–164) wurde die Drau 811 von Karl dem Großen als Diözesangrenze festgelegt, wodurch die Gebiete nördlich der Drau (somit auch Karantanien) Salzburg, jene südlich der Drau Aquileia unterstanden (vgl. Krahwinkler 1992: 167, Katičić 1980: 31, Zagiba 1964: 276); Sakramentarien des Patriarchats Aquileia wurden jedoch weiterhin im Gebiet der Erzdiözese Salzburg verwendet (vgl. Zagiba 1961: 31–34, 36).

Althochdeutschen und wurden im 8. und 9. Jh. durch die vorkyrillomethodianische Missionierung von Salzburg, Regensburg oder Passau aus (vgl. Matasović 2008: 52, Auty 1976: 171, Zagiba 1961: 10), oder durch die fränkische Missionierung aus Aquileia (vgl. Kap. 4.1) übermittelt. Vereinzelt beginnt eine bairische Missionierung der Slawen bereits im 7. Jh. (vgl. Zagiba 1961: 3). Viele der betreffenden Lehnwörter finden sich in der gesamten Slavia und sind durch die Expansion des Fränkischen Reichs (und davor der röm.-kath. Kirche) Richtung Osten bedingt (vgl. Pronk-Tiethoff 2012: 236). Die Weitervermittlung althochdeutscher (aber auch romanischer) Lehnwörter ist der kyrillomethodianischen Mission zu verdanken, da diese als kirchl. Termini auch ins Kirchenslawische übernommen wurden (vgl. Zagiba 1964: 291; 1961: 1).

4.3. Iranische Sprachen

Bereits seit der Ausgliederung der einzelnen Sprachzweige aus der idg. Grundsprache bestanden zwischen den iran. und den slaw. Stämmen enge Kulturkontakte, die sich auch später in der ethnischen Vielfalt der Černjachov-Kultur ablesen lassen, welche Iraner, Dako-Mysier, Germanen (v.a. Goten) und Slawen umfasste (hierbei ist aber nur ein Teil des slaw. Sprachraums betroffen) (vgl. Katičić 1999: 110–111). Die engen Beziehungen des Slaw. zu den iran. Sprachen sind durch Kontakte slawischsprachiger Stämme zu iran. Reiternomaden am Rand der pontischen Steppenzzone begründet (vgl. Katičić 1999: 106–107) und führten zu einer Reihe von Entlehnungen aus den iran. Sprachen (so ev. auch der Name der Kroaten selbst, wenn er nicht awar. Herkunft ist, vgl. Katičić 1999: 194, 195). Zur Diskussion iran. Lehnwörter im Slaw. vgl. a. Matasović (2008: 47–48). Lt. Meillet (1926: 166–167) gibt es im Slaw. und Indo-Iranischen eine Reihe lexikalischer Gemeinsamkeiten, welche das Balt. nicht teilt, und die Lt. Meillet statt auf Entlehnungen auf eine „[...] voisinage ancien des futurs parlars slaves et des futurs parlars aryens dans cette communauté“ (Meillet 1926: 174) zurückzuführen sind. Dies betrifft gerade den Bereich der religiösen Terminologie, wobei v.a. das Slaw. altes idg. religiöses Vokabular konserviert hat (vgl. a. Katičić 1999: 106) und dieses wiederum viele Übereinstimmungen mit den iran. Sprachen aufweist (vgl. Meillet 1926: 169). Auch wenn diese Terminologie durch die Christianisierung größtenteils wieder obsolet wurde, ist in Hinsicht auf die lexikalischen Gemeinsamkeiten mit den iran. Sprachen „[...] le nombre de ces concordances [...] digne de remarque. Il a sûrement une signification et montre que, à date ancienne, les termes religieux étaient à peu près les mêmes chez les Slaves et chez leurs voisins Aryens“ (Meillet 1926: 171).

4.4. Griechisch, Türkisch, Ungarisch, Slawisch

Entlehnungen aus dem Türkischen (hier meist arab. Herkunft) fanden seit Beginn der osmanischen Herrschaft am Balkan ab dem 15. Jh. statt und finden sich auch weiterhin in der heutigen Standardsprache

(vgl. Matasović 2008: 311–312). Etwa zur selben Zeit wurden auch viele Wörter aus dem Neugriechischen entlehnt (vgl. Matasović 2008: 312); durch kirchl. Vermittlung (vorrangig der kyrillomethodianischen Mission und durch das Aksl., vgl. a. Miklosich 1876) kommt es zudem zur Übernahme einer großen Anzahl an Gräzismen. Durch die direkte Nachbarschaft zum Ungarischen kommt es auch hier zu Entlehnungen, die betreffenden Wörter finden sich aber hpts. in kr. Dialekten (vgl. Matasović 2008: 314). Magyarismen werden jedoch auch durch kirchl. Vermittlung über das Zagreber Bistum aufgenommen (vgl. u. a. ERHSJ II: 42, s.v. kapitul). Entlehnungen aus anderen slaw. Sprachen sind im Kr. nur schwer zu erkennen, da die betreffenden Lehnwörter nur selten auch lautlich (z. B. *-št-* statt zu erwartendem *-ć-*, vgl. sr. *opština* aber kr. *općina*) zu identifizieren sind (vgl. Matasović 2008: 314), v.a. wenn diese „kroatisiert“ wurden (vgl. Kap. 6.2). Bis ins 19. Jh. (vermehrt Anfang des 17. Jh. zur Zeit der Reformierung des Ksl.) werden zudem viele Russismen über das Russ.-Ksl. aufgenommen (vgl. Matasović 2008: 314).

5. Lehn- und Erbwortschatz des Kroatischen in der semantisch-lexikalischen Gruppe „Religion und Glaube“

Für die in Kap. 4 ermittelten kr. Entsprechungen der Bedeutungen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ wird im Folgenden der Lehnwortstatus festgestellt (wobei auch Calques, mit entlehnten Basen gebildete Lexeme sowie Bedeutungserweiterungen und -übertragungen miteinbezogen werden, vgl. Kap. 2) und der Übersicht halber nochmals in einer Tabelle zusammengefasst. Auf das Verhalten des jeweiligen Lehnworts in der Nehmersprache (Ersetzung, Einfügung, Koexistenz, vgl. Haspelmath/Tadmor 2009b: 16) wird nicht weiter eingegangen.

5.1. Analysierbarkeit und Lehnwortstatus

Aufgrund der in Kap. 2 besprochenen Problematik betreffs lexikalischer und semantischer Entlehnungen wird die den Lehnwortstatus eines Wortes betreffende Skala bei Haspelmath/Tadmor (2009b: 139) übernommen (*no evidence for borrowing* = 5, *very little evidence for borrowing* = 4, *perhaps borrowed* = 3, *probably borrowed* = 2, *clearly borrowed* = 1⁶⁰), jene die Analysierbarkeit betreffende Skala (*unanalyzable*, *semi-analyzable*, *analyzable derived*, *analyzable compound*, *analyzable phrasal* bei Haspelmath/Tadmor (2009b: 12) jedoch dahingehend adaptiert, dass nicht nur dem damit zusammenhängenden (morphologischen und phonologischen) Integrationsgrad der betreffenden

⁶⁰ Auch (gesicherte oder potenzielle) Entlehnungen aus vorursl. Zeit wurden mit *clearly*, *probably* oder *perhaps borrowed* versehen, da sie im Kr. zwar aus dem Ursl. ererbt sind, es sich aber dennoch um Lehnwörter handelt. Dasselbe gilt für ins morphologische System integrierte oder mit entlehnten Basen gebildete Entsprechungen und für Calques; Bedeutungserweiterungen oder -übertragungen werden mit *probably* oder *perhaps borrowed* bzw. *very little evidence for borrowing* versehen.

Lehnwörter, sondern auch semantischen und strukturellen Entlehnungen Rechnung getragen wird: unanalysierbar (A), ins morphologische System (mittels Suffigierung o.Ä.) integriert (B), mit entlehnter Basis gebildet (C), Calque (D), Bedeutungserweiterung bzw. -übertragung, d. h. semantische Entlehnungen (E).

Tab. 1: Bedeutungen der „LWT meaning list“, kr. Entsprechungen, Lehnwortstatus, Analysierbarkeit, Herkunft

| LWT meaning list | Kroatisch | Lehnwortstatus | Analysierbarkeit | Herkunft |
|------------------|---|----------------|------------------|---|
| religion | <i>rèligija</i> ‘Glaube’ | 1 | B | < lat. <i>religiō</i> |
| | <i>vjerozákon</i> ‘Religion (als Glaubenssystem)’ | 3 | D (?) | ursl. <i>*wērā</i> , idg. <i>*ueh₁-r-eh₂-</i> ; nachursl. <i>začęti</i> (Calque?) |
| | <i>vjěra</i> ‘Glaube, Vertrauen’ | 5 | - | ursl. <i>*wērā</i> , idg. <i>*ueh₁-r-eh₂-</i> |
| | <i>pōbožnōst</i> ‘Religiosität’ | 1 | C | ursl. <i>*baga</i> (< iran. <i>baga-</i> ?) |
| | <i>ispovijēd</i> , <i>ispovijēst</i> (arch.) ‘Bekenntnis’ | 1 | D | idg. <i>*uejd-</i> (Calque lat. <i>confessio</i>) |
| | <i>vjeroispovijēst</i> , <i>veroispovijēd</i> ‘Konfession, Glaube’ | 3 | D (?) | ursl. <i>*wērā</i> , idg. <i>*ueh₁-r-eh₂-</i> ; idg. <i>*uejd-</i> (tlw. Calque) |
| god | <i>konfēsija</i> ‘Glaube; Glaubensbekenntnis’ | 1 | B | < lat. <i>confessio</i> |
| | <i>Bōg</i> ‘Gott’ | 1 | A (D) | ursl. <i>*baga</i> (< iran. <i>baga-</i> ?) |
| | <i>božanstvo</i> ‘Gottheit’ | 1 | C | ursl. <i>*baga</i> (< iran. <i>baga-</i> ?) |
| temple | <i>Gōspōd</i> ‘Herr’ | 5 | - | idg. <i>*g^host(i)-pot-</i> |
| | <i>hrām</i> ‘Tempel’ (‘Kirche’) | 5 | - | idg. <i>*(s)ker-</i> |
| | <i>sakrarij</i> (selten) ‘Heiligtum, Tempel’ | 1 | B | < lat. <i>sacrarium</i> |
| | <i>svetište</i> ‘Heiligtum, Tempel’ | 3 | E | ursl. <i>*swentu</i> , idg. <i>*kuejt-</i> (Bedeutungsübertragung) |
| | <i>idòlija</i> (selten) ‘heidn. Tempel’ | 1 | B | < gr. <i>εἰδωλεῖον</i> |
| church | <i>pānteōn</i> (arch.) ‘Tempel (aller Götter)’ | 1 | A | < gr. <i>πάνθειον</i> |
| | <i>črkva</i> ‘Kirche (sowohl Bauwerk als auch Organisation)’ | 1 | B (D) | < westgerm. <i>*kirikō</i> (ahd. <i>kirihha</i> , <i>ckīlīhha</i> , bair. dial. <i>chirhha</i> , got. <i>*kyrikō</i> ? < gr. <i>κυρικόν</i> (Bedeutungserweiterung) |
| | <i>bogòmolja</i> ‘Bethaus (Kirche, Halle, Tempel usw.)’ | 1 | C | ursl. <i>*baga</i> (< iran. <i>baga-</i> ?); ursl. <i>*modliiti</i> , idg. <i>*meld^h-</i> |
| | <i>sinagóga</i> ‘Synagoge, Bethaus’ | 1 | A | < lat. <i>synagōga</i> < gr. <i>συναγωγή</i> |
| mosque | <i>zbórnica</i> (arch.) ‘*Tempel, Synagoge’, ‘Versammlungsort’ | 1 | D | nachursl. <i>*bьrati</i> , bsl. <i>*ber-</i> , idg. <i>*b^her-e/o-</i> (Calque von lat. <i>synagōga</i> oder gr. <i>συναγωγή</i>) |
| | <i>džámija</i> ‘Moschee’ | 1 | B | < türk. <i>cami</i> < arab. <i>ġami</i> |
| | <i>mošėja</i> (selten) ‘Moschee’ | 1 | A (B) | < dt. <i>Moschee</i> < ital. <i>moschea</i> oder frz. <i>mosquée</i> < türk. < arab. <i>mäsġid</i> |
| altar | <i>měcet</i> , <i>mečit</i> ‘Moschee’; <i>měsdžid</i> , <i>měsdžet</i> ‘islam. Bethaus ohne Minarett’ | 1 | A | < türk. <i>mesced</i> , <i>-id</i> < arab. <i>mäsġid</i> |
| | <i>òltār</i> ‘Altar’ | 1 | A | < ahd. <i>altari</i> , mgr. <i>ἀλτάριον</i> , lat. <i>altāria</i> , <i>altāre</i> |
| | <i>àltār</i> (čak., selten) ‘Altar’ | 1 | A | < ahd. <i>altari</i> , mgr. <i>ἀλτάριον</i> , lat. <i>altāria</i> , <i>altāre</i> |
| sacrifice | <i>žrtvenik</i> ‘heidn. Altar’ | 1 | C | < (russ. <i>žertva</i> ?), ursl. <i>*žer-ti</i> , idg. <i>*g^herH-</i> |
| | <i>žrtva</i> ‘Opfer’ | 1 | A (B) | < (russ. <i>žertva</i> ?), ursl. <i>*žer-ti</i> , idg. <i>*g^herH-</i> |
| | <i>àldōv</i> (kajk.) ‘Opfer’ | 1 | B | < ung. <i>aldó</i> |
| | <i>kurbān</i> ‘Opfer’ | 1 | A | < türk. <i>kurban</i> < arab. <i>kurban</i> |

| | | | | |
|----------------------------------|--|-------|---|---|
| | <i>òblānda, òblānta</i> (čak.) 'Oblate, Hostie' | 1 | A | friul. (<i>òblānda</i>) bzw. dalm. (<i>òblānta</i>) < lat. <i>oblata</i> |
| | <i>ofar</i> (čak.) 'Opfer, Gabe' | 1 | A | < ahd. <i>offar</i> < lat. <i>offere</i> (ev. Überkreuzung mit lat. <i>operāri</i>) |
| | <i>ostija</i> 'Hostie' | 1 | A (B) | < ital. <i>ostia</i> oder lat. <i>hostia</i> |
| | <i>hostija</i> 'Hostie' | 1 | A (B) | < lat. <i>hostia</i> |
| | <i>prosfor, pròsfora</i> 'Hostie' | 1 | A | < gr. <i>προσφορά</i> |
| | <i>proskura, proskomidija</i> 'Hostie' | 1 | A | < ngr. <i>προσκομιδή</i> |
| | <i>svetba</i> 'Sakrament, Opfer, Hostie' | 5 | - | ursl. * <i>swentu</i> , idg. * <i>k̑eǵt-</i> |
| | <i>pāljenica</i> 'religiöses (Brand-)Opfer' | 5 | - | ursl. * <i>pol-</i> , idg. * <i>pōlh₁-</i> bzw. * <i>pel-</i> |
| | <i>posvećenje</i> 'dargebrachtes Opfer' | 5 | - | ursl. * <i>swentu</i> , idg. * <i>k̑eǵt-</i> |
| to worship | <i>posvétiti</i> 'heiligen, ehren' | 5 | - | ursl. * <i>swentu</i> , idg. * <i>k̑eǵt-</i> |
| | <i>obožávati</i> 'als Gottheit verehren' | 1 | C | ursl. * <i>bagu</i> (< iran. <i>baga-</i> ?) |
| | <i>poštòvati, štòvati</i> 'ehren' | 5 | - | idg. * <i>k^heǵt-</i> |
| | <i>slāviti</i> 'feiern, preisen' | 5 | - | ursl. * <i>slāvā</i> , idg. * <i>k̑leu-os-</i> |
| | <i>poklòniti se</i> 'huldigen' | 5 | - | idg. * <i>k̑leǵ-</i> (?) |
| | <i>divinizírati</i> 'als Gottheit verehren' | 1 | C | < ital. <i>divino</i> < lat. <i>dīvinus</i> |
| | <i>díviti se</i> 'anbeten; bewundern' | 5 | - | nachursl. * <i>dívьnъ</i> , idg. * <i>d^hHi-uo-</i> / * <i>d^heiH-uo-</i> |
| | <i>adorírati, adòrati, adorávati</i> 'verehren' | 1 | C | < ital. oder lat. <i>adōrāre</i> |
| | <i>uvažávati</i> '(ver-)ehren' | 1 | C | < ahd. <i>vaga</i> |
| | <i>uznòsiti, ùznijēti</i> 'anpreisen, verherrlichen' | 5 | - | nachursl. * <i>nositi</i> , idg. * <i>h₁nek-</i> |
| | <i>glorificírati</i> 'verehren' | 1 | C | < lat. <i>glōrificāre</i> |
| <i>veličati</i> 'lobpreisen' | 5 | - | nachursl. * <i>velikъ</i> , idg. * <i>uelh₁-</i> | |
| <i>hváliti</i> 'loben, preisen' | 5 | - | ursl. * <i>xwālins</i> | |
| to pray | <i>mòliti (se)</i> 'beten' | 5 (4) | E (?) | ursl. * <i>modliti</i> , idg. * <i>meld^h-</i> (Bedeutungsübertragung?) |
| priest | <i>svećenik</i> 'Priester' | 1 | D | ursl. * <i>swentu</i> , idg. * <i>k̑eǵt-</i> (Calque von gr. <i>ἱερέυς, ἱερός</i>) |
| | <i>jèrej</i> 'orthodoxer Priester' | 1 | B | < gr. <i>ἱερέυς</i> |
| | <i>glagòljāš</i> (arch.) 'kath. Priester, Glagolit' | 5 | - | nachursl. * <i>golgolъ</i> , idg. * <i>gal-</i> |
| | <i>latināš</i> 'Priester, der die Messe in lateinischer Sprache hält' | 1 | B | < lat. <i>latīnus</i> |
| | <i>dušòbrižnik</i> 'Seelsorger' | 1 | D | ursl. * <i>dōxu</i> (Calque nhd. <i>Seelsorger</i>) |
| | <i>dùhovnik</i> 'Geistlicher' | 1 | D | ursl. * <i>dōxu</i> (Calque nhd. <i>Geistlicher</i> oder gr. <i>πνευματικός</i>) |
| | <i>pòp</i> (pejor., kolloquial) 'Priester' | 1 | A | < ahd. <i>phapho, phaffo</i> , got. * <i>papa</i> , gr. <i>παπᾶς</i> ? |
| | <i>mantijāš</i> (pejor., kolloquial) 'Priester, Pfarrer, jmd. der ein Priesterkleid trägt' | 1 | B | < gr. <i>μαντίον</i> |
| | <i>přvad</i> (čak.) '(höchster Rang eines) Diakon, Priester' | 1 | A | < afriul. * <i>preved(e)</i> < vlat. <i>praebyster</i> < lat. <i>presbyter</i> < gr. <i>πρεσβύτερος</i> |
| | <i>prèzbiter</i> '(höchster Rang eines) Diakon, Priester' | 1 | A | < lat. <i>presbyter</i> < gr. <i>πρεσβύτερος</i> |
| | <i>podreka</i> (arch.) 'Oberhaupt der Kirche' | 1 | A | < friul. <i>padriarca</i> < lat. <i>patriarcha</i> |
| <i>mìsnik</i> (dial.) 'Priester' | 1 | C | < lat. <i>mìsa</i> | |

| | | | | |
|-----------|---|---|-------|--|
| | <i>mašnik</i> (kajk.) ‘Priester’ | 1 | C | < friul. < lat. <i>mīsa</i> |
| | <i>rābīn</i> ‘Rabbi, hebr. Priester’ | 1 | B (?) | < hebr. <i>rabbi</i> |
| | <i>hōdža</i> ‘muslim. Priester; Glaubenslehrer’ | 1 | A | < türk. <i>koca/koç</i> < pers. <i>haḡe</i> < <i>hāwaḡa</i> |
| | <i>šāmān</i> ‘Priester bei Tataren, Mongolen usw.; Schamane’ | 1 | A | < nhd. <i>Schamane</i> < russ. < sanskr. <i>sehama</i> |
| | <i>žākan</i> ‘Diakon, Kleriker, niedrigster Priesterrang’ | 1 | A | < istr./dalm. < lat. <i>diāconus</i> < gr. <i>διάκονος</i> |
| | <i>đākon</i> ‘Diakon, Hilfspriester’ | 1 | A | < lat. <i>diāconus</i> < gr. <i>διάκονος</i> |
| | <i>dijakon</i> ‘Diakon, Hilfspriester’ | 1 | A | < ung. < lat. <i>diāconus</i> < gr. <i>διάκονος</i> |
| | <i>prēpošt</i> (kajk.) ‘Aufseher, Vorsteher’ | 1 | A | < ung. (oder venez./ital.) < lat. <i>praepos(it)us</i> |
| | <i>prepōzīt</i> ‘Aufseher, Vorsteher, Prior’ | 1 | A | < lat. <i>praepositus</i> |
| | <i>vikār</i> ‘Stellvertreter eines Bischofs’ | 1 | A | < lat. <i>vicārius</i> |
| | <i>kapèlān</i> ‘(arch.) Priester, der sich um die Pfarre kümmert’ bzw. ‘Heeres-, Spitals-, Hilfspriester’ | 1 | A | < klat. <i>capellanus</i> |
| | <i>kaplan</i> (kajk.) ‘Kaplan’ | 1 | A | < nhd. <i>Kaplan</i> < klat. <i>capellanus</i> |
| | <i>küstōd</i> ‘Küster’ | 1 | A | < lat. <i>custōdis</i> |
| | <i>kustoš</i> (kajk.) ‘einer der Chorherren, Kanoniker’ | 1 | A | < ung. < lat. <i>custōs</i> |
| | <i>pāstīr</i> ‘Hirte; Priester’ | 5 | - | nachursl. <i>*pasti</i> , idg. <i>*peh₂s-</i> |
| | <i>pāstor</i> ‘protestantischer Priester, Pfarrer’ | 1 | A | < lat. <i>pāstor</i> |
| | <i>isповјèdник</i> ‘Beichtvater’ | 1 | C | idg. <i>*uejd-</i> (vgl. <i>īspovijēd</i>) |
| | <i>nadarbènik</i> ‘Priester’ | 1 | D | nachursl. <i>*dāti</i> , bsl. <i>*doʔ-</i> ; <i>*doʔd-mi</i> ; idg. <i>*deh₃-</i> ; <i>*didh₃-</i> (Calque von lat. <i>praebēnda</i>) |
| | <i>prebēndār</i> ‘kath. Priester, der vom Ertrag des Landes oder einer Schenkung lebt’ | 1 | A | < lat. <i>praebendarius</i> |
| ‘Pfarrer’ | <i>pārōh, parohijān</i> ‘Pfarrer; orthodoxer Priester’ | 1 | A (B) | < gr. <i>παροικία</i> |
| | <i>parok</i> (selten) ‘Pfarrer; orthodoxer Priester’ | 1 | A | < klat. <i>parochia</i> < gr. <i>παροικία</i> |
| | <i>žūpnīk</i> ‘Pfarrer’ | 1 | B | ursl. <i>*džawpānu</i> < awar. |
| | <i>plòvan</i> (arch., kolloquial) ‘Pfarrer’ | 1 | A | < friul./dalm. <i>plevan</i> < lat. <i>plebānus</i> |
| | <i>pleban</i> (kolloquial) ‘Pfarrer’ | 1 | A | < lat. <i>plebānus</i> |
| | <i>plèbānuš (plebānuš)</i> (kajk., kolloquial) ‘Pfarrer’ | 1 | A | < ung. < lat. <i>plebānus</i> |
| holy | <i>svēt</i> ‘heilig’ | 5 | - | ursl. <i>*swentu</i> , idg. <i>*kuejt-</i> |
| | <i>sàkrālan</i> ‘sakral, heilig’ | 1 | B | < mlat. <i>sacer</i> |
| to preach | <i>propovijèdati</i> ‘predigen, verkünden’ | 1 | D | idg. <i>*uejd-</i> (Calque lat. <i>praedicāre</i>) |
| | <i>navijèstīti, naviještati</i> ‘predigen, verkünden’ | 5 | - | idg. <i>*uejd-</i> |
| | <i>predikati, predikovati, prodikovati</i> ‘predigen’ | 1 | C | < ahd. <i>bredigōn, predigon, prediga</i> < lat. <i>praedicāre</i> |
| | <i>moralizīrati</i> (pejor.) ‘predigen, moralisieren’ | 1 | C | < lat. <i>morālis</i> |
| | <i>pròdika, prèdika, prìdika</i> ‘Predigt’ | 1 | A | < ahd. <i>bredigōn, predigon, prediga</i> < lat. <i>praedicāre</i> |

| | | | | |
|----------|--|---|-------|---|
| | <i>pröpovijēd</i> ‘Predigt’ | 3 | C | idg. * <i>uejd-</i> (Ableitung eines Calques) |
| | <i>hōmīlija</i> ‘(Teil der) Predigt’ | 1 | A (B) | < lat. <i>homīlia</i> < gr. <i>ὁμιλία</i> |
| | <i>mīsa</i> ‘kath. Messe’ | 1 | A | < lat. <i>missa</i> |
| | <i>māša</i> (kajk.), <i>měša</i> (čak.) ‘Messe’ | 1 | A | < friul. oder ahd. <i>měssa, missa</i> < lat. <i>missa</i> |
| | <i>slūžba bōžjā</i> oder <i>bogoslužja</i> ‘Gottesdienst’ | 1 | D | * <i>klou-</i> (?); ursl. * <i>bagu</i> (< iran. <i>baga-</i> ?) (Calque gr. (<i>θεία</i>) <i>λειτουργία</i> oder nhd. <i>Gottesdienst</i> ?) |
| | <i>òbred</i> ‘Gesamtheit der religiösen Tätigkeiten; Zeremonie’ | 5 | - | nachursl. * <i>rědbъ</i> , bsl. * <i>rind-</i> , idg. ev. * <i>ar-</i> (?) |
| | <i>litūrgija, lejturgija, letūrgija</i> ‘Messe’ | 1 | A (B) | < gr. <i>λειτουργία, λειτουργῶ</i> |
| | <i>oficij, oficije</i> (čak.), <i>ofičije, ofičji</i> (dial.) ‘Messe, Gottesdienst’ | 1 | A (B) | < lat. <i>officium</i> bzw. ital. oder venez. |
| to bless | <i>blagoslōviti</i> ‘segnen’ | 1 | D | ursl. * <i>balgu</i> , idg. * <i>b^herg^h</i> oder * <i>bhel(g)-</i> (?); ursl. * <i>slawa</i> , idg. * <i>klej-</i> (Calque gr. <i>εὐλογέω</i> oder lat. <i>benedīcere</i>) |
| to curse | <i>prokleti, proklīnjati</i> ‘verfluchen, verwünschen’ | 5 | - | nachursl. * <i>kleti</i> , idg. * <i>klej-</i> (?) |
| | <i>ūklēti</i> ‘verfluchen’ | 5 | - | nachursl. * <i>kleti</i> , idg. * <i>klej-</i> (?) |
| | <i>kunuti, kumući</i> ‘fluchen’ | 5 | - | nachursl. * <i>kleti</i> , idg. * <i>klej-</i> (?) |
| | <i>anatemizirati, anatemati, anatemisati</i> ‘verfluchen’ | 1 | C | < gr. <i>ἀνάθεμα</i> |
| | <i>proklétstvo</i> ‘Fluch’ | 5 | - | nachursl. * <i>kleti</i> , idg. * <i>klej-</i> (?) |
| | <i>anàtema, ànàtema</i> ‘Fluch’, auch ‘Kirchenbann, Exkommunikation’ | 1 | A | < gr. <i>ἀνάθεμα</i> |
| | <i>ùrok</i> ‘Verwünschung, Verzauberung’ | 5 | - | nachursl. * <i>rekti</i> , idg. * <i>rek-</i> |
| to fast | <i>pōst</i> ‘Fasten’ | 1 | A | < ahd. <i>fasta</i> < lat. <i>fastus</i> ; got. <i>fastubni</i> ? |
| | <i>pòstiti</i> ‘fasten’ | 1 | B (C) | < ahd. <i>fastēn</i> < lat. <i>fastus</i> ; got. (<i>ski</i>) <i>fastan</i> ? |
| | <i>žěžīn</i> (čak.) ‘Fasten’ | 1 | A | < lat. <i>iēiūnāre, iēiūnus</i> |
| | <i>žěžīnāt</i> (čak.) ‘fasten’ | 1 | B (C) | < lat. <i>iēiūnāre, iēiūnus</i> |
| | <i>gòvjeti</i> (arch.) ‘*fasten’ | 5 | - | ursl. * <i>gawē iēj</i> , idg. * <i>g^{uh}om-</i> (oder * <i>g^{uh}ed^h-</i> ?) |
| | <i>kòrizma</i> ‘40-tägiges Fasten’ | 1 | A | < istr./friul. < vlat. < lat. <i>Quadragesima</i> (< Calque gr. <i>τεσσαρακοστή</i>) |
| | <i>korisma</i> (dial.) ‘40-tägiges Fasten’ | 1 | A | < dalm. < vlat. < lat. <i>Quadragesima</i> (< Calque gr. <i>τεσσαρακοστή</i>) |
| | <i>četrdesètnica</i> ‘40-tägiges Fasten’ | 1 | D | Calque lat. <i>quadragesima</i> (* <i>kjetur-</i> + * <i>dekm-t-</i>) |
| | <i>nēmrs</i> ‘Tag, an dem bzw. Zeit, in der man kein Fleisch isst (im Gegensatz zum Fasten)’ | 5 | - | (ungesichert, vgl. aksl. <i>mr̥siti sę</i>) |
| | <i>ùmjerenōst</i> ‘Enthaltsamkeit’ | 5 | - | nachursl. * <i>měra</i> , idg. * <i>meh₁-</i> |
| | <i>suzdržavanje</i> (kolloquial) ‘Entziehung, Mäßigung’, <i>suzdržati</i> ‘zurückhalten’ | 5 | - | nachursl. * <i>držati</i> , idg. * <i>dher-</i> |
| heaven | <i>něbo (neběsa)</i> ‘Firmament; Himmelreich’ | 4 | E (?) | ursl. * <i>neba</i> , idg. * <i>neb^h-es-</i> , * <i>neb^hos-</i> (Bedeutungsübertragung?) |
| | <i>rāj</i> ‘Paradies, Aufenthaltsort der Seligen’ | 1 | A | (ungesichert, wohl Entlehnung aus einer iran. Sprache) |
| | <i>džènet</i> ‘Paradies’ | 1 | A | < türk. <i>cennet</i> < arab. <i>ġennet</i> |
| | <i>paradis</i> ‘Paradies’ | 1 | A | < gr. <i>παράδεισος</i> |
| hell | <i>pàkao</i> ‘Hölle, Inferno’ | 3 | E | idg. * <i>pejH-</i> (Bedeutungsübertragung ahd.) |

| | | | | |
|---|--|-------|---|---|
| | <i>pandēmōnīj</i> (<i>pandemōnīj</i>) 'Sitz der Dämonen, Hölle' | 1 | A (B) | < gr. <i>πανδαμόνιον</i> |
| | <i>dže(h)ènem</i> , <i>džehènnem</i> , <i>džènem</i> 'Hölle' | 1 | A | < türk. < arab. <i>ǧehenem</i> < hebr. <i>gē ben Hinnom</i> |
| demon / devil | <i>zlòduh</i> 'Dämon, böser Geist' | 5 | - | ursl. <i>*dōxu</i> , idg. <i>*ǵ^huer-</i> |
| | <i>démon</i> 'böser Geist' | 1 | A | < (k)lat. <i>daemon</i> < gr. <i>δαίμων</i> |
| | <i>génīj</i> 'guter/böser Genius; Schutzgeist' | 1 | A | < lat. <i>genius</i> |
| | <i>džin</i> 'Dämon; guter oder böser Geist' | 1 | A | < türk. <i>cin</i> < arab. <i>ǧinn</i> |
| | <i>Levijàtan</i> '(bibl.) böser Geist' | 1 | A | < hebr. <i>liwjatan</i> |
| | <i>vàmpīr</i> , <i>upir</i> 'Vampir, Blutsauger' | 5 (1) | - (A) | < nordtürk. <i>ubyr</i> , <i>ubyrly</i> oder idg. <i>*(s)per-</i> |
| | <i>nēmān</i> 'großes, furchterregendes Phantasietier aus Märchen' | 5 | - | idg. <i>*mā-</i> |
| | <i>vrâg</i> 'Teufel' | 3 | E | ursl. <i>*wargu-</i> , idg. <i>*(h₁)- uorg-o-</i> (Bedeutungsübertragung) |
| | <i>đavao</i> , <i>ďävō</i> 'Teufel, Satan' | 1 | A | < gr. <i>διάβολος</i> , lat. <i>diabolus</i> , ahd. <i>tiufal</i> ? |
| | <i>satan</i> 'Teufel' | 1 | A | < gr. <i>σατᾶν</i> , <i>σατανᾶς</i> < hebr. <i>sātān</i> |
| | <i>sòtona</i> (dial.), <i>sotòna</i> 'Teufel' | 1 | A | < gr. <i>σατανᾶς</i> < hebr. <i>sātān</i> (bzw. über lat. <i>satanas</i>) |
| | <i>šéjtan</i> (bosn.) 'Teufel' | 1 | A | < türk. <i>şeytan</i> < arab. <i>šaitan</i> < äthiop. <i>sajeṭān</i> |
| | <i>hūd</i> , <i>hudoba</i> , <i>hūdīc</i> (čak., arch.) 'Teufel' | 3 | E | idg. <i>*k^(u)seud-</i> (Bedeutungsübertragung) |
| | <i>zlòca</i> 'Teufel' | 5 | - | idg. <i>*ǵ^huer-</i> |
| | <i>paklenica</i> 'Teufel, Teufelin' | 5 | - | idg. <i>*pejH-</i> |
| | <i>pàklenjāk</i> 'Teufel; Verurteilter in der Hölle' | 5 | - | idg. <i>*pejH-</i> |
| | <i>kūsī</i> (sr.) 'Teufel' | 5 | - | nachursl. <i>*kōsъ</i> , bsl. <i>*kanъd-</i> , idg. <i>*k^(u)ond-</i> oder <i>*k^(u)onHd^(h)-</i> |
| | <i>àntikrist</i> 'Antichrist, Teufel' | 1 | A (C?) | < lat. <i>antichristus</i> < gr. <i>ἀντί, χριστός</i> (Bestandteile kommen auch einzeln im Kr. vor) |
| | <i>Màmon</i> '(bibl.) Satan, Widersacher, böser Gott' | 1 | A | < lat. < gr. < hebr. < aram. <i>mammōn</i> oder <i>māmōna</i> |
| | <i>bèlāj</i> 'Unglück, Leid', 'Teufel' | 1 | A | < türk. <i>bela</i> < arab. <i>belā</i> |
| | <i>belzèbub</i> 'Teufel' (vgl. a. <i>Vřzin</i>) | 1 | A | klat. <i>Beelzebūb</i> < gr. <i>Βεελζεβοὺλ</i> < hebr. <i>Ba'al zewūl</i> oder <i>Ba'al zewūw</i> |
| | <i>nāpasnik</i> , <i>nāpast</i> 'Versucher, Teufel' | 5 | - | nachursl. <i>*pāsti</i> , idg. <i>*pōd-</i> |
| | <i>pali andeo</i> 'gefallener Engel' | 5 | - | nachursl. <i>*pāsti</i> , idg. <i>*pōd-</i> ; gr. <i>ἄγγελος</i> |
| <i>nečāstivī</i> , <i>nečāstīvník</i> 'Teufel' | 5 | - | nachursl. <i>čьtо, čisti</i> , idg. <i>*k^hejt-</i> | |
| <i>rèpat</i> 'mit einem langen Schwanz oder Schweif versehen' | 5 | - | nachursl. <i>репъ</i> (idg. Wurzel ungesichert) | |
| <i>cřnī</i> 'Teufel' | 5 | - | nachursl. <i>*čьrнъ</i> , bsl. <i>*kirsnos</i> , idg. <i>*krs-no-</i> | |
| idol | <i>kumir</i> 'Götzenstatue, Idol' | 2 | A | < ev. Russismus < aksl. <i>kumirъ</i> (ev. < syr. <i>kumrā</i> 'Priester' oder arm. <i>k'urm</i>) |
| | <i>kīp</i> 'Gestalt; Form, Bild; Idol, Götze' | 1 | A | < ung. <i>kép</i> |
| | <i>fètiš</i> 'verehrtes heiliges | 1 | A | < frz. <i>fétiche</i> < port. <i>feitiço</i> < lat. <i>facitūsus</i> |

| | | | | |
|------------------------------|---|---|---|---|
| | Objekt mit übernatürlichen Kräften' | | | |
| | <i>simulàkr(um)</i> 'Götze, Statue' | 1 | A | < lat. <i>simulacrum</i> |
| | <i>poklónac</i> '(verehrte) Statue' | 5 | - | idg. * <i>k̑lej-</i> (?) |
| | <i>tòtēm</i> 'verehrtes Totemtier bzw. dessen Symbol' | 1 | A | < amerik. Engl. < indigene amerik. Sprache |
| | <i>ìdol</i> 'Götze' | 1 | A | < lat. <i>ìdolum</i> < gr. <i>εἶδωλον</i> |
| magic | <i>bàjati</i> 'zaubern, Magie ausüben, Geister beschwören' | 5 | - | idg. * <i>b^heh₂-</i> |
| | <i>bajanje</i> (pejor.) 'Zauberei' | 5 | - | idg. * <i>b^heh₂-</i> |
| | <i>bàhoriti</i> 'zaubern' | 5 | - | idg. * <i>b^heh₂-</i> |
| | <i>bahorenje</i> 'Zauberei' | 5 | - | idg. * <i>b^heh₂-</i> |
| | <i>vràcati</i> 'Zaubersprüche hersagen' | 5 | - | idg. * <i>uer-</i> , *(<i>h</i> ₁)- <i>uorg-o-</i> ? |
| | <i>vračanje, vràdžbina, vračarija</i> 'Zauberei' | 5 | - | idg. * <i>uer-</i> , *(<i>h</i> ₁)- <i>uorg-o-</i> ? |
| | <i>čâr</i> 'Zauber, Zaubersprüche hersagen'; <i>čaròlija, čaròbija</i> 'Zauberei, Hexerei', <i>čârnost, čârnošt, čârka</i> 'Zauber' | 5 | - | idg. * <i>k^uer-/k^uēr-o-</i> |
| | <i>čárati</i> 'zaubern' | 5 | - | idg. * <i>k^uer-/k^uēr-o-</i> |
| | <i>čîni</i> 'Zauberei, Verwünschung, Zaubersprüche hersagen' | 5 | - | nachursl. * <i>čîmъ</i> , idg. * <i>k^uei-no-</i> |
| | <i>opčínjavanje</i> 'Zauberei', <i>opčínjavati</i> 'bezaubern', <i>opčíniti</i> 'verzaubern' | 5 | - | nachursl. * <i>čîmъ</i> , idg. * <i>k^uei-no-</i> |
| | <i>gátati</i> 'zaubern, wahrsagen' | 5 | - | idg. * <i>g^hed^h-</i> |
| | <i>magija, màdija</i> 'Magie, Zauberei' | 1 | B | < gr. <i>μαγεία</i> oder lat. <i>magia</i> |
| | <i>còprati</i> 'zaubern' | 1 | B (C) | < ahd. <i>zoubar</i> (nhd. <i>Zauber</i>) |
| | <i>còprija (coprija)</i> (pejor.) 'Zauberei' | 1 | B (C) | < ahd. <i>zoubar</i> (nhd. <i>Zauber</i>) |
| | <i>okultizam</i> 'Okkultismus' | 1 | A (B) | < lat. <i>occultus</i> |
| | <i>negromàntija, negromàncija, nekromantija</i> 'Zauberei, Wahrsagerei' | 1 | C (B) | < mlat. <i>necromantes</i> < gr. |
| | <i>spiritizam</i> 'Spiritismus; Glaube an Kommunikation m. Geistern oder übernatürlichen Wesen' | 1 | A (B) | < lat. <i>spiritismus</i> |
| | <i>opsjenárstvo</i> (pejor.) 'Scharlatanerie', <i>òpsjena</i> 'Täuschung' | 5 | - | nachursl. * <i>sěmъ</i> , idg. * <i>skoHi-n(-i)-</i> |
| | <i>ùrok (ùroci)</i> 'Zauberei, Verwünschung' | 5 | - | nachursl. * <i>rekti</i> , idg. * <i>rek-</i> |
| | <i>pròreći, proricati</i> 'weissagen, orakeln' | 5 | - | nachursl. * <i>rekti</i> , idg. * <i>rek-</i> |
| | <i>proricanje</i> 'Wahrsagerei' | 5 | - | nachursl. * <i>rekti</i> , idg. * <i>rek-</i> |
| | <i>zazivanje</i> 'Beschwörung' | 5 | - | nachursl. * <i>zъvati</i> , idg. * <i>g^hueH-</i> |
| <i>zazívati</i> 'beschwören' | 5 | - | nachursl. * <i>zъvati</i> , idg. * <i>g^hueH-</i> | |
| <i>čùdo</i> 'Wunder' | 5 | - | nachursl. * <i>čùdo</i> , idg. * <i>keud-os-</i> | |
| sorcerer or witch | <i>màdionĭk, mađionĭčar, magičar, mág</i> 'Zauberer' | 1 | C | < gr. <i>μαγεία</i> oder lat. <i>magia</i> |

| | | | | |
|---------------|--|---|--------|---|
| | <i>mađionica, mađioničarka</i> 'Zauberin' | 1 | C | < gr. <i>μαγεία</i> oder lat. <i>magia</i> |
| | <i>vrāc</i> 'Wahrsager, Magier' | 5 | - | idg. * <i>uer-</i> , *(<i>h</i> ₁)- <i>uorg-o-</i> ? |
| | <i>vrāčara, vračarica</i> 'Wahrsagerin' | 5 | - | idg. * <i>uer-</i> , *(<i>h</i> ₁)- <i>uorg-o-</i> ? |
| | <i>gatāra, gātālica</i> 'Wahrsagerin' | 5 | - | idg. * <i>g^hed^h-</i> |
| | <i>gātālac</i> 'Wahrsager' | 5 | - | idg. * <i>g^hed^h-</i> |
| | <i>čarobnjak, čarobnik</i> 'Zauberer' | 5 | - | idg. * <i>k^her-</i> /* <i>k^hēr-o-</i> |
| | <i>čarobnica</i> 'Zauberin' | 5 | - | idg. * <i>k^her-</i> /* <i>k^hēr-o-</i> |
| | <i>čaratār, čaratārica</i> 'Zauberer, Zauberin' | 1 | B (C) | < ital. <i>cerretano</i> (ev. Überkreuzung m. ital. <i>ciarlatano</i>) |
| | <i>žrēc</i> 'Schamane, Zauberer' | 1 | A | < russ. <i>žrec</i> |
| | <i>vilēnjāk</i> 'Hexer, Zauberer' | 5 | - | idg. * <i>uieh₁-</i> (?) oder idg. * <i>ueid-</i> (?) |
| | <i>iluziōnist</i> 'Magier' | 1 | C (B) | < lat. <i>illūsiō</i> |
| | <i>iluziōnistkinja</i> 'Magierin' | 1 | C | < lat. <i>illūsiō</i> |
| | <i>šāmān</i> 'Schamane, jmd. der Macht über andere besitzt; Spiritist' | 1 | A | < nhd. <i>Schamane</i> < russ. < sanskr. <i>sehama</i> |
| | <i>opsjēnār</i> (pejor.) 'Zauberer' | 5 | - | nachursl. * <i>sěnb</i> , idg. * <i>skoHi-n(-i)-</i> |
| | <i>negrōmant, nekrōmant</i> 'Zauberer' | 1 | A | < mlat. <i>necromantes</i> / ital. <i>negromante</i> < gr. |
| | <i>negrōmantkinja, negrōmantica</i> 'Zauberin' | 1 | C | < mlat. <i>necromantes</i> / ital. <i>negromante</i> < gr. |
| | <i>okūltist (okultist)</i> 'jmd. der Okkultismus betreibt' | 1 | A | < lat. <i>occultus</i> |
| | <i>sprītist (spiritist)</i> 'jmd. der sich mit Spiritismus beschäftigt' | 1 | A | < lat. <i>spiritismus</i> |
| | <i>sprītistkinja (spiritistica)</i> 'jmd. der sich mit Spiritismus beschäftigt' | 1 | C | < lat. <i>spiritismus</i> |
| | <i>švarckinstler</i> (kolloquial) 'Zauberer' | 1 | A | < nhd. <i>Schwarzkünstler</i> |
| 'Hexe' | <i>vjěštica</i> 'Hexe' | 5 | - | idg. * <i>ueid-</i> |
| | <i>vjěštac</i> 'Hexer' | 5 | - | idg. * <i>ueid-</i> |
| | <i>cōprnica (cōprnjica), cōparica, cōparnjica</i> (kolloquial) 'Hexe' | 1 | B | < nhd. <i>Zauberin</i> bzw. ahd. <i>zoubar</i> |
| | <i>štrīga</i> 'Hexe' | 1 | A | < ital. <i>strega</i> , friul. <i>stria</i> oder dalm. < lat. <i>strīga</i> |
| | <i>babarōga</i> 'Hexe; erfundenes Wesen, vor dem sich Kinder fürchten' | 5 | - | nachursl. * <i>bāba</i> , bsl. * <i>bá?ba?</i> , kr. <i>rōg</i> , idg. * <i>(h₂)reg^h-</i> |
| fairy or elf | <i>vīla</i> 'Fee' | 5 | - | idg. * <i>uieh₁-</i> (?) oder idg. * <i>ueid-</i> (?) |
| | <i>vilēnica</i> 'Fee', 'Ort, an dem Feen wohnen' | 5 | - | idg. * <i>uieh₁-</i> (?) oder idg. * <i>ueid-</i> (?) |
| | <i>rūsālka</i> 'Fee; Luftgeist', <i>rūsāljka</i> 'Fee; jmd. der bei religiösen Riten in Trance fällt' | 1 | A | < russ. < balkanrom. <i>rosalis</i> < lat. |
| | <i>gōrkinja</i> (kolloquial) 'Bergfee' | 5 | - | nachursl. * <i>gorà</i> , bsl. * <i>gor-/gir-</i> , idg. * <i>g^hōrHs</i> |
| | <i>jězērkinja</i> 'Fee (in Volkssagen)' | 5 | - | nachursl. * <i>ězero/ězerb</i> , bsl. * <i>ězero</i> , idg. * <i>h₁éǵ^h-(e)r-o-</i> |
| | <i>vodārica</i> 'Wasserfee' | 5 | - | ursl. * <i>wadā</i> , idg. * <i>uód-r</i> |
| ghost / angel | <i>āvēt, avētinja</i> 'Geist, Gespenst, Phantom' | 1 | A (C?) | < türk. < arab. <i>afēt</i> (oder nachursl. * <i>aviti</i> , bsl. * <i>āv-</i> als Entlehnung aus einer iran. Sprache, idg. * <i>h₂ēu-ēi-</i> oder * <i>h₁eu-</i>) |

| | | | | |
|---------------------|---|-------|--------|---|
| | <i>utvara</i> ‘Gespenst’ | 5 | - | idg. * <i>t₁erH-</i> |
| | <i>băuk</i> ‘Gespenst’ | 5 | - | idg. Schallwurzel (vgl. kr. <i>bău bău</i>) |
| | <i>voděnjāk</i> ‘Wassergeist; Wassermann’ | 5 | - | ursl. * <i>wadā</i> , idg. * <i>uód-r</i> |
| | <i>prividěnje</i> ‘Erscheinung, Geist’ | 5 | - | idg. * <i>uejd-</i> |
| | <i>vedenie</i> (kirchl.) ‘Vision’ | 4 | C (?) | idg. * <i>uejd-</i> (Überkreuzung mit lat. <i>vedea</i>) |
| | <i>săblast</i> ‘Erscheinung’ | 5 | - | (ungesichert, vgl. kr. <i>blăzan</i>) |
| | <i>pričina</i> (selten) ‘Gespenst, Erscheinung’ | 5 | - | nachursl. * <i>čīnъ</i> , idg. * <i>k^uei-no-</i> |
| | <i>prikaza, prikazánje</i> ‘Erscheinung, Geist’ | 5 | - | nachursl. * <i>kāzāti</i> , idg. * <i>k^ueĕ-</i> |
| | <i>fāntōm</i> ‘Geist, Phantom’ | 1 | A | < frz. <i>fantôm</i> < gr. <i>φάντομα</i> |
| | <i>fantăzma, fantăzam</i> ‘Erscheinung, Geist’ | 1 | A | < gr. <i>φάντομα</i> |
| | <i>simulăkr(um)</i> ‘Erscheinung, Geist’ | 1 | A | < lat. <i>simulacrum</i> |
| | <i>snòvid, snòviděnje</i> ‘Erscheinung im Traum, Gespenst’ | 5 | - | nachursl. * <i>snъnъ</i> , bsl. * <i>su(o)pnum</i> , idg. * <i>sup-n-o</i> ; idg. * <i>uejd-</i> |
| | <i>dūh</i> (kirchl.) ‘Seele, Geist’ | 5 | - | ursl. * <i>dōxu</i> , idg. * <i>d^hues-</i> |
| | <i>āńdeo</i> ‘Engel’ | 2 | B | < gr. <i>ἄγγελος</i> |
| | <i>serafim, serafin</i> ‘sechsflügeliger Engel’ | 1 | A | < lat./gr. < hebr. <i>serāfim</i> |
| | <i>heruvim</i> ‘dreiflügeliger Engel’ | 1 | A | < klat. <i>cherūbīm</i> < hebr. <i>kerūbhīm</i> |
| | <i>kerūbīn</i> ‘dreiflügeliger Engel’ | 1 | A | < ital. <i>cherubino</i> < mlat. <i>cherubinus</i> < klat. <i>cherūbīm</i> < hebr. <i>kerūbhīm</i> |
| omen | <i>znāměnje</i> ‘Zeichen; Omen’ | 5 | - | idg. * <i>ġneh₃-men-</i> |
| | <i>prědzńāk</i> ‘Vorzeichen’ | 5 | - | idg. * <i>ġneh₃-</i> |
| | <i>pretkazanje</i> ‘Vorzeichen, Omen’ | 5 | - | nachursl. * <i>kāzāti</i> , idg. * <i>k^ueĕ-</i> |
| | <i>pòjava</i> ‘Vorzeichen, Omen’ | 5 (3) | - (B?) | nachursl. * <i>aviti</i> , bsl. * <i>āv-</i> als Entlehnung aus einer iran. Sprache, idg. * <i>h₂ēu-ēi-</i> oder * <i>h₁eu-</i> |
| | <i>pěgula, pěgola</i> (čak.) ‘Pech, schlechtes Omen’ | 1 | A | < ital. <i>pégola</i> < lat. <i>pīcula</i> |
| | <i>omen</i> ‘Omen’ | 1 | A | < lat. <i>ōmen</i> |
| | <i>kōb</i> ‘Omen’ | 4 | A (?) | nachursl. * <i>kobъ</i> , ursl. * <i>kabi</i> (< voridg. Substrat?) |
| circumcision | <i>sūnet, sūnečěnje</i> ‘Beschneidung’ | 1 | A | < türk. < arab. <i>sūnnet</i> |
| | <i>sūnetiti</i> ‘beschneiden’ | 1 | C | < türk. < arab. <i>sūnnet</i> |
| | <i>ōbrezati</i> ‘beschneiden’ | 5 | - | idg. * <i>ureh₁ġ-</i> |
| | <i>obrezivanje</i> ‘Beschneidung’ | 5 | - | idg. * <i>ureh₁ġ-</i> |
| | <i>cirkumcīzija</i> ‘Beschneidung (religiös und medizinisch)’ | 1 | B | < lat. <i>circumcīsio, circumcidere</i> |
| initiation ceremony | <i>konfirmacija</i> (selten) ‘Firmung’ | 1 | A (B) | < lat. <i>confīrmatio</i> |
| | <i>krizma</i> ‘Firmung’ | 1 | A | < friul. < klat. <i>chrisma</i> < gr. <i>χρῖσμα</i> |
| | <i>krīsmā</i> (dial.) ‘Firmung’ | 1 | A | < dalm. < klat. <i>chrisma</i> < gr. <i>χρῖσμα</i> |
| | <i>sveta pōtvrda</i> bzw. <i>sakrament potvrde</i> ‘Konfirmation’ | 5 | - | nachursl. * <i>tvъrdъ</i> , idg. * <i>t₁orH-o-</i> ; ursl. * <i>swentu</i> |
| | <i>fīrma</i> ‘Firmung’ | 1 | A | < ital. <i>fīrma, fīrmare</i> < lat. <i>fīrmus</i> |
| | <i>fērma, bērma, bīrma</i> (čak.) ‘Firmung’ | 1 | A | < aital. <i>ferma, fermare</i> < lat. <i>fīrmus</i> |

5.2. Ausmaß von Lehn- und Erbwörtern

Eine genaue Auswertung des Ausmaßes an Lehn- und Erbwörtern im Kr. in der semantisch-lexikalischen Gruppe „Religion und Glaube“ ist erst dann sinnvoll, wenn auch Daten zur Vorkommenshäufigkeit der jeweiligen kr. Wörter vorliegen. Allerdings kann schon jetzt der Schluss gezogen werden, dass sich Lehnwörter eher bei Bezeichnungen finden, die mit dem institutionell verankerten Glauben oder dem bibl. Vokabular in Verbindung stehen und durch christl. Missionierung vermittelt wurden (obwohl sie im puristischen Kr. teilweise wieder ersetzt wurden, vgl. Kap. 6.2). Bei Bezeichnungen, die eher den abstrakten Glauben betreffen und die aus dem slaw. Heidentum übernommen und durch das Christentum entsprechend umgedeutet wurden, handelt es sich zumeist um Erbwörter. Auch Pronk-Tiethoff (2012: 236) konnte in ihrer Untersuchung der germ. Lehnwörter im (Nach-)Ursl. feststellen, dass „[...] the religious loanwords from germanic are clearly words relating to the Christian faith. This is in contrast with the Proto-Slavic loanwords from Iranian, which comprise more mythological and abstract religious terms, e.g. PSl. **rajъ* ‘heaven’, **bogъ* ‘god’ [...]” (Pronk-Tiethoff 2012: 236).

6. Probleme bei der Ermittlung des Lehnwortanteils des Kroatischen in der semantisch-lexikalischen Gruppe „Religion und Glaube“

Die Untersuchung des Lehn- und Erbwortanteils der kr. Entsprechungen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ zeigt (neben der in Kap. 2 und 5 besprochenen engen Lehnwortdefinition) Schwierigkeiten auf drei Ebenen. Die erste betrifft die Wortlisten an sich und die in bereits vorliegenden Untersuchungen angeführten Entsprechungen, die zweite das Problem der (auch durch Purismus bedingten) Synonymie, die dritte eine Identifizierung derjenigen Entsprechung, welche aufgrund ihrer Vorkommenshäufigkeit als Grundlage für den Vergleich mit anderen Sprachen herangezogen werden können.

6.1. Wortlisten und Bedeutungsentsprechungen in bereits vorliegenden Untersuchungen

Auf die Arbitrarität bei der Zuordnung von Bedeutungen zu, sowie auf den divergierenden Umfang von semantisch-lexikalischen Gruppen wurde bereits in Kap. 3.1 eingegangen. Hier sei noch angemerkt, dass es sich um die bei Buck (1949/³1971) angeführten Entsprechungen aufgrund der indogermanistischen Ausrichtung eher um Erbwörter handelt; Miklosich (1876) zieht v.a. ksl. Material (und somit viele Gräzismen) heran; in den in der vorliegenden Arbeit verwendeten Einzelstudien liegt der Schwerpunkt zudem meist auf Lehnwörtern aus einer bestimmten Sprache, wobei die tatsächliche heutige Verwendung der betreffenden Wörter bewusst außer Acht gelassen wird.

6.2. Synonymie und Purismus

Bei der Untersuchung des Lehn- und Erbwortanteils der kr. Entsprechungen in der semantisch-lexikalischen Gruppe „Religion und Glaube“ stellt sich das primäre Problem, dass hier in fast keinem Fall 1:1-Entsprechungen vorliegen, sondern sich oft eine Vielzahl an Synonymen⁶¹ im Kr. findet (welche tlw. nur in čak. und kajk. Dialekten auftreten oder tlw. – auch wiederbelebte – Archaismen sind).

Dieser Umstand dürfte zu einem Teil auch dem in der kr. Standardsprache vorherrschenden Sprachpurismus geschuldet sein, im Zuge dessen Lehn- durch Erbwörter ersetzt wurden (vgl. Wingender 1997)⁶². Dieser Purismus ist im Kr. bereits ab dem späten 16. Jh. durch die Abneigung kr. Lexikographen gegenüber Lehnwörtern und tlw. auch Internationalismen bei der Erstellung der ersten kr. Wörterbücher ersichtlich (vgl. Thomas 1988: 47, 83, 88). Im 17. und 18. Jh. entstanden hier viele Calques und (teils eigenwillige) Neologismen⁶³, die zunächst am Lat. und Ital. orientiert waren (vgl. Thomas 1988: 43), ab dem 19. Jh. wurden Calques von Internationalismen oder Germanismen, Bedeutungserweiterungen und -übertragungen präferiert (vgl. Thomas 1988: 83–87, 100–101). Lt. Thomas (1988: 100–103) ist die Sprachpolitik der illyrischen Periode eher pragmatisch statt dogmatisch, was sich auch an der Bevorzugung von Calques (aufgrund ihrer transparenten Bedeutung und der einfachen Bildung), Entlehnungen aus anderen slaw. Sprachen und Internationalismen ablesen lässt. Der illyrische Purismus manifestiert sich großteils in der Ablehnung volkssprachlicher Elemente (Dialektalismen, Regionalismen) und ist insgesamt zwar extensiv, aber moderat (vgl. Thomas 1988: 103).

In der NDH-Zeit (Unabhängiger Staat Kroatien 1941–1945) setzte sich (als Antwort auf die vorangehende unifizierende Serbokroatisierungspolitik) im Kr. ein v. a. die allgemeine Lexik, aber auch verschiedene Terminologien betreffender Sprachpurismus fort (vgl. Wingender 1997: 376). „Das sprachpolitische Ziel, der „čisti hrvatski jezik“, sollte durch das Ersetzen von Fremdwörtern, die Verbannung von Serbismen und das Wiederbeleben alter kroatischer Wörter erreicht werden.“

⁶¹ Deren Vorhandensein ist lt. Thomas (1988: 43–46) auch dem Umstand geschuldet, dass eine Fülle an Synonymen für kr. Lexikographen seit dem Barock bis ins 19. Jh. mit Sprachreichtum gleichbedeutend war. Interessant ist hierbei auch die Tatsache, dass lt. Thomas (1988: 100) die Vermeidung bzw. Ablehnung von Germanismen, Magyarismen, Turzismen und Italianismen während der illyrischen Periode zur Folge hat, dass diese in der heutigen kr. Standardsprache großteils als kolloquial erachtet werden.

⁶² Wingender setzt sich mit den Prinzipien des Umbaus der kr. Standardsprache auseinander, wobei ihr Schwerpunkt auf den in der Zagreber Tageszeitung *Vjesnik* seit 1994 erscheinenden (der Sprachpflege, der Unterscheidung des Kr. und Sr. v.a. auf lexikalischer Ebene und dem Kampf gegen Internationalismen und Fremdwörter dienenden, vgl. Wingender 1997: 382, 385–386) sog. Sprachratgebern und auf Standardisierungsvorschlägen in sprachwissenschaftlichen Zeitschriften wie *Jezik* liegt (vgl. Wingender 1997: 373). Auch zur Zeit der Illyrischen Bewegung (1830–1843) dienten die der *Danica ilirska* (die wöchentlich erscheinende Literaturbeilage der kr. Zeitung *Novine Hrvatske*) im Appendix angefügten Spezialglossare (*Sbirka*) zur Einführung und Erläuterung neuer Wörter der Stabilisierung des neuen Vokabulars (vgl. Thomas 1988: 28, 83, 105).

⁶³ Neologismen sind vermehrt vor und nach der illyrischen Periode anzutreffen und lt. Thomas eine extreme Form von Sprachpurismus (vgl. Thomas 1988: 71–72, 93). Sie sind z. B. im Dt. ein beliebtes Mittel der Puristen und wurden von den Protagonisten der illyrischen Periode wohl aus pragmatischen Gründen abgelehnt, da Neologismen oft idiosynkratisch gebildet werden und eine klare semantische Motivation in den meisten Fällen fehlt (vgl. Thomas 1988: 101–102).

(Wingender 1997: 377) Anfang der 70er Jahre wird die sog. (Re)Kroatiasierung („pohrvaćivanje“) der kr. Standardsprache durch sprachplanerische Tendenzen (Ablehnung von Serbismen, Russismen und Internationalismen, Wiederbelebung obsolet gewordener kr. Wörter⁶⁴) der Zeitschrift *Jezik* weitergeführt (vgl. Wingender 1997: 378–380). Lt. Wingender (1997: 387) äußert sich die durch Stigmatisierung gewisser sprachlicher Einflüsse geprägte Sprachpolitik in Kroatien weiterhin in (teilweise archaisierenden) puristischen Tendenzen.

Auch im Niedersorbischen herrscht ähnlich wie im Kr. ein Sprachpurismus vor, wobei in der Standardsprache Germanismen durch osorb. Äquivalente, nsorb. Neologismen oder wiederbelebte Erbwörter ersetzt werden. Da lt. Bartels (2009a) eine volle Ersetzung der Germanismen (in nsorb. Wörterbüchern meistens als kolloquial markiert) durch (ererbte oder entlehnte) slaw. Wörter jedoch in den wenigsten Fällen erfolgreich war, findet sich im Nsorb. in 40% aller Entsprechungen der „LWT meaning list“ ein synonyme Germanismus, in 8% aller Fälle eine synonyme Entlehnung aus dem Osorb. (vgl. Bartels 2009a: 316–317). Dem trägt Bartels (2009b) Rechnung, indem er in der WOLD auch die jeweiligen Synonyme anführt. Aufgrund der in Kap. 2 angesprochenen engen Definition eines Lehnworts im LWT-Projekt wurden einige möglicherweise aus dem Osorb. entlehnte Wörter bei Bartels jedoch nicht als Lehnwörter erachtet (vgl. Bartels 2009a: 311, Anm. 4).

Interessant ist in diesem Zusammenhang auch die Tatsache, dass Schuhmann (2009b) für das Althochdeutsche zwar fünf Lehnwörter (für ‘temple’, ‘church’, ‘altar’, ‘to preach’ und ‘to bless’, vgl. Schuhmann 2009a: 335) identifiziert hat, für jede Bedeutung in der semantisch-lexikalischen Gruppe „Religion und Glaube“ aber nur jeweils eine ahd. Entsprechung anführt. Lt. Sonderegger (2007: 45) liegen jedoch im Ahd. zu den Bedeutungen ‘church’, ‘sacrifice’, ‘to preach’, ‘demon’ und ‘god’ oft bis zu fünf (teils ererbte, teils entlehnte) Synonyme vor (z. B. *diabolus*, *tiuval/tievel*, *fiant/altfiant*, *widarwerto/widerwarto* und *unholdo* für ‘demon’). Die Aufnahme von Latinismen und Gräzismen in den Wortschatz des Ahd. war lt. Sonderegger (1997: 44) mit ein Hauptgrund für dessen Vereinheitlichung (lexikalischer Ausgleich im Spätahd.). Leider geht Schuhmann nicht weiter auf diese Uneinheitlichkeit der v.a. frühen ahd. Dialekte in Bezug auf Varianten ein und gibt auch nicht bekannt, anhand welchen Kriteriums die ahd. Entsprechungen ausgewählt wurden.

⁶⁴ Dies steht im klaren Kontrast zur illyrischen Periode, in welcher Entlehnungen aus anderen slaw. Sprachen – sofern sie „kroatiasiert“, also Lautsubstitutionen unterzogen und so der kr. Sprache angepasst wurden – als wichtige Impulse angesehen wurden (vgl. Thomas 1988: 73–74, 101). Dies ist darin begründet, dass der Illyrismus einerseits als lokale Manifestation eines Pan-Slawismus verstanden werden kann und v.a. das Č. Quelle und Inspiration der neuen kulturellen Identifikation war (vgl. Thomas 1988: 59), andererseits um die Abgrenzung gegenüber Österreich-Ungarn bemüht war (vgl. Thomas 1988: 17–18, 73). Internationalismen (v.a. Graeco-Latinismen) wurden als Alternative zu kr. Wörtern toleriert, da sie ein wichtiges Element für den Anschluss an die internationale Gelehrsamkeit darstellten (vgl. Thomas 1988: 87–89, 100, 103).

6.3. Vorkommenshäufigkeit, Archaismen und Kolloquialismen

Wie anhand der in Kap. 5.1 dargestellten Tabelle ersichtlich ist, sind für eine Reihe von Bedeutungen der „LWT meaning list“ in der semantisch-lexikalischen Gruppe „Religion und Glaube“ meist eine Vielzahl an kr. Entsprechungen anzuführen, von denen viele jedoch nicht standardsprachlich sind. Durch den in Kroatien vorherrschenden Sprachpurismus kommt es zudem zur Herausbildung vieler (tlw. stilistisch markierter) Synonyme.

In Anbetracht dessen stellt (neben der Identifizierung von Archaismen, Kolloquialismen und Dialektalismen) die Ermittlung der Vorkommenshäufigkeit jeder kr. Entsprechung ein wichtiges Forschungsdesiderat dar. Somit könnte auf dieser Grundlage mittels der jeweils frequentesten Entsprechungen auch der Lehnwortanteil in der semantisch-lexikalischen Gruppe „Religion und Glaube“ im Sinne des LWT-Projekts ermittelt und mit jenem der anderen Sprachen der WOLD verglichen werden.

Hierfür würden sich der *Hrvatski Čestotni Rječnik* (Moguš/Bratanić/Tadić 1999) und der *Hrvatski Nacionalni Korpus*⁶⁵ anbieten. Jedoch sei schon an dieser Stelle kritisch vorweggenommen, dass sich das auf dem Korpus des HČR (das sich recht ausgewogen aus Zeitungen, Dramen, Prosa, Gedichten und Lehrbüchern zusammensetzt, vgl. Moguš/Bratanić/Tadić 1999: 6, 11) basierendene Onlinekorpus *HNK* größtenteils aus Zeitungs- und Zeitschriftenartikeln (neben Belletristik) zusammensetzt⁶⁶ und sich hier die Frage der Repräsentativität stellt, wenn sprachplanerische Tendenzen in Kroatien vorrangig durch kr. Tageszeitungen getragen werden.

7. Literatur

- ANIĆ, ŠIME; KLAJČIĆ, NIKOLA; DOMOVIĆ, ŽELIMIR (2002): *Rječnik stranih riječi. Tuđice, posuđenice, izrazi, kratice i fraze*. Zagreb: Sani-plus.
- ANIĆ, VLADIMIR (2003): *Veliki rječnik hrvatskoga jezika*. Zagreb: Novi Liber.
- AUTY, ROBERT (1976): Lateinisches und Althochdeutsches im altkirchenslavischen Wortschatz. In: *Slovo* 25/26, 169–174.
- BARTELS, HAUKE (2009a): Loanwords in Lower Sorbian, a Slavic language of Germany. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter; 304–329.
- BARTELS, HAUKE (2009b): Lower Sorbian vocabulary. In: Haspelmath, Martin; Tadmor, Uri (eds.): *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Online-Ressource: <<http://wold.livingsources.org/vocabulary/10>> [aufgerufen am 14.03.2015])
- BUCK, CARL DARLING (1971)³: *A Dictionary of selected Synonyms in the principal Indo-European Languages. A Contribution to the History of Ideas*. Chicago – London: University of Chicago Press.

⁶⁵ <<http://www.hnk.ffzg.hr>>, das Korpus ist erreichbar unter <http://filip.ffzg.hr/bonito2/run.cgi/first_form> [aufgerufen am 14.03.2015].

⁶⁶Vgl. <<http://www.hnk.ffzg.hr/korpus.html>>, <<http://www.hnk.ffzg.hr/struktura.html>> [aufgerufen am 14.03.2015].

- DERKSEN, RICK (2008): *Etymological Dictionary of the Slavic Inherited Lexicon* (= Leiden Indo-European Etymological Dictionary Series 4). Leiden – Boston: Brill.
- DWDS = *Digitales Wörterbuch der deutschen Sprache* (Berlin-Brandenburgische Akademie der Wissenschaften). (Online-Ressource: <<http://www.dwds.de>> [aufgerufen am 27.7.2015])
- ERHSJ (Bd. I-III) = Skok, Petar (1971–1973): *Etimologijski rječnik hrvatskoga ili srpskoga jezika*. Zagreb: JAZU.
- HASPELMATH, MARTIN; TADMOR, URI (2009a) (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter.
- HASPELMATH, MARTIN; TADMOR, URI (2009b): The Loanword Typology project and the World Loanword Database. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 1–34.
- HASPELMATH, MARTIN (2009): Lexical borrowing: Concepts and issues. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter; 35–54.
- HČR = Moguš, Milan; Bratanić, Maja; Tadić, Marko (1999): *Hrvatski čestotni rječnik*. Zagreb: Školska knjiga – Zavod za lingvistiku filozofskog fakulteta sveučilišta u Zagrebu.
- HNK = Hrvatski Nacionalni Korpus v3.0 (<<http://www.hnk.ffzg.hr>>). Korpus: <http://filip.ffzg.hr/bonito2/run.cgi/first_form> [aufgerufen am 14.3.2015]
- HOLZER, GEORG ET AL. (2010): Lautgeschichtliches Glossar zum Neuštokavischen. In: *Wiener Slavistisches Jahrbuch* 56, 39–60.
- HOLZER, GEORG (2011): *Glasovni razvoj hrvatskoga jezika*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- HOLZER, GEORG ET AL. (2012): Lautgeschichtliches Glossar zum Neuštokavischen (weitere Lemmata). In: *Ricerche Slavistiche*, Nuova Serie 10, (56); 5–44.
- KATIČIĆ, RADOSLAV (1980): Slavica Foroiuliensia. In: *Wiener Slavistisches Jahrbuch* 26, 28–32.
- KATIČIĆ, RADOSLAV (1999): *Literatur- und Geistesgeschichte des kroatischen Frühmittelalters* (= Schriften der Balkan-Kommission. Philologische Abteilung 40). Wien: ÖAW.
- KRAHWINKLER, HARALD (1992): *Friaul im Frühmittelalter. Geschichte einer Region vom Ende des fünften bis zum Ende des zehnten Jahrhunderts* (= Veröffentlichungen des Instituts für österreichische Geschichtsforschung 30). Wien – Köln – Weimar: Böhlau.
- LIV = Rix, Helmut (2001): *Lexikon der indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen*. 2., erw. u. verb. Aufl. Wiesbaden: Reichert.
- MARCHETTI, GIUSEPPE (1985)⁴: *Lineamenti di grammatica friulana*. Udine: Società Filologica Friulana.
- MATASOVIĆ, RANKO (2008): *Poredbenopovijesna gramatika hrvatskoga jezika*. Zagreb: Matica Hrvatska.
- MEILLET, ANTOINE (1926): Le vocabulaire slave et le vocabulaire indo-iranien. In: *Revue des Études Slaves* 6, 3/4;165–174.
- MIKLOSICH, FRANZ (1876): Die christliche Terminologie der slavischen Sprachen: eine sprachgeschichtliche Untersuchung. In: *Denkschriften der kaiserlichen Akademie der Wissenschaften, Philosophisch-Historische Classe* 24, 1–58.

- PRONK-TIETHOFF, SASKIA (2012): *The Germanic loanwords in Proto-Slavic. Origin and accentuation*. Leiden: Univ. Diss.
- SCHUHMAN, ROLAND (2009a): Old High German. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 330–337.
- SCHUHMAN, ROLAND (2009b): Old High German vocabulary. In: Haspelmath, Martin; Tadmor, Uri (eds.): *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Online-Ressource: <<http://wold.clld.org/vocabulary/11>> [aufgerufen am 31.3.2015])
- SKOK, PETAR (1953): Uslovi života glagoljice. In: *Slovo* 3, 50–63.
- SONDEREGGER, STEFAN (1997): *Althochdeutsch als Anfang deutscher Sprachkultur* (= Vorträge Wolfgang Stammeler Gastprofessur für Germanische Philologie 2). Freiburg: Univ.-Verl.
- ŠARIĆ, LJILJANA; WITTSCHEN, WIEBKE (2008): *Rječnik sinonima hrvatskoga jezika*. Zagreb: Jesenski i Turk.
- TADMOR, URI (2009): Loanwords in the world's languages: Findings and results. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 55–75.
- THOMAS, GEORGE (1988): *The Impact of the Illyrian Movement on the Croatian Lexicon*. (= Slavistische Beiträge 223). München: Sagner.
- WINGENDER, MONIKA (1997): Sprachpolitik in Kroatien. Eine exemplarische Analyse der Sprachratgeber im Vjesnik. In: Schulze, Jana; Werner, Eduard (eds.): *Linguistische Beiträge zur Slavistik. V. JungslavistInnen-Treffen Bautzen 1996*. München: Sagner, 372–392.
- ZAGIBA, FRANZ (1961): Die bairische Slavenmission und ihre Fortsetzung durch Kyrill und Method. In: *Jahrbücher für Geschichte Osteuropas*. (N.F. Bd. 9/1961), 1–56.
- ZAGIBA, FRANZ (1964): Die Missionierung der Slaven aus „Welschland (Patriarchat Aquileja) im 8. und 9. Jahrhundert.“ In: Hellmann, Manfred et al. (eds.): *Cyrillo-Methodiana. Zur Frühgeschichte des Christentums bei den Slaven 863–1963*. Köln – Graz: Böhlau, 274–311.

Saskia Pronk-Tiethoff (Leiden)

The Methodology of the *Loanword Typology Project*: Opportunities for the Investigation of Loanwords in Slavic

Abstract

Der vorliegende Aufsatz vergleicht die Methodologie einzelner Datenbanken des *Loanword Typology Project* (Haspelmath/Tadmor 2009a) insbesondere des *Lower Sorbian vocabulary* (Bartels 2009) und des *Dutch vocabulary* (Van der Sijs 2009). Es wird versucht festzustellen, welche methodischen Ansätze für die Untersuchung von Lehnwörtern in den slawischen Sprachen geeignet sind. Zum Abschluss wird die semantische Einordnung der frühesten germanischen Lehnwörter im Urslavischen (Pronk-Tiethoff 2013) den semantischen Feldern und Wortklassen des *Loanword Typology Project* gegenüber gestellt (vgl. Tadmor 2009).

1. Introduction

The *Loanword Typology Project* (LTP) is an empirical study of lexical borrowability on the basis of a worldwide sample of 41 languages and a fixed list of 1460 lexical meanings. The project results in a separate database for each of the 41 languages included. The authors of these databases indicated for each word the likelihood that it was borrowed. In further consequence, the databases formed the basis for typological research into borrowability (Haspelmath/Tadmor 2009). The present paper investigates aspects of the project's methodology by comparing some of the databases, in particular the Lower Sorbian and Dutch vocabularies. I will discuss differences in the approaches and methods of working of the individual contributors to the project, which have resulted in diverging interpretations of the material. While this may not greatly influence the results of worldwide typological research such as the *Loanword Typology Project*, it may be more problematic when studying a number of closely related languages such as the Slavic languages. I will conclude by briefly comparing the findings of my research into the semantics of the Germanic loanwords in Proto-Slavic to the semantic classification of the *Loanword Typology Project*.

2. The *Loanword Typology Project*

Contributors to the project were given a list of almost 1,500 lexical meanings, for which they were asked to provide correspondences in the language under investigation (Haspelmath/Tadmor 2009b). The contributors then had to assess for each word:

1. the loanword status. The authors were requested to indicate whether a given word can be assumed to be a loanword. At that five degrees of certainty are distinguished:
 1. (4.) clearly borrowed

2. (3.) probably borrowed
 3. (2.) perhaps borrowed
 4. (1.) very little evidence for borrowing
 5. (0.) no evidence for borrowing.¹
2. the analyzability. If a word is morphologically analysable, it is considered to be almost certainly created by the speakers of the respective language. It is therefore less likely to be a loanword.
 3. the age of a word. The authors had to provide for each word the earliest date for which it is attested or can be reconstructed (i.e., the date of borrowing for loanwords, and the date of creation for non-loanwords) (Haspelmath/Tadmor 2009b: 10–14).

In the online glossary to the World Loanword Database a loanword is defined as follows:

A loanword is a word that was copied from another language, either by adoption or by retention, at some point in the history of the language. Even if a loanword is fully integrated, it is still a loanword, and a loanword never ceases to be a loanword.²

For Lower Sorbian (LS) as representative of the Slavic languages in the LTP, this definition quite correctly implies that words borrowed as early as Proto-Slavic are rated as loanwords in Lower Sorbian.

The *Loanword Typology Project* only takes lexical borrowing into account and explicitly excludes calques and neologisms created from foreign material. These words were created in the recipient language, instead of being transferred as a whole from a donor language. Calques and neologisms are labelled in the databases under the heading “no evidence for borrowing” and therefore belong to the same category as the inherited words. To me, this heterogeneous character of the category words with “no evidence for borrowing” is one of the major disadvantages of the database. In the Lower Sorbian database, this category contains words that can be shown to go back to Proto-Slavic, such as: *bratš* ‘brother’, *sotša* ‘sister’, *maś* ‘mother’. But it also contains many words that entered the language at a later date and under the influence of German: *starjejšej* ‘parents’ (a calque from German *Eltern*), *štifnan* ‘stepfather’ (borrowed + inherited element), *enkelsyn* ‘grandson’ (borrowed + inherited element), *šwigerstarjejšej* ‘parents-in-law’ (borrowed + calqued element). According to the rules of the project, these formations are placed in a different category than words like: *šwigerfoter* ‘father-in-law (of a man)’, *šwigermutter(ka)* ‘mother-in-law (of a woman)’, *grosmuter(ka)* ‘grandmother’. The latter group of words consist of two borrowed elements and are therefore considered to be loanwords. The databases show many such examples (cf. Tab. 1).

¹ The book and the database use different numberings to indicate the loanword status. The numbering in brackets refers to the numbering in the book.

² Cf. Glossary to the *World Loanword Database* <<http://wold.clld.org/terms>> [accessed on 2015-03-17].

Tab. 1: Examples of inconsequent labelling³

| 1. “clearly borrowed” | 5. “no evidence for borrowing” |
|---|---|
| LS <i>haj</i> ‘shark’ < G <i>Hai</i> ‘shark’ | LS <i>hajryba</i> ‘shark’ and similarly LS <i>walryba</i> ‘whale’, both of which consist of a borrowed and a calqued element. |
| LS <i>postkórta</i> ‘postcard’ < G <i>Postkarte</i> ‘postcard’ | LS <i>postowa kórtka</i> ‘postcard’ |
| LS <i>šejdunga</i> ‘divorce’ < G <i>Scheidung</i> ‘divorce’ | LS <i>šejdowanje</i> ‘divorce’ |
| OHG <i>mûl</i> , E <i>mule</i> , LS <i>mul</i> < Lat. <i>mulus</i> ‘mule’ | Du. <i>muilezel</i> (mule-ass), <i>muildier</i> (mule-animal) |
| OHG <i>pînboum</i> ‘pine’ < Lat. <i>pînus</i> ‘pine’ | Du. <i>pijnboom</i> ‘pine’ |

The attribution of calques and neologisms to the category of non-borrowed words significantly influences the general picture of foreign influence on a language. Bartels also considers the rules for defining a loanword as “relatively restrictive” (2009: 317).

The classification of words as loanwords or non-loanwords deserves attention in future projects along the lines of the *Loanword Typology Project*. It is clear that the calques, hybrids and words derived on a loan basis cannot be headed under the label “clearly borrowed” in a project aimed to investigate lexical borrowing. Nevertheless, classifying these words into a separate category with a heading such as “no evidence for borrowing, but not inherited either” may be useful. This would enable researchers working with the databases to separate the inherited lexicon from newer formations in a language.

On the basis of the number of pure lexical borrowings, the editors of the *Loanword Typology Project* consider Lower Sorbian to be an average borrower, with 22.4% of borrowed words. Lower Sorbian thus belongs to the same category as, e.g., Dutch, which has a percentage of 19.1% borrowed words (Tadmor 2009: 57).

The approach of taking only lexical borrowings into account is understandable, but it may create a false picture of the general amount of foreign influence on a language. The German influence on Lower Sorbian goes far beyond lexical borrowings. Its grammar includes many contact-induced phenomena, such as the regular use of articles, the use of verbal particles in word formation (LS *sobu-želiš* < G *mitteilen*) and the formation of the passive with the auxiliary verb *wordowaś* (Bartels 2009: 322–323). All in all, Lower Sorbian is a language heavily influenced by German, with respect to both its lexicon and its grammar. In the 19th century, the literary language was partly cleared of German lexical influence for purist reasons. This was done by the replacement of Germanisms either by neologisms or by Upper Sorbian loanwords (idem: 316). According to the rules of the project, the Upper Sorbian loanwords are regarded as loanwords, whereas the neologisms constructed from indigenous material are regarded as non-loanwords. Both the Lower Sorbian neologisms and the Upper Sorbian loanwords were introduced into 19th-century Lower Sorbian as a result of the same process. We can therefore assume a certain extent

³ Du. = Dutch, Lat. = Latin, LS = Lower Sorbian, OHG = Old High German.

of coincidence which word was replaced by an Upper Sorbian loanword and which by a neologism. In order to enable future researchers to collect from a database all words, be it a loanword or a neologism, that entered a language at a certain stage, I consider it important that the contact situation in which a word was borrowed or created should be included in the database.

The contributors to the LTP were asked to provide a dating for the words in the database. This has been specified as the *time of borrowing* for loanwords and the earliest *attestation or reconstruction* for non-loanwords. The age of a word in a language is important for a general investigation into borrowability because it has been supposed that the longer a word exists in a language, the more chance it has had to be replaced by another word. Older words form therefore more reliable evidence for resistance against borrowing than new words (Haspelmath/Tadmor 2009b: 13–14).

Significant methodological differences exist between the authors with regard to the dating of the loanwords. Bartels differentiates between words that go back to Proto-Slavic (400–1000 AD) and later formations, and he dates newer words approximately to the century of their entrance into Sorbian. These dates seem to be based on 1) comparative Slavic evidence (especially for the older loanwords), 2) the historical evidence we have on the contact situation in which a word was borrowed, and 3) the date of earliest attestation⁴.

Van der Sijs' approach for the Dutch database is quite different because she primarily departs from the earliest attestation of a word (Van der Sijs 2009).⁵ The date provided in the Dutch database is therefore the date of earliest attestation in a Dutch text, regardless of the etymology of the word. Dutch is reliably attested from the 13th century onwards. This means that almost no entries receive a dating⁶ earlier than the 13th century. Words with cognates in other Germanic languages, but which happen to be attested late in Dutch written sources, are given a late dating, even though the word may have existed in Proto-Germanic already. The word *donkey*, Dutch *ezel*, for example, was borrowed from Latin into Proto-Germanic, but on the basis of its earliest attestation it has been given the dating 1240 in the Dutch database. The Old High German database⁷ dates the same word to the 9th century (801–900). The Lower Sorbian database dates the word, which the Slavs borrowed from Germanic at an early stage, to Proto-Slavic (400–1000). Such differences between the different contributors to the project naturally influence the comparative findings and results of the project.

⁴ See Bartels 2009 <<http://wold.clld.org/vocabulary/10>> [accessed on 2015-03-17].

⁵ “My dating of the words is largely based on the information provided by the EWB and van der Sijs 2001. Most of the dates mentioned in these works come from the quotations given in the three big historical dictionaries of the Dutch language, the WNT, the MNW and the VMNW.” (for details see under ‘Field descriptions’ at <http://wold.clld.org/vocabulary/12>).

⁶ See Van der Sijs 2009 <<http://wold.clld.org/vocabulary/12>> [accessed on 2015-03-17].

⁷ Cf. Schuhmann 2009 <<http://wold.clld.org/vocabulary/11>> [accessed on 2015-03-17].

The word *klok* ‘clock’ is the only word in the Dutch database that has been attributed to Old Irish influence. It was introduced by Irish missionaries who were active in the Low Countries. This etymology dates the entrance of the word into Dutch to the 7th century, but the word nevertheless receives the dating 1237 in the database⁸, because of its earliest date of attestation.

In order to draw reliable conclusions on the basis of the material, it is important that all contributors to a larger project assess the material in the same manner. With regard to the dating of the loanwords, I consider Bartels’ etymology-based approach the most adequate.⁹

One of the aims of the project has been to find statistical evidence for long-standing questions relating to the borrowability of words, which until then could only be judged on the basis of pre-existing so-called “impressionistic observations” (Haspelmath/Tadmor 2009b: 1). These impressionistic observations include:

- body part terms are unlikely to be borrowed (ibid.),
- terms for new artefacts are often borrowed (ibid.),
- content words are more borrowable than function words (Tadmor 2009: 59),
- nouns are more borrowable than verbs (Tadmor 2009: 61),
- basic (“fundamental”) vocabulary is much less susceptible to borrowing than non-basic (“cultural”) vocabulary (Tadmor 2009: 65).

The *Loanword Typology Project* showed that these observations are generally correct, although the list of most borrowable words turned out to include more interrogatives and indicators for spatial organization, time and persons, than expected (Tadmor 2009: 66). The pre-existing observations are certainly true for the Germanic loanwords in Proto-Slavic. In my monograph (Pronk-Tiethoff 2013), I investigated the Germanic and German loanwords that entered Proto-Slavic until its final disintegration in the early 9th century AD. This corpus consists of 76 certain Germanic loanwords and 42 possible Germanic loanwords. It contains:

- no words for body parts,
- many terms for new artefacts,
- only content words,
- just nine verbs,
- and no words belonging to the basic vocabulary as defined by the final Swadesh 100-word list (1971).

⁸ Cf. Van der Sijs 2009 <<http://wold.clld.org/vocabulary/12>> [accessed on 2015-03-17].

⁹ At the workshop *Lehnwörter im Slawischen: Empirische und crosslinguistische Perspektive* in September 2014 in Vienna, Prof. Newerkla pointed out the importance of including the earliest attestations of a word as well because this may give important information about the use and spread of a word in early times.

3. Semantics of the Germanic loanwords in Proto-Slavic

Research into the semantic classification of the Germanic loanwords in Proto-Slavic provides information about the nature of the contacts between Proto-Slavs and Germanic peoples. Words are borrowed from other languages for numerous reasons, for example out of lexical need or due to social factors. The semantics of the loanwords therefore gives information about the social and cultural circumstances in which the words were borrowed. The Germanic loanwords in Proto-Slavic can be divided into the following semantic categories:

- power and warfare,
- skills and mental concepts,
- technical terminology,
- trade (divided into general trading terms, money and buying, words for containers),
- Christian terminology,
- words relating to the yard and home grown/made products,
- remaining words (Pronk-Tiethoff 2013: 251–255).

The semantic classification used in the *Loanword Typology Project* consists of 24 semantic fields (Haspelmath/Tadmor 2009b: 7–8). Some of these categories have a one-to-one correspondence to the categories I distinguished for the Germanic loanwords in Proto-Slavic, whereas in other cases words from a single semantic category in my classification correspond to different categories in the project's classification. The semantic classification in the *Loanword Typology Project* primarily departs from the meaning of the word, whereas I tried to connect the meaning of the word as far as possible to a type of linguistic exchange or specific contact situation.

The animals PS1. **osъlbъ* 'donkey', **velъblъqъdъ* 'camel' and **lvъbъ* 'lion', for example, belong in the project's classification to the same category "3. Animals". The Slavs were likely to have become acquainted with these animals under different circumstances. The Slavs came to know the donkey because of trading: the word for 'donkey' was borrowed from Latin into Germanic because the Romans used donkeys (and mules) to transport their wares overland and thus introduced the animal into northern Europe (Green 1998: 204).¹⁰ The word was subsequently borrowed from a Germanic language, probably Gothic, into Proto-Slavic as **osъlbъ* 'donkey'. The grounds for borrowing the word 'lion' into Slavic are unclear, but cannot be related to trade. The lion might have entered the Slavic lexicon along with the Bible translations. The contact situation in which the Slavs borrowed the word **velъblъqъdъ* 'camel' is unclear, but it has been thought that the Goths became acquainted with domesticated camels in the Lower

¹⁰ Compare in this respect also Croatian *tòvar* 'load, shipment', which dialectally means 'donkey'.

Danube region in the 4th century and subsequently introduced them in neighbouring areas (Hoops 1989: 141–142).

Another Germanic loanword belonging to the semantic category “Animals” in the project’s classification is PSI. **xlěvъ* ‘cattle shed, stable’. This is one of the words for a type of structure borrowed from Germanic, the other being PSI. **xyzъ, *xyžъ* ‘small house, cottage’. The latter word belongs according to the project’s semantic classification to the category “7. The house”. Both words represent types of structures the Slavs came to know and took over when they came to live in close proximity to Germanic peoples. I therefore think these words should be assigned to the same semantic category. In my research, I have therefore attributed both words to the semantic group of words relating to the farm yard. As mentioned before, I consider it important to relate the loanwords to the contact situation in which they were borrowed.

4. Germanic loanwords in Proto-Slavic according to the LTP semantic categories:

The following overview (cf. Tab. 2) shows the Germanic loanwords in Proto-Slavic according to the semantic categories of the *Loanword Typology Project*:¹¹

Tab. 2: Germanic loanwords in Proto-Slavic

| LTP semantic category | No. of Germanic loanwords | Germanic loanwords in Proto-Slavic ¹² | |
|-------------------------|---------------------------|--|---|
| 5. Food and drink | 15 | <i>*bljudo</i> ‘plate, dish’ <i>*bъdъnja, *bъdъnjъ</i> ‘tub’ <i>*kotъlъ</i> ‘kettle’ <i>*kusiti</i> ‘to try, taste’ <i>*kъbъlъ</i> ‘tub, quantity of grain’ <i>*lagy</i> ‘bottle, cask’ <i>*lukъ</i> ‘chive, onion’ <i>*qborъ(kъ)</i> ‘bucket; quantity of grain’ | <i>*orky</i> ‘box’ <i>*ovotjъ, *ovotje</i> ‘fruit’ <i>*redъky, *rъdъky</i> ‘radish, <i>Raphanus sativus</i> ’ <i>*stopa</i> ‘pestle, mortar’ <i>*vino</i> ‘wine’ <i>*vinogordъ</i> ‘vineyard’ <i>*xlěbъ</i> ‘loaf, bread’ |
| 11. Possession | 9 | <i>*dъlgъ</i> ‘debt’ <i>*gobina/*gobino</i> ‘wealth, abundance’ <i>*gobъdъžъ</i> ‘wealth, abundance’ <i>*kupiti</i> ‘to buy’ <i>*myto</i> ‘toll, payment’ | <i>*pěnędžъ</i> ‘penny, coin’ <i>*skъlędžъ, *stъlędžъ, *štъlędžъ</i> ‘coin’ <i>*vaga</i> ‘weight, scales’ <i>*lixva</i> ‘interest, usury’ |
| 22. Religion and belief | 8 | <i>*cъrky</i> ‘church’ <i>*goneznqiti</i> ‘to recover’ <i>*gonožiti</i> ‘to save’ <i>*krъstъ</i> ‘cross, Christ, | <i>*popъ</i> ‘clergyman, (Orthodox) priest’ <i>*postiti se</i> ‘to fast’/ <i>*postъ</i> ‘fast, Lent’ |

¹¹ The categories in Tab. 2 are arranged according to the number of words they contain. For a detailed discussion of the etymology and German equivalents of these words, I refer to Pronk-Tiethoff (2013).

¹² Words that can be placed in several categories are placed between brackets.

| | | | |
|------------------------------------|---|--|---|
| | | baptism’ | * <i>xr̥st̥</i> ‘cross, Christ, baptism’ |
| 16. Emotions and values | 7 | *(<i>u-</i>) <i>žasn̥oti</i> ‘to terrify, frighten’ * <i>l̥st̥</i> ‘cunning (trick)’ * <i>užas̥</i> ‘horror, amazement’ * <i>x̥dog̥</i> ‘skill (?)’ | * <i>xula</i> /* <i>xuliti</i> ‘(to) abuse, revile’ (* <i>gorazd̥</i> ‘experienced, able’) (* <i>duma</i> ‘advice, thought, opinion’) |
| 3. Animals | 6 | * <i>l̥v̥</i> ‘lion’ * <i>nuta</i> ‘cow, cattle’ * <i>os̥l̥</i> ‘donkey’ | * <i>skot̥</i> ‘livestock’ * <i>vel̥bl̥d̥</i> ‘camel’ * <i>xl̥v̥</i> ‘cattle shed, stable’ |
| 9. Basic actions and technology | 6 | * <i>lug̥</i> ‘lye, caustic soda’ * <i>neboz̥ž̥</i> /* <i>naboz̥ž̥</i> ‘wood drill’ * <i>pila</i> ‘saw, file’ | * <i>st̥klo</i> ‘glass (ware)’ (* <i>pet̥lja</i> ‘noose, snare’) (* <i>ret̥d̥ž̥</i> ‘chain(s)’) |
| 20. Warfare and hunting | 6 | * <i>br̥nja</i> ‘harness, suit of armour’ * <i>r̥lk̥</i> ‘regiment, crowd’ * <i>š̥lm̥</i> ‘helmet’ | * <i>x̥osa</i> ‘robbery, trap’ (* <i>pet̥lja</i> ‘noose, snare’) (* <i>ret̥d̥ž̥</i> ‘chain(s)’) |
| 8. Agriculture and vegetation | 5 | * <i>buk̥</i> ‘beech’ * <i>gr̥del̥j̥</i> ‘plough-beam, axis’ * <i>plug̥</i> ‘plough’ | * <i>tyn̥</i> ‘fence’ * <i>v̥rtogord̥</i> ‘garden’ |
| 19. Social and political relations | 5 | * <i>c̥sar̥j̥</i> , * <i>cesar̥j̥</i> , * <i>č̥sar̥j̥</i> ‘(Roman) emperor’ * <i>kor̥lj̥</i> ‘king’ * <i>k̥n̥d̥ž̥</i> ‘prince, ruler’ | * <i>vit̥d̥ž̥</i> ‘hero, knight’ * <i>vol̥x̥</i> ‘Romance-speaking person’ |
| 1. The physical world | 3 | * <i>kold̥d̥ž̥</i> ‘well, spring’ * <i>per̥gyn̥ja</i> ‘impenetrable covert (?)’ * <i>x̥lm̥</i> ‘hill’ | |
| 18. Speech and language | 3 | * <i>buky</i> ‘beech(nut), letter, book, document’ * <i>lik̥</i> ‘choir (?)’ * <i>tr̥oba</i> ‘trumpet’ ¹³ | |
| 6. Clothing and grooming | 2 | * <i>skut̥</i> ‘hem, clothing covering the legs’ * <i>user̥d̥ž̥</i> ‘earring’ | |
| 17. Cognition | 2 | (* <i>gorazd̥</i> ‘experienced, able’) (* <i>duma</i> ‘advice, thought, opinion’) | |
| 4. The body | 1 | * <i>l̥k̥</i> ‘medicine’ | |
| 7. The house | 1 | * <i>x̥yz̥/-a</i> , * <i>x̥ys̥/-a</i> , * <i>x̥yž̥/-a</i> , * <i>x̥yš̥/-a</i> ‘small house, cottage’ | |
| 21. Law | 1 | (* <i>duma</i> ‘advice, thought, opinion’) | |

¹³ The World Loanword Database assigns musical instruments to the semantic category “18. Speech and language” (Haspelmath/Tadmor 2009: 32).

It follows from the overview that the category 5. *Food and drink* contains the largest number of loanwords (15). This category includes the numerous borrowed designations for different types of containers, which were most likely borrowed because of extensive trade between Germanic and Slavic people. Many of these words are also dry or liquid measures in Slavic, such as PSI. **кѣбѣль* ‘tub; quantity of grain’. The words belonging to the semantic group 11. *Possession* include words for currencies, scales, toll and debt, which are also primarily borrowed as a result of trade and money dealing between Slavs and Germans. Another prominent semantic category is 22. *Religion and belief* (8 words), which contains the words that entered Proto-Slavic along with Christianity.

In the LTP-database of Lower Sorbian loanwords, the loanword-rate is highest in the following semantic fields: *Clothing and grooming*, *Religion and belief*, *The modern world*, *The house* and *Agriculture and vegetation*. Out of all words relating to *The modern world* almost 66% are loanwords, while the remaining ones are likely to be calques and neologisms. This semantic group is understandably absent from the corpus of loanwords in Proto-Slavic. Both Proto-Slavic and Lower Sorbian have a large number of loanwords in the semantic field *Religion and belief*. Whereas the groups *Clothing and grooming* and *The house* are prominent semantic groups for loanwords into Sorbian, there are almost no words in these categories borrowed into Proto-Slavic. Apparently, the Slavs did not borrow many designations for cloths and garments or furniture from Germanic in the early medieval period.

5. Conclusion

The comparative investigation of the loanwords in the individual Slavic languages is likely to be an interesting and fruitful enterprise. The meaning list from the project *Loanwords in the world's languages* and the project's methodology would serve this aim well. With the help of the meaning list, it will be possible to identify many loanwords in the core vocabulary, although the peripheral lexicon of the languages will of course contain many more loanwords. In my view, it would be especially interesting to compare different contact situations and their impact on the individual Slavic languages with each other, rather than just focus on the number of lexical borrowings. When investigating closely related languages, such as the Slavic languages, it is especially important that researchers approach and evaluate the material in the same way in order to avoid inconsistencies such as the ones discussed in this paper.

6. References

BARTELS, HAUKE (2009): Lower Sorbian vocabulary. In: Haspelmath, Martin; Tadmor, Uri (eds.) *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 1765 entries. <<http://wold.cild.org/vocabulary/10>> [accessed on 2015-03-17]

- BARTELS, HAUKE (2009): Loanwords in Lower Sorbian, a Slavic language in Germany. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 304–323.
- GREEN, DENNIS H. (1998): *Language and history in the early Germanic world*. Cambridge: Cambridge University Press.
- HASPELMATH, MARTIN; TADMOR, URI (eds.) (2009a): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter.
- HASPELMATH, MARTIN; TADMOR, URI (2009b): The Loanword Typology project and the World Loanword Database. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 1–34.
- HASPELMATH, MARTIN (2009): Lexical borrowing: Concepts and issues. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 35–54.
- HOOPS, JOHANNES; BECK, HEINRICH (eds.) (1989): *Reallexikon der germanischen Altertumskunde, Band 7*. Berlin: de Gruyter.
- PRONK-TIETHOFF, SASKIA (2013): *The Germanic loanwords in Proto-Slavic* (= Leiden Studies in Indo-European 20). Amsterdam – New York: Rodopi.
- SCHUHMANN, ROLAND (2009): Old High German vocabulary. In: Haspelmath, Martin; Tadmor, Uri (eds.): *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 1258 entries. <<http://wold.cld.org/vocabulary/11>> [accessed on 2015-03-17]
- TADMOR, URI (2009): Loanwords in the world's languages: Findings and results. In: Martin Haspelmath und Uri Tadmor (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 55–75.
- VAN DER SIJS, NICOLINE (2009): Dutch vocabulary. In: Haspelmath, Martin; Tadmor, Uri (eds.): *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 1588 entries. <<http://wold.cld.org/vocabulary/12>> [accessed on 2015-03-17]

Stefan Engelberg, Peter Meyer (Mannheim)

Das *Lehnwortportal Deutsch* als kontaktlinguistisches Forschungsinstrument

Abstract

The web portal *Lehnwortportal Deutsch* <lwp.ids-mannheim.de>, developed at the Institute for the German Language (IDS), aims to provide unified access to a growing number of lexicographical resources on German loanwords in other languages. This paper discusses different possibilities of creating an onomasiological access structure for portal users. We critically examine the meaning list of the “World Loanword Database” project (Haspelmath/Tadmor 2009a) as well as WordNet-based taxonomies and propose a new way of inductively creating a semantic classification scheme that takes both hyperonymic relations and semantic fields into account. We show how such a classification can be integrated into the underlying graph-based data representation of the *Lehnwortportal* and thus be exploited for advanced onomasiological search options.

1. Das *Lehnwortportal Deutsch*

Die am Institut für Deutsche Sprache entwickelte und gepflegte, frei zugängliche Online-Anwendung „*Lehnwortportal Deutsch*“¹ ermöglicht Nutzern einen einheitlichen Zugriff auf eine wachsende Anzahl von Wörterbüchern zu deutschen Lehnwörtern in anderen Sprachen. Damit fokussiert das Portal – anders als in der Sprachkontaktlexikographie üblich – nicht auf die Zielsprache der Entlehnungsprozesse, sondern auf die Quellsprache. Jedes in einem beliebigen Lehnwörterbuch des Portals gebuchte deutsche Etymon (das häufig einer älteren Sprachstufe oder einem Dialekt angehört) wird manuell einem deutschen Portal-Metalemma, das von diasystematischer Variation abstrahiert, zugeordnet. Abgesehen von den üblichen, insbesondere lemmabasierten, makrostrukturellen Zugriffsmöglichkeiten auf die Einzelwörterbücher werden durch diese portalweite lexikographische Vernetzung zum einen komplexe wörterbuchübergreifende Suchfunktionen möglich (vgl. Meyer 2013a). Zum anderen fungieren die – auch untereinander aufgrund von Derivations- und Kompositionsbeziehungen verknüpften – Metalemmata als Lemmata eines „umgekehrten Lehnwörterbuchs“ (vgl. Engelberg 2010, Meyer/Engelberg 2011), also eines automatisch erzeugten zusätzlichen Portalwörterbuchs der deutschen Herkunftswörter, dessen Einträge die Entlehnungswege deutscher Etyma in die verschiedenen Nehmersprachen der integrierten Lehnwörterbücher nachvollziehbar machen.

Eine der zentralen Aufgaben beim strukturellen Ausbau des Portals ist die Schaffung eines onomasiologischen Zugangs, der es erlauben soll, Abfragen über Sprachen, Zeiten und formale Eigenschaften mit solchen über semantisch-diskursive Eigenschaften von Herkunfts- und Lehnwörtern zu

¹ Vgl. <<http://lwp.ids-mannheim.de/>> [aufgerufen am 04.07.2015].

verknüpfen. In diesem Beitrag sollen die Vor- und Nachteile verschiedener onomasiologischer Konzeptionen und deren Umsetzung in eine onomasiologische Zugriffstruktur diskutiert werden.

2. Onomasiologische Strukturierungsoptionen

2.1. Anforderungen an ein onomasiologisches System

Unter einem onomasiologischen System soll hier ein Klassifikationssystem verstanden werden, das Wortbedeutungen² nach semantischen, pragmatischen und/oder diskursiven Eigenschaften für lexikographische Zwecke klassifiziert. Die Überlegungen dazu bewegen sich damit im Umfeld von Konzepten wie Sinnbezirken, Ontologien, Sachgruppen, Sachbereichen, lexikalischen Netzen und Thesauri. Ziel der hier formulierten Überlegungen ist es primär, verschiedene Optionen für solche Systeme hinsichtlich ihrer lexikographischen Praktikabilität zu sichten. Grundlegende semantische Überlegungen müssen weitgehend außen vor bleiben.

Über seine Nützlichkeit für die Bereitstellung entsprechender Suchfunktionen im Lehnwortportal hinaus sollte das verwendete onomasiologische System im Idealfall die folgenden Anforderungen erfüllen:

1. Es sollte in seinen Konzepten und Relationen semantisch gut begründet und definiert sein.
2. Es sollte in der Annotation möglichst einfach und eindeutig anwendbar sein.
3. Es sollte für Lehnwörterbücher verschiedener Typen und Zielsprachen verwendbar sein.

Bedingung 1 verlangt zum einen, dass die einzelnen Klassen hinreichend präzise definiert sind, zum anderen aber auch, dass die grundlegenden semantisch-diskursiven Konzepte – seien es nun hyperonymische Relationen, Sinnbezirke oder Diskursbereiche – hinreichend geklärt sind. Bedingung 2 ist dadurch operationalisierbar, dass eine manuelle semantische Annotation möglichst wenig Zeit erfordert und gleichzeitig mit einer hohen Annotatorenübereinstimmung einhergeht. Bedingung 3 schließlich hebt erstens darauf ab, dass Lehnwörterbücher bezüglich des lexikalischen Bereichs, den sie erfassen, oft bestimmte Fokussierungen oder Einschränkungen aufweisen, etwa hinsichtlich der Aufnahme von Internationalismen, fachsprachlich-wissenschaftlicher Lexik oder sogar Eigennamen. Zweitens unterscheiden sich die Kontaktsituationen, in denen einzelne Sprachen aus dem Deutschen entlehnt haben, erheblich, und damit auch der entlehnte Wortschatz. So lassen sich unterscheiden: (i) langandauernder Kontakt an Bevölkerungsgrenzen; Entlehnungen z. B. ins Slowenische, Polnische, Niederländische, Sorbische; (ii) Kontakt durch Emigration mit Sprachinselbildung; Entlehnungen z. B. ins Rumänische; (iii) Kontakt durch Elitenaustausch; Entlehnungen z. B. ins Japanische, Russische,

² Um die Redeweise von einem onomasiologischen System zu rechtfertigen, das ja einen Weg von lexikalischen Inhalten zu lexikalischen Formen nahelegt, müssen Wortbedeutungen hier genauer als disambiguierte Lexeme verstanden werden.

Englische; (iv) Kontakt in kolonialen oder missionarischen Kontexten; Entlehnungen z. B. ins Bàsàa (Kamerun), Samoanische, Labrador-Inuit, Yapesische (Mikronesien) oder Kuanua (Papua-Neuguinea). Dazu kommen indirekte Entlehnungen über Mittlersprachen, also Entlehnungen ohne direkten Kontakt mit dem Deutschen, so etwa über die Entlehnungskette Polnisch > Weißrussisch > Russisch ins Usbekische und andere Sprachen Zentralasiens oder über das Niederländische in die Sprachen Südasiens und in Niederländisch basierte Kreolsprachen.

Über die drei genannten Bedingungen hinaus sollte das onomasiologische System natürlich möglichst große Teile des verzeichneten Wortschatzes im Portal abdecken. Es ist allerdings nicht erforderlich, dass der Lemmabestand vollständig semantisch-thematisch erschlossen wird. Schließlich geht es nicht darum, die exhaustive Anwendbarkeit eines onomasiologischen Systems auf den Wortschatz einer Sprache nachzuweisen, sondern einen onomasiologischen Zugriff zu entwickeln, der lehnwortlexikalisch relevante Abfragen über semantisch-thematische Bereiche erlauben soll. Dabei ist auch zu berücksichtigen, dass mit Sachgruppenwörterbüchern wie etwa dem Dornseiff (2004), in denen onomasiologische Systeme und Zugriffe ihr lexikographisches Hauptanwendungsgebiet finden, ganz andere Nutzungsszenarien verbunden sind als mit onomasiologischen Zugriffen in Lehnwörterbüchern. Während jene vor allem die Ausdrucksfindung bei der Sprachproduktion befördern möchten, sollen letztere die Beantwortung von sprachhistorischen, sprachkontakttheoretischen oder kulturgeschichtlichen Fragen unterstützen. Eine Übernahme der entsprechenden Systeme aus der Sachgruppenlexikographie ist also nicht unbedingt naheliegend und ist im Rahmen des hier beschriebenen Projekts auch (noch) nicht verfolgt worden.

Im Folgenden sollen drei Optionen für die ontologische Strukturierung des Wortbestands des Portals diskutiert werden: (i) eine Klassifizierung auf der Basis der Bedeutungsklassen der Loanword Typology Meaning List des Leipziger Lehnwortprojekts, (ii) eine Anbindung an das Netz semantischer Relationen in GermaNet und (iii) der Versuch einer induktiven Entwicklung eines onomasiologischen Systems aus den Lehnwortdaten. Im Zuge dieser Diskussion werden wir Daten aus folgenden Wörterbüchern diskutieren:

- *W-Pol*: „Wörterbuch der deutschen Lehnwörter in der polnischen Schrift- und Standardsprache“ (de Vincenz/Hentschel/Brüggemann 2010; im Lehnwortportal enthalten); verzeichnet keine Lehnwörter, für die das Deutsche lediglich als Mittlersprache diente.
- *W-Frz*: „Das deutsche Lehngut im Französischen als Zeugnis für den Wissenstransfer im 20. Jahrhundert“ (Sarcher 2001); starker Anteil an wissenschaftlichem Wortschatz aus dem 20. Jh.

- *W-Slv*: „Deutsche Lehnwörter im Slovenischen“ (Striedter-Temps 1963; im Lehnwortportal enthalten); breite semantische Entlehnungsbereiche über einen langen Entlehnungszeitraum.
- *W-Pzf*: „Wörterbuch deutscher Lehnwörter in den Sprachen des Südpazifiks“ (in Vorbereitung am Institut für deutsche Sprache); Lehnwörter größtenteils aus der deutschen Kolonialzeit (v.a. in den Bereiche Schule, Religion, Verwaltung, Alltagsgegenstände).

Die folgenden Bemerkungen repräsentieren den gegenwärtigen Stand der noch nicht abgeschlossenen konzeptionellen Überlegungen zu einem onomasiologischen Zugriff für das Lehnwortportal.³

2.2. Bedeutungsklassen der „Loanword Typology Meaning List“

Die „World Loanword Database“ (Haspelmath/Tadmor 2009a, 2009b) basiert auf der sogenannten „Loanword Typology Meaning List“ (LTML), einer Liste von 1460 Konzepten, die 24 semantischen Klassen zugeordnet werden. Für diese Konzepte wurde zu 41 Zielsprachen ermittelt, ob und aus welcher Sprache die sie ausdrückenden Lexeme entlehnt wurden. Die 24 semantischen Klassen sind die folgenden:

| | | |
|--------------------------------|----------------------------------|------------------------------------|
| 01: THE PHYSICAL WORLD | 09: BASIC ACTIONS AND TECHNOLOGY | 17: COGNITION |
| 02: KINSHIP | 10: MOTION | 18: SPEECH AND LANGUAGE |
| 03: ANIMALS | 11: POSSESSION | 19: SOCIAL AND POLITICAL RELATIONS |
| 04: THE BODY | 12: SPATIAL RELATIONS | 20: WARFARE AND HUNTING |
| 05: FOOD AND DRINK | 13: QUANTITY | 21: LAW |
| 06: CLOTHING AND GROOMING | 14: TIME | 22: RELIGION AND BELIEF |
| 07: THE HOUSE | 15: SENSE PERCEPTION | 23: MODERN WORLD |
| 08: AGRICULTURE AND VEGETATION | 16: EMOTIONS AND VALUES | 24: MISCELLANEOUS FUNTION WORDS |

Wir haben diese Klassifikation auf 720 Lehnwörter aus 30 Sprachen aus dem Südpazifik-Wörterbuch *W-Pzf*, angewendet. Dabei tauchten zwei hauptsächliche Probleme auf. Das erste Problem betrifft die geringe Differenziertheit der Klassifikation. Die Lehnwörter verteilen sich erwartungsgemäß sehr

³ Aus expositorischen Gründen konzentrieren wir uns im Folgenden auf die Klassifizierung von Lehnwortbedeutungen; dies schließt selbstverständlich eine analoge onomasiologische Behandlung von Herkunftswörtern nicht aus.

ungleich auf die 24 Klassen, zwischen 0 (KINSHIP) und 105 (RELIGION AND BELIEF). Für die stark frequentierten Klassen läge hier eine größere Differenzierung nahe. Als problematisch erweist sich hier vor allem die Fokussierung der LTML auf prämoderne Konzepte, die zu einer geringen Differenzierung in Bereichen des modernen Lebens führen (Technik, Wissenschaft, Gesellschaft). Dies wirkt sich natürlich auch erheblich in Lehnwörterbüchern wie *W-Frz* aus, das nur Entlehnungen aus dem 20. Jh. dokumentiert.

Während dieses Problem gegebenenfalls durch eine eigene Subklassifikation behoben werden könnte, ist das zweite Problem substantiellerer Natur. Ein kohärenter Handlungs- und Diskursbereich, der Entlehnungen aus dem Deutschen in deutsch-kolonialen Zusammenhängen befördert hat, ist die Schule. Z. B. sind hier *Bleistift, malen, Griffel, Schule, Papier, auswendig, Tafel* und viele andere deutsche Etyma zur Quelle von Entlehnungen geworden. Ein Feld SCHULE kommt in der Leipziger Klassifikation aber nicht vor und ließe sich auch nicht als Subfeld eines schon bestehenden Feldes konstruieren. Vielmehr fallen Lehnwörter aus dem Schulkontext in verschiedene Felder der LTML: 17 COGNITION (hier wären *Schule* und *auswendig* einzuordnen) / 18 SPEECH AND LANGUAGE (*Papier, Bleistift, Buch*) / 23 MODERN WORLD (*Brief*).

Insgesamt ließen sich zwei Drittel der Lehnwörter in der LTML-Klassifikation gut, wenn eben auch nicht besonders differenziert, verorten. Für das andere Drittel mussten weitere Klassen wie etwa SCHOOL, TRANSPORT oder ADMINISTRATION hinzugefügt werden, um zu einer Klassifikation zu gelangen, die die für die spezifische Kontaktsituation relevanten Bereiche angemessen differenziert berücksichtigt und semantisch-diskursiv zusammengehörige Lehnwortgruppen nicht klassifikatorisch auseinanderreißt. Letztlich zeigt sich aber, dass aus der Geeignetheit der Feldgliederung der „Loanword Typology Meaning List“ für typologische Fragestellungen nicht die Geeignetheit für konkrete historisch und sprachlich spezifische Lehnwortschätze folgt.

2.3. Anbindung an GermaNet

Während die onomasiologische Ordnung der LTML auf der Zuordnung von Wortbedeutungen zu semantischen Feldern basiert, vernetzen lexikalisch-semantische Wortnetze wie das englischsprachige WordNet (Fellbaum 1998) und das deutschsprachige GermaNet (Hamp/Feldweg 1997, Henrich/Hinrichs 2010) den Wortschatz auf der Basis semantischer Relationen zwischen Wörtern, wobei in unserem Zusammenhang vor allem die das Wortnetz dominierenden hyperonymischen Relationen von Interesse sind.

Die Probleme bei den Versuchen, Lehnwortbedeutungen an die hyperonymische Struktur eines solchen Wortnetzes anzubinden, sollen im Folgenden am Beispiel des polnischen Lehnwörterbuchs *W-Pol*, erläutert werden. Hier stellt sich zunächst die Frage nach der Auswahl einer geeigneten WordNet-

Ressource. Vor dem Hintergrund, dass später Abfragen über unterschiedliche Wörterbücher und Nehmersprachen hinweg möglich sein sollen, ist die Verwendung je verschiedener einzelsprachlicher Wortnetze nicht sinnvoll. Der Verwendung einer Ressource wie EuroWordNet (Vossen et al. 1998), die über einen „interlingualen Index“ je eigenständige einzelsprachliche Wortnetze aufeinander abbildet, steht unter anderem die kleine Anzahl der in diesem mittlerweile abgeschlossenen Projekt berücksichtigten Sprachen entgegen, aber natürlich auch die grundsätzlichen Problematik der Verknüpfung von womöglich sehr unterschiedlich strukturierten Wortnetzen. Unter rein praktischen Gesichtspunkten ist für das Lehnwortportal Deutsch die sprachübergreifende Zuordnung von Wortbedeutungen zu „passenden“ GermaNet-Synsets vermutlich die einzige gangbare, wenngleich theoretisch nicht ganz befriedigende Vorgehensweise.

Es erweist sich dabei, dass aus verschiedenen Gründen Lehnwortbedeutungen oft mehrere Hyperonyme in GermaNet zugeordnet werden müssen. Das ist insofern ein Problem, als die Zuordnung eines Wortes zu zwei Hyperonymen außer in Fällen, in denen es in GermaNet Mehrfachvererbung gibt, auch zwei wohlunterschiedene Bedeutungspositionen des Wortes erwarten lässt. Diese Voraussetzung ist aber nicht immer gegeben. Zum einen verzichten Wörterbücher häufig auf Lesartendifferenzierungen, die für die Zuordnung zu GermaNet-Hyperonymen erforderlich sind. So wird *rajsplata* in *W-Pol* nur in einer Bedeutungsposition dargestellt, die mit der Beschreibung ‘Arbeitstisch, gußeiserne, glatte, waagerechte Tischplatte’ versehen ist. Das bedingt die doppelte Einbindung in GermaNet als |>Tisch2 >Möbel1 >Artefakt1| und als |>Tischplatte1 >Platte5 >Brett2 >Artefakt1>|. ⁴ Mehrfache Einordnungen in die GermaNet-Hierarchie treten des Weiteren auch bei systematischer Polysemie auf, soweit diese nicht im Wörterbuch abgebildet ist: *jarmuż* ‘Grünkohl’ ist sowohl als |>Grünkohl1 >... >Pflanze1| wie als |>Grünkohl2 >... >Gemüse1 >... >Nahrung1> ...| zu klassifizieren. Drittens treten Mehrfachzuordnungen bei unklaren semantischen Differenzierungen in GermaNet auf. So könnte *śrubsztak*, mit der Bedeutungsbeschreibung ‘Vorrichtung zum Festhalten von zu bearbeitenden Gegenständen, Schraubstock’ je nach Interpretation der GermaNet-Hyperonyme als |>Spannvorrichtung1 >Vorrichtung1> ...|, als |>Haltevorrichtung1 >Vorrichtung1> ...| oder als |>Spezialwerkzeug1 >Werkzeug1> ...| eingeordnet werden.

Ein zweiter Problemkomplex bei der Verwendung von GermaNet entsteht dadurch, dass die Hierarchisierungen in GermaNet zum Teil den semantischen Differenzierungen im Lehnwortbereich nicht angemessen sind. So ist bei Verben oft nur eine sehr grobe Einordnung in GermaNet möglich: *takielowác*, in der Bedeutung ‘die Takelage eines Schiffes vorbereiten’ ließe sich sowohl unter |>einrichten3|, |>vorbereiten1| oder |>fertig machen1| einordnen, während es sehr wohl ein substantivisches Synset

⁴ Die Zahlen indizieren Bedeutungspositionen der Hyperonyme in GermaNet. Die Beispiele beziehen sich auf Version 7.0 von GermaNet.

*Takelage*1 gibt. Dabei ist zudem nicht klar, welche Benutzerinteressen Abfragen befriedigen könnten, die auf solchen Zuordnungen basieren. Neben mangelnder Spezifik in einigen Bereichen ist das Wortnetz in anderen Teilen wiederum sehr filigran strukturiert. Das macht die Einordnung von Fachwortschatz, die ja lediglich aufgrund der in den Wörterbüchern gegebenen Paraphrasen erfolgen kann, oft schwierig. Ein weiteres Problem ist die Behandlung von prämodernem Wortschatz, der, ebenso wie fachsprachliche Bereiche, unter Umständen aufwändige Erweiterungen des Wortnetzes erforderlich machen kann.

Die vorgetragenen Bedenken legen nahe, Lehnwortbedeutungen nicht direkt semantisch passenden Synsets in GermaNet oder auch nur möglichst weit unten in der Hierarchie angeordneten Hyperonymen zuzuordnen. Auch für Benutzer des Lehnwortportals wäre es im Einzelfall eine anspruchsvolle Aufgabe, ein zur Problemstellung gut passendes Synset in den verzweigten Hierarchien eines Wortnetzes überhaupt ausfindig zu machen. Eine alternative Vorgehensweise besteht also darin, die Wortnetzhierarchie in einer bestimmten Tiefe abzuschneiden und Lehnwortbedeutungen auf eine begrenzte Anzahl von Synsets einer mittleren oder oberen Hierarchieebene abzubilden. Solche Synsets müssten auch in irgendeinem für die Sprachkontaktforschung relevanten Sinne besonders grundlegende semantische Konzepte repräsentieren. Die zunächst vielleicht naheliegende Verwendung nur der hierarchisch höchsten Konzepte (entsprechend etwa den *unique beginners* oder *supersenses* im Princeton-WordNet) scheitert jedoch letztlich an deren zu hohem Abstraktionsgrad, ähnlich wie Versuche, eine formale und daher für maschinelles logisches Schlussfolgern ausgelegte Top-Level-Ontologie wie SUMO (Niles/Pease 2001) auf eine WordNet-artige Ressource abzubilden, wie das für das Princeton-WordNet durchgeführt worden ist (Niles/Pease 2003).

Vielversprechender sind Versuche, nicht (nur) die am deutlichsten allgemeinen, sondern die in einem näher zu definierenden Sinne für onomasiologische Zwecke am ehesten relevanten Synsets auszuwählen. Zwei Ansätze, die wir jedoch bislang nicht systematisch auf ihre Tragfähigkeit im vorliegenden Kontext geprüft haben, sind die folgenden:

- Im EuroWordNet-Projekt wurden wortnetzübergreifend etwa 1000 sogenannte Base Concepts definiert, die den für die semantischen Lexikalisierungsmuster der einzelnen Wortnetzeentralen Synsets entsprechen und dementsprechend möglichst zugleich hoch in der Netzhierarchie und reich an Relationen zu anderen Synsets sein sollen (Vossen et al. 1998). Davon kommen ca. 160 sogenannte Core Base Concepts in mindestens drei von den sieben Wortnetzen des Projekts (Niederländisch, Italienisch, Spanisch, Deutsch, Französisch, Tschechisch, Estnisch) vor.
- Linguistisch womöglich noch interessanter sind Versuche, unter den Synsets solche auszuwählen, die *basic-level concepts* im Sinne von Rosch (1977) entsprechen, also unter Umständen eine basalere kognitive Relevanz haben, da sie semantisch

verhältnismäßig homogen und zugleich deutlich abgegrenzt von auf derselben Hierarchieebene benachbarten Konzepten sind. Die eben erwähnten Base Concepts sind in den meisten Fällen eine oder zwei Hierarchieebenen oberhalb von typischen basic-level concepts, liegen also eher auf der Ebene von *Möbel/Einrichtungsgegenstand* als auf der von *Tisch* oder *Spiegel*. Es gibt Versuche, mit automatisierten Verfahren eine eng begrenzbare Anzahl von Kandidaten für WordNet-Synsets auf dem basic level ausfindig zu machen (vgl. Izquierdo/Suárez/Rigau 2007, Lin 2010), wobei die Güte der Verfahren typischerweise mit ihrer Eignung für maschinelle Sprachverarbeitungsaufgaben wie etwa Lesartendisambiguierung beurteilt wird.

In jedem Fall wird für die Zwecke der Einrichtung einer onomasiologischen Suche ein im weitesten Sinne induktives, an den jeweils betrachteten Einzelfällen orientiertes Vorgehen erforderlich sein, da der Umfang der hyperonymischen Klassen handhabbar und die Taxonomie selber den Besonderheiten der jeweiligen Lehnwortschätze angemessen sein muss.

2.4. Eine induktive Ontologie

Die beiden vorgestellten Lösungen zur Unterstützung onomasiologischer Zugänge sind von zwei unterschiedlichen semantischen Konzepten ausgegangen, semantischen Feldern einerseits und hyperonymischen Relationen andererseits. Beide Arten inhaltlicher Zuordnung können natürlich für eine lexikalische Einheit vorgenommen werden und sind als Grundlagen onomasiologischer Zugriffsstrukturen konzipierbar: So lässt sich *Straßenbahn* in einer Taxonomie |<Schienenfahrzeug<Fahrzeug<Artefakt| hyperonymisch verankern ebenso wie es sich in einem semantisch-diskursiven Feld TRANSPORT & VERKEHR verorten lässt.

Angesichts der Probleme, die die Anwendung von LTMN und GermaNet auf unseren Lehnwortbestand begleiteten, soll im Folgenden versucht werden, eine Klassifikation induktiv aus den spezifischen lexikalischen Daten der Lehnwörterbücher zu gewinnen. Klassifiziert werden sollen dabei disambiguierte Lexeme, also Wortbedeutungen, insofern sie in den integrierten Wörterbüchern als Bedeutungspositionen ausgezeichnet sind. Die Wortbedeutungen sollen sowohl bezüglich hyperonymischer Relationen als auch in thematische Felder eingeordnet werden. Das Wesen der Hyperonymie als semantischer Relation lässt sich logisch-semantisch über Implikationen (und strukturelle Zusatzannahmen) bestimmen und soll hier nicht weiter diskutiert werden. Weitgehend unbeantwortet und in der Literatur kaum diskutiert ist die Frage, welche Relation eigentlich zwischen einem Wort bzw. einer Wortbedeutung und einem semantischen Feld, thematischen Feld, Sinnbezirk etc. besteht. Wir konzipieren diese Beziehung hier als eine Relation zwischen einer Wortbedeutung W^b als semantischer

Einheit und einem thematischen Feld T als einer textuell-pragmatischen Einheit. Ein thematisches Feld lässt sich dabei als ein thematisch gebundenes virtuelles Textkorpus verstehen und wäre damit ähnlich operationalisiert wie ein Diskurs in der Auffassung von Busse/Teubert (1994: 15). Die Relation zwischen W , B und T sei wie folgt definiert: Ein Wort W in der Bedeutung B gehört einem thematischen Feld T dann an, wenn W^B typischerweise in T vorkommt. An diese Zuordnung sind die folgenden Erwartungen geknüpft:⁵

- W^B ist mit T stärker kognitiv assoziiert als mit den weitaus meisten anderen thematischen Feldern.⁶
- W^B ist stark mit vielen anderen T zugeordneten lexikalischen Einheiten kognitiv assoziiert.
- W^B tritt in T relativ zur jeweiligen T -Korpusgröße häufiger auf als in den weitaus meisten anderen thematischen Feldern.
- W^B kookkuriert überzufällig oft mit vielen anderen T zugeordneten lexikalischen Einheiten.⁷

Die doppelte Kodierung von Wortbedeutungen in Bezug auf hyperonymische Relationen einerseits und thematische andererseits erlaubt es nun, wichtige Differenzierungen vorzunehmen. Dies soll an einem Wortschatzausschnitt aus *W-Pzf* illustriert werden. In den Sprachen des Südpazifiks findet sich eine größere Anzahl von Lehnwörtern aus dem Tierbereich, die unter anderem auf folgende deutsche Etyma zurückgehen: *Biene, Büffel, Esel, Frosch, Gans, Hund, Kakerlake, Kalb, Kamel, Katze, Kuh, Lamm, Löwe, Ochs, Rindvieh, Ross, Schaf, Schlange, Taube, Wachtel* und *Ziege*. Aus der Perspektive semantischer Relationen lassen sich alle diese Wörter als Hyponyme von *Tier* beschreiben. Thematisch gehören die entsprechenden Lehnwörter aber ganz unterschiedlichen Bereichen an. Während einige der Wörter im Kontext des Lebensumfelds (HAUSTIERE & VIEHZUCHT) Verwendung finden (z. B. *Gans, Katze*), treten andere Wörter als Konsequenz von Bibelübersetzungen in die Sprachen des Südpazifiks primär im Kontext RELIGION auf (z. B. *Esel, Schlange, Kamel, Löwe*). Die sich hyperonymisch einheitlich verhaltende Wortgruppe fällt also thematisch in zwei verschiedene Felder.

Ausgehend von diesen Überlegungen haben wir versucht, ein thematisches Klassifikationsschema induktiv aus Lehnwortdaten zu gewinnen. Induktiv soll hier nicht heißen, dass das Wissen über bestehende Systeme (wie etwa LTM, Dornseiff 1934, 2004, Hallig/Wartburg 1952, Archer et al. 2004) vollständig ausgeblendet wird, sondern dass thematische Felder auf der Basis systematischen

⁵ Vgl. zu ähnlichen Überlegungen zur Rolle von Kookkurrenzen bei der Zuordnung von lexikalischen Einheiten zu Feldern/Sachgruppen auch Wiegand (2004) und Quasthoff (2004).

⁶ Die Formulierung „mit den weitaus meisten“ beinhaltet, dass es möglich ist, eine Wortbedeutung mehr als nur einem thematischen Feld zuzuordnen.

⁷ Diese Konzeptionalisierung und Operationalisierung des Begriffs „thematisches Feld“ ist eher grundsätzlicher Natur. Die Operationalisierungen werden im Einzelfall im Rahmen der hier anvisierten lexikographischen Annotation nicht praktikierbar sein.

Analysierens von Lehnwortbedeutungen etabliert werden. Dazu wurde eine Liste von 450 Lehnwortbedeutungen zusammengestellt, die sich aus jeweils einer Zufallsauswahl von 150 Einheiten aus *W-Slv*, *W-Frz* und *W-Pzf* zusammensetzte. Um die thematische Annotation nicht zu erschweren und bei onomasiologischen Abfragen nicht zu kleine Treffermengen zu bieten, sollte die entstehende Klassifikation nicht zu feingliedrig sein. Angestrebt wurde eine maximal zweistufige Klassifikation, die auf der obersten Stufe nicht mehr als einige Dutzend Klassen enthält. Aus der sukzessiven Bearbeitung der Lehnwortliste entwickelten wir eine Klassifikation, die als alphabetisch geordnete Themenliste so beginnt:

| | |
|---------------------|-------------------------|
| 01: BAUWESEN | 09: HAUSHALT |
| 02: BERGBAU | 10: HIMMEL |
| 03: CHARAKTER | 11: JAGD & FISCHFANG |
| 04: ESSEN & TRINKEN | 12: KLEIDUNG & SCHMUCK |
| 05: FAMILIE | 13: KOMMUNIKATION |
| 06: GELD | 14: KRANKHEIT & HEILUNG |
| 07: GESELLSCHAFT | 15: KULTURRAUM |
| 08: HANDWERK | 16: ... |

Einige der Themen erlauben für einen größeren Teil der ihnen zugeordneten lexikalischen Einheiten eine Klassifikation auf zweiter Stufe, z. B.

LANDWIRTSCHAFT

LANDWIRTSCHAFT > ACKERBAU & OBST- UND GEMÜSEANBAU

LANDWIRTSCHAFT > HAUSTIERE & VIEHZUCHT

LANDWIRTSCHAFT > WEINBAU

WISSENSCHAFT

WISSENSCHAFT > BIOLOGIE

WISSENSCHAFT > CHEMIE

WISSENSCHAFT > PHYSIK

WISSENSCHAFT > PSYCHOLOGIE

WISSENSCHAFT > ...

Die Themenbezeichnungen wurden dabei jeweils mit Annotationshinweisen versehen, die den Umfang des Themas illustrieren, z. B.:

BAUWESEN Bezeichnungen für Gebäude, Gebäudeteile, Räumlichkeiten,
Baumaterialien, Handlungen beim Errichten von Gebäuden

| | |
|---------------------|--|
| HAUSHALT | Bezeichnungen für Gegenstände und Handlungen, die primär dem privaten Haushalt zuzurechnen sind (einschließlich Kochen und Küche, aber ohne KLEIDUNG & SCHMUCK und ohne ESSEN & TRINKEN) |
| KRANKHEIT & HEILUNG | Bezeichnungen für Krankheiten und Ausdrücke für körperliche Unwohlseinsempfindungen; Bezeichnungen für Heilmittel, Berufe im Gesundheitsbereich, Handlungen des Heilens |

Letztlich konnten über diese Vorgehensweise etwa 90% der Lehnwortbedeutungen einem oder mehreren Themen zugeordnet werden.⁸ Hier ein Auszug aus der Lehnwortliste mit den entsprechenden Zuordnungen:

| Wortbedeutung | Zielsprache | Lehnwort | Quelle | Thematisches Feld |
|-----------------------|-------------|-------------------|--------------|---|
| ... | | ... | ... | ... |
| Reis | Tok Pisin | <i>rais</i> | <i>W-Pzf</i> | (i) Landwirtschaft > Ackerbau, (ii) Essen & Trinken |
| Relativität(stheorie) | Französisch | <i>relativité</i> | <i>W-Frz</i> | Wissenschaft > Physik |
| Rollmops | Französisch | <i>rollmops</i> | <i>W-Frz</i> | Essen & Trinken |
| Rotpunkt | Französisch | <i>rot-punkt</i> | <i>W-Frz</i> | Spiel & Sport |
| Sägemehl | Tok Pisin | <i>sikmel</i> | <i>W-Pzf</i> | Handwerk |
| Satan | Nauruisch | <i>Satan</i> | <i>W-Pzf</i> | Religion |
| Schaf | Nauruisch | <i>Schaf</i> | <i>W-Pzf</i> | (i) Landwirtschaft > Haustiere-&-Viehzucht, (ii) Religion |
| Schaf, graue Geiß | Slovenisch | <i>birka</i> | <i>W-Slv</i> | Landwirtschaft > Haustiere-&-Viehzucht |
| scheiden | Französisch | <i>scheider</i> | <i>W-Frz</i> | Bergbau |
| ... | | ... | ... | ... |

Bei den verbleibenden ca. 45 Lemmata ergaben sich Zuordnungsprobleme verschiedener Art:

⁸ Eine Überprüfung der Annotatorenübereinstimmung bei der Kodierung steht noch aus.

- Die lexikographische Bedeutungsbeschreibung des Lehnworts war unklar, z. B. Slov. *ahker* ‘die Warte’ < Kärntnerisch *árker* (Niederhochdeutsch *Erker*).⁹
- Das Lehnwort war nicht hinreichend disambiguiert, z. B.: Yapesisch *silinder* < Dt. *Zylinder*. Es ist unklar, ob es sich um das Kleidungsstück oder das Maschinenteil handelt.
- Es ließ sich kein geeignetes Feld konstruieren, z. B. Slov. *ajntajlung* < Dt. *Einteilung*.

Auch wenn mit dem induktiven Klassifikationsversuch insgesamt ein befriedigendes Ergebnis erzielt wurde, das eine für onomasiologische Abfragen vermutlich interessante Klassenbildung ermöglicht, sind natürlich auch erhebliche Schwächen eines solchen Verfahrens zu konstatieren:

- i.) Die Setzung von Themen ist zwar durch die Daten geleitet, aber nur schwach kriteriengesteuert und nicht operationalisiert.
- ii.) Die Klassifikation ist spezifisch für unseren Lehnwortschatz und nicht ohne Weiteres auf andere Klassifikationen abbildbar, was wiederum vergleichende Studien (etwa mit der World Loanword Database) einschränkt.
- iii.) Das Inventar an Klassen erfordert vermutlich mit der Erweiterung des Lehnwortbestands auch Erweiterungen oder Änderungen des Themeninventars. Damit werden wiederum möglicherweise Reklassifizierungen des lexikalischen Bestands erforderlich.

Überprüft wurde die Lehnwortliste auch hinsichtlich hyperonymischer Relationen. Insgesamt war der Bestand an lexikalischen Einheiten, die aus Sicht einer onomasiologischen Suche sinnvollerweise einem Hyperonym zugeordnet werden konnten, geringer als bei der thematischen Zuordnung. Insbesondere die Hyperonyme *Fahrzeug*, *Gebäude*, *Zahlungsmittel*, *Kleidungsstück*, *Lebensmittel*, *Maßeinheit*, *Möbelstück*, *Person hinsichtlich ihres Berufs*, *Pflanze*, *Tier*, *Waffe* und *Werkzeug* subsumierten jeweils eine größere Anzahl an Lehnwortbedeutungen. Grundsätzlich könnten auch meronymische Relationen für eine onomasiologische Abfrage von Interesse sein. In dem untersuchten Lehnwortbestand fanden sich vereinzelt solche Relationen, etwa bei den Lehnwörtern zu den Etyma *Stachelbeere* als Meronym von *Stachelbeerstrauch* > *Pflanze* oder *Erker* als Meronym zu *Gebäude*. Wir haben diese Option hier nicht weiter verfolgt.

3. Onomasiologische Zugriffsstruktur im *Lehnwortportal Deutsch*

⁹ Die im vorliegenden Fall vermutlich einschlägige Bedeutung „Wachturm“ o.ä. ist als Wortbedeutung von *Warte* im heutigen Deutsch außerhalb von Toponymen kaum mehr gebräuchlich. Dass die Lesart „erhöhter Standpunkt“ nicht gemeint war, lässt sich dem Wörterbuchartikel, der die alleinige Grundlage für die ontologische Zuordnung darstellt, nicht explizit entnehmen, sondern erschließt sich dem philologisch geschulten Leser allenfalls aufgrund der weiteren etymologischen Erläuterungen im Artikel.

Im Folgenden soll geklärt werden, wie sich ein onomasiologisches System sinnvoll in die Datenstrukturen des Lehnwortportals integrieren und für onomasiologische Suchen nutzbar machen lässt. Der einfacheren Exposition halber wird im Folgenden lediglich das Problem thematischer Kodierungen von Wortbedeutungen berücksichtigt; hyperonymische Information lässt sich völlig analog modellieren.

3.1. Zur graphenbasierten Datenorganisation im *Lehnwortportal Deutsch*

Mit Ausnahme der Artikel des „umgekehrten Lehnwörterbuchs“ ist jeder Lehnwörterbucheintrag im Lehnwortportal Deutsch datenbankseitig zunächst als eigenständiges XML-Dokument repräsentiert, aus dem, außer bei imagedigitalisierten Werken, mit Standardtechnologien wie XSLT auch die gewöhnliche HTML-basierte Online-Ansicht des jeweiligen Artikels generiert wird. Nach einem für jedes Wörterbuch separat zu definierenden Verfahren (vgl. Meyer 2014b) wird zusätzlich aus jedem der XML-Dokumente automatisch die Menge der in ihm repräsentierten *Relationen* zwischen gebuchten Wortformen (Lehnwörtern und Etyma sowie davon abgeleitete Lexeme, insbesondere Derivate) extrahiert. Wichtige Typen von Relationen zwischen Lexemen sind etwa „y ist Lehnwort zum Etymon x“, „y ist Derivat zu x“, „y ist Kompositum mit Kompositionsglied x“, „y ist (phonologische, diasystematische, ...) Variante zu x“. Hinzu kommen Beziehungen des Typs „Etymon y ist dem Portal-Metalemma x zugeordnet“, die – mitsamt den Metalemmata selbst – beim Prozess der Integration der Lehnwörterbücher in das Portal mit eigens entwickelten Softwarewerkzeugen manuell definiert werden müssen und zugleich als Datengrundlage für das „umgekehrte Lehnwörterbuch“ dienen. So entsteht ein komplexes Netz von Beziehungen zwischen Wörtern, das formal als gerichteter Graph beschreibbar ist und im Portal im Rahmen einfacher Visualisierungen auch dem Benutzer anschaulich vermittelt wird. Nachstehende Abb. 1 zeigt den vom Metalemma *Waage* ausgehenden Teilgraphen, wie er unter <http://lwp.ids-mannheim.de/art/meta/lemma/Waage> angezeigt wird, und deutet zusätzlich die Quellen der visualisierten Informationen an.

Die Graphenrepräsentation wird datenbankseitig separat verwaltet, derzeit in einer relationalen Datenbank und im parallelen Testbetrieb auch in einer dedizierten Graphendatenbank. Diese Datenbanken werden bei der Erstellung der Artikel des „umgekehrten Lehnwörterbuchs“ verwendet, insbesondere aber auch für die sehr komplexen, z.T. wörterbuchübergreifenden Rechercheoptionen des Portals (vgl. Meyer 2013a, 2013b für weitere Details).

3.2. Integration eines onomasiologischen Systems in die Datenmodellierung

Jedem als Knoten im Graphen repräsentierten Wort/Lexem lassen sich Eigenschaften (Attribute) zuordnen, etwa zur orthografischen Form sowie zu grammatischen Eigenschaften und zur diastematischen Einordnung der betreffenden Form. Im Lehnwortportal Deutsch wird ein „Steckbrief“ der mit einem Knoten assoziierten Eigenschaften in der bereits erwähnten Graphenvisualisierung angezeigt, wenn man den Mauszeiger über dem betreffenden Knoten ruhen lässt. Insbesondere wird derzeit eine Liste aller zu einem Wort angegebenen lexikographischen Bedeutungsdefinitionen als Knotenattribut geführt und ermöglicht es Benutzern, in den fortgeschrittenen Suchfunktionen auch nach Zeichenketten in Bedeutungsangaben zu suchen. Da vor allem bei Komposita und produktiv gebildeten Derivaten in der Zielsprache in Wörterbüchern gerne auf eine explizite Bedeutungsangabe verzichtet wird, ist dieses Attribut in nicht wenigen Fällen jedoch leer bzw. nicht vorhanden.

Um die einzelnen Wortbedeutungen thematischen Feldern zuzuordnen, könnte man die Bedeutungsangaben selber als Listen komplex strukturierter Datenobjekte – statt einfach als Listen von Zeichenketten – modellieren. Ein Attribut eines solchen Datenobjektes wäre dann eine nicht leere Liste von Codes für die zugeordneten thematischen Felder. In der derzeit im Portal verwendeten relationalen Implementierung des Graphen ist ein solches Vorgehen technisch ohne weiteres möglich. Der bereits vorhandenen 1:n-Relation zwischen einem Knoten und seinen Bedeutungen wäre eine m:n-Relation zwischen diesen Bedeutungen und der Liste der thematischen Felder hinzuzufügen. In üblichen Graphendatenbanken stößt man mit diesem Vorgehen allerdings an eine Grenze, da beliebig komplex strukturierte Knotenattribute im Allgemeinen nicht möglich und auch technisch nicht sinnvoll sind. Der naheliegende Ausweg besteht darin, die Bedeutungen eines Wortes ebenfalls graphentheoretisch zu modellieren, also einen neuen Knotentyp einzuführen, der kein Wort bzw. Lexem, sondern eine einzelne Wortbedeutung repräsentiert und als Attribute zum einen die im jeweiligen Wörterbuch angegebene Bedeutungsdefinition, sofern vorhanden, und zum anderen eine Liste der zugeordneten thematischen Felder aufweist. Von einem Wortknoten können dann eine oder mehrere gerichtete Kanten zu den zugehörigen Wortbedeutungsknoten führen. Ein technisch-konzeptueller Vorteil dieser Modellierung ist die einfachere und besser indizierbare Abfragbarkeit; es gibt aber auch lexikographische Vorteile einer ‚Reifizierung‘ von Wortbedeutungen als Knoten, die es insbesondere ermöglicht, beliebige Relationen

zwischen Wortbedeutungen zu modellieren. Wenn beispielsweise bei einem Derivat, das die Nehmersprache nach einem produktiven (und dem intendierten menschlichen Adressaten vertrauten) Wortbildungsmuster von einem Lehnwort gebildet hat, die Bedeutungsangabe fehlt, kann in der graphenbasierten Modellierung ein „Platzhalter“-Bedeutungsknoten angesetzt werden, der durch Kanten eines speziellen, verweisenden Typs mit den im Wörterbuch definierten Bedeutungen des Simplex verbunden ist. Interessanter sind Beziehungen zwischen den Bedeutungen von Lehnwörtern in Entlehnungsketten. So werden in einem laufenden Projekt zu polnisch vermittelten Germanismen in den ostslavischen Sprachen, deren lexikographische Erfassung im Lehnwortportal publiziert werden soll (Meyer 2014c), systematisch die Beziehungen von Wortbedeutungen der ostslavischen Lehnwörter zu den Bedeutungen korrespondierender Lehnwörter laut dem bereits im Portal republizierten *W-Pol* erfasst; es ist daher denkbar, dass solche Beziehungen zumindest für diesen Teilbereich des Portals in die Graphendatenbank aufgenommen werden und für Benutzer abfragbar sind.

3.3. Onomasiologische Suchfunktionen im Lehnwortportal

Schon jetzt sind im *Lehnwortportal Deutsch* wörterbuchübergreifende Suchen nach Zeichenketten in Bedeutungsangaben möglich, die zudem mit Booleschen Operatoren miteinander verknüpfbar sind. In der formularbasierten erweiterten Suche¹⁰ kann so beispielsweise nach Wörtern gesucht werden, bei denen in der Definition einer und derselben Wortbedeutung sowohl die Zeichenkette „Gewicht“ als auch die Zeichenkette „Last“ vorkommt; alternativ ist aber auch eine Suche nach Wörtern möglich, bei denen sowohl das Wort „Gewicht“ als auch das Wort „Wichtigkeit“ in je mindestens einer Lesart, aber nicht notwendigerweise in derselben, vorkommen. In einem um thematische Felder erweiterten Suchformular könnte man dann zusätzlich eine Auswahlliste der in der thematischen Klassifikation verwendeten Felder anbieten, so dass der Nutzer der Suchfunktion Wortbedeutungen sowohl unter Angabe einer in der Definition enthaltenen Zeichenkette als auch anhand zugeordneter thematischer Felder spezifizieren kann.

Für wissenschaftliche Zwecke besonders interessant sind die bereits jetzt möglichen Abfragen mit einer eigens für das Lehnwortportal konzipierten, deklarativen Abfragesprache¹¹ die es ermöglicht, explizit nach beliebig komplexen Konstellationen im gerichteten Graphen zu suchen.¹² Werden die Wortbedeutungen wie beschrieben als Knoten kodiert, lässt sich die Abfragesprache auf naheliegende

¹⁰ Vgl. <<http://lwp.ids-mannheim.de/search/meta>> [aufgerufen am 04.07.2015].

¹¹ Vgl. <<http://lwp.ids-mannheim.de/search/prof>> [aufgerufen am 04.07.2015].

¹² Vgl. hierzu die allgemeinen Überlegungen in Meyer (2013a) sowie die Dokumentation des Abfragesystems unter <<http://lwp.ids-mannheim.de/doc/meta/graphsearch>> [aufgerufen am 04.07.2015].

Weise so erweitern, dass auch die thematische Klassifikation von Bedeutungen in Suchen einbezogen werden kann. Ein hypothetisches¹³ Beispiel:

```
/* deklariere und benenne Knoten (Wörter) */
suche etymon ety.
suche lehnwort lw.
suche bedeutung lw_bed.
suche lehnwort deriv.
suche bedeutung deriv_bed.

/* lege Eigenschaften der Knoten (Wörter) fest */
ety ist substantiv.
lw_bed gehoert zum feld 'Bewertung & Klassifikation'.
deriv ist adjektiv.
deriv_bed enthaelt 'wichtig'.

/* lege Eigenschaften der Kanten (Relationen) fest */
lw ist lehnwort zu ety.
lw_bed ist bedeutung von lw.
deriv ist derivat zu lw.
deriv_bed ist bedeutung von deriv.
```

Mit der gezeigten Abfrage werden Teilgraphen mit folgender Konstellation gesucht: Ein Lehnwort *lw* zu einem deutschen substantivischen Etymon *ety* bildet nehmersprachlich ein adjektivisches Derivat *deriv*. Mindestens eine Wortbedeutung *lw_bed* von *lw* ist dem thematischen Feld ‚Bewertung & Klassifikation‘ zugeordnet; die Definition mindestens einer Wortbedeutung *deriv_bed* von *deriv* enthält die Zeichenkette *wichtig*. Einer solchen Suchanfrage entspricht als Suchergebnis allgemein eine Liste von passenden Knoten-n-Tupeln. Im vorliegenden Beispiel könnte ein einzelnes konkretes Suchergebnis etwa so dargestellt werden:

- *ety*: Etymon *Waage* (Subst., f.) im Artikel → *waga* (W-Pol)
- *lw*: Lehnwort *waga* (Subst., f.) im Artikel → *waga* (W-Pol)
- *lw_bed*: Bedeutung ‘Wichtigkeit, Bedeutung von etw.’; thematische Felder: BEWERTUNG & KLASSIFIKATION, ...; hyperonymische Relationen: ... im Artikel → *waga* (W-Pol)
- *deriv*: Derivat *ważny* (Adj.) im Artikel → *waga* (W-Pol)

¹³ Lässt man im Beispiel die drei Bedingungen zu Bedeutungsknoten fort, entsteht eine bereits jetzt im Portal zulässige Suchanfrage.

- *deriv_bed*: Bedeutung ‘wichtig; gewichtig’; thematische Felder: BEWERTUNG & KLASSIFIKATION, ...; hyperonymische Relationen: ... im Artikel → *waga (W-Pol)*

4. Literatur¹⁴

- ARCHER, DAWN; RAYSON, PAUL; PIAO, SCOTT; MCENERY, TONY (2004): Comparing the UCREL Semantic Annotation Scheme with Lexicographical Taxonomies. In: Williams, Geoffrey; Vessier, Sandra (eds.): *Proceedings of the Eleventh EURALEX International Congress. EURALEX 2004, Lorient, France. July 6–10, 2004. Volume III*. Lorient Cedex: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud, 817–827.
- BUSSE, DIETRICH; TEUBERT, WOLFGANG (1994): Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In: Busse, Dietrich; Herrmanns, Fritz; Teubert, Wolfgang (eds.): *Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik*. Oplanden: Westdeutscher Verlag, 10–28.
- DORNSEIFF, FRANZ (1934): *Der deutsche Wortschatz nach Sachgruppen*. Berlin: de Gruyter.
- DORNSEIFF, FRANZ (2004): *Der deutsche Wortschatz nach Sachgruppen*. 8., völlig neubearbeitete und mit einem vollständigen alphabetischen Zugriffsregister versehene Auflage von Uwe Quasthoff. Berlin: de Gruyter.
- ENGELBERG, STEFAN (2010): An inverted loanword dictionary of German loanwords in the languages of the South Pacific. In: Dykstra, Anne; Schoonheim, Tanneke (eds.): *Proceedings of the XIV EURALEX International Congress (Leeuwarden, 6–10 July 2010)*. Ljouwert (Leeuwarden): Fryske Akademy, 639–647.
- FELLBAUM, CHRISTIANE (ed.) (1998): *WordNet: An Electronic Lexical Database*. Cambridge: MA: MIT Press.
- HALLIG, RUDOLF; VON WARTBURG, WALTHER (1952): *Begriffssystem als Grundlage für die Lexikographie. Versuch eines Ordnungsschemas*. Berlin: Akademie-Verlag.
- HAMP, BIRGIT; FELDWEG, HELMUT (1997): GermaNet – a Lexical-Semantic Net for German. In: *Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications. ACL/EACL-97 Workshop Proceedings. July 12th 1997. Madrid*. Somerset: Association for Computational Linguistics, 9–15.
- HASPELMATH, MARTIN; URI, TADMOR (eds.) (2009a): *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <<http://wold.clld.org>>
- HASPELMATH, MARTIN; TADMOR, URI (2009b): The Loanword Typology project and the World Loanword Database. In: Martin Haspelmath; Tadmor, Uri (eds.) *Loanwords in the world's Languages. A Comparative Handbook*. Berlin: de Gruyter: 1–34.
- HENRICH, VERENA; HINRICHS, ERHARD (2010): GernEdiT – The GermaNet Editing Tool. In: *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC10). Valletta, Malta. May 17–23, 2010*. Valletta, Malta: European Language Resources distribution agency, 2228–2235.
- IZQUIERDO BEVIÁ, RUBÉN; SUÁREZ CUETO, ARMANDO; RIGAU CLARAMUNT, GERMAN (2007): Exploring the Automatic Selection of Basic Level Concepts. In: *International Conference. Recent*

¹⁴ Alle im Folgenden angegebenen URL's wurden am 29.5.2015 aufgerufen.

- Advances in Natural Language Processing. Proceedings. Borovets, Bulgaria.* Shoumen: INCOMA Ltd., 298–302.
- LIN, SHU-YEN [林淑晏] (2010): *A Computational Study of the Basic Level Nouns in English.* Unpublished doctoral dissertation, National Taiwan Normal University. <http://ir.lib.ntnu.edu.tw/retrieve/51190/metadata_02_02_s_05_0155.pdf>
- MEYER, PETER; ENGELBERG, STEFAN (2011): Ein umgekehrtes Lehnwörterbuch als Internetportal und elektronische Ressource: Lexikographische und technische Grundlagen. In: Hedeland, Hanna; Schmidt, Thomas; Wörner, Kai (eds.): *Multilingual Resources and Multilingual Applications.* Hamburg: Universität Hamburg, 169–174.
- MEYER, PETER (2013a): Advanced graph-based searches in an Internet dictionary portal. In: Kosem, Iztok et al. (eds.): *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17–19 October 2013, Tallinn, Estonia.* Ljubljana/Tallinn, 488–502.
- MEYER, PETER (2013b): Ein Internetportal für deutsche Lehnwörter in slavischen Sprachen. Zugriffsstrukturen und Datenrepräsentation. In: Kempgen, Sebastian et al. (eds.): *Deutsche Beiträge zum 15. Internationalen Slavistenkongress, Minsk 2013* (= Die Welt der Slaven. Sammelbände 50). München: Sagner, 233–242.
- MEYER, PETER (2014a): Graph-Based Representation of Borrowing Chains in a Web Portal for Loanword Dictionaries. In: Abel, Andrea; Vettori, Chiara; Ralli, Natascia (eds.): *Proceedings of the XVI EURALEX International Congress: The User in Focus. 15–19 July 2014, Bolzano/Bozen.* Bolzano/Bozen: EURAC research, 1135–1144.
- MEYER, PETER (2014b): Von XML zum DAG: Der lexikographische Prozess bei der Erstellung eines graphenbasierten Wörterbuchportals. In: Domínguez Vázquez, María José; Mollica, Fabio; Nied Curcio, Martina (eds.): *Zweisprachige Lexikographie zwischen Translation und Didaktik* (= Lexicographica, Series Maior 147). Berlin: de Gruyter, 303–321.
- MEYER, PETER (2014c): Entlehnungsketten in einem Internetportal für Lehnwörterbücher. IT-Infrastruktur und computerlexikographischer Prozess in einem Projekt zu polnisch vermittelten Germanismen im Ostslavischen. In: Mann, M. (ed.): *Digitale Lexikographie. Ein- und mehrsprachige elektronische Wörterbücher mit Deutsch: aktuelle Entwicklungen und Analysen* (= Germanistische Linguistik 223/224). Hildesheim u. a.: Olms, 97–132.
- NILES, IAN; PEASE, ADAM (2001): Towards a standard upper ontology. In: Welty, Chris; Smith, Barry (eds.): *2nd International Conference on Formal Ontology in Information Systems. FOIS 2001, Ogunquit, Maine, USA. October 17–19, 2001. Proceedings.* New York, NY: ACM Press, 17–19.
- NILES, IAN; PEASE, ADAM (2003): Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In: Arabnia, Hamid R. (ed.): *Proceedings of the 2003 International Conference on Information and Knowledge Engineering IKE'03. June 23–26, 2003. Las Vegas, Nevada, USA. Volume II.* Athens, GA: CSREA Press, 412–416.
- QUASTHOFF, UWE (2004): Methodologische Einführung. In: Dornseiff, Franz: *Der deutsche Wortschatz nach Sachgruppen.* 8., völlig neubearbeitete und mit einem vollständigen alphabetischen Zugriffsregister versehene Auflage von Uwe Quasthoff. Berlin u. a.: de Gruyter.
- ROSCH, ELEANOR (1977): Human Categorisation. In: Warren, Neil (ed.): *Studies in Cross-Cultural Psychology*, Vol. I. London: Academic Press, 1–49.
- SARCHER, WALBURGA (2001): *Das deutsche Lehngut im Französischen als Zeugnis für den Wissenstransfer im 20. Jahrhundert.* Hamburg: Kovač.

- STRIEDTER-TEMPS, HILDEGARD (1963): *Deutsche Lehnwörter im Slovenischen*. Wiesbaden: Harrassowitz.
- VINCENZ, ANDRÉ DE; GERD HENTSCHEL; MARK BRÜGGEMANN (2010): *Wörterbuch der deutsche Lehnwörter in der polnischen Schrift- und Standardsprache. Von den Anfängen des polnischen Schrifttums bis in die Mitte des 20. Jahrhunderts* (= *Studia Slavica Oldenburgensia* 20). Oldenburg: BIS-Verlag. <<http://diglib.bis.uni-oldenburg.de/bis-verlag/wdlp>>
- VOSSEN, PIEK ET AL. (1998): *The EuroWordNet Base Concepts and Top Ontology. Deliverable D017, D034, D036, WP5, EuroWordNet (LE2-4003)*. University of Amsterdam.
- WIEGAND, HERBERT ERNST (2004): Lexikographisch-historische Einführung. In: Dornseiff, Franz (2004): *Der deutsche Wortschatz nach Sachgruppen*. 8., völlig neubearbeitete und mit einem vollständigen alphabetischen Zugriffsregister versehene Auflage von Uwe Quasthoff. Berlin: de Gruyter: 9–91.

Jürgen Fuchsbauer (Regensburg, Wien)

Gräzisierung versus Sprachpurismus

Zum Fremdgut im Wortschatz des bulgarischen Kirchenslawisch des 14. Jahrhunderts

Abstract

The present paper focuses on a certain aspect of the treatment of borrowings from Greek in Middle Bulgarian Church Slavonic, namely on the replacement of established loanwords with calques. This puristic tendency, which appears to be characteristic of a number of translators, redactors, and scribes working presumably on Mount Athos and in Tárnovo in the first half of the 14th century, constitutes a deliberate approach to the treatment of language. It can, therefore, be seen in the context of the reformation of Church Slavonic during the times of the Second Bulgarian Empire, thus giving insight into the linguistic reasoning of some of the protagonists of this crucial stage in the history of Church Slavonic.

Bekanntermaßen erfolgte in mittelbulgarischer Zeit eine Erneuerung des kirchenslawischen Schrifttums, die weit über den bulgarischen Sprachraum hinaus wirkte und dabei einen prägenden Einfluss auf die Schriftkultur Serbiens, der Moldau und der Rus' ausübte.¹ Ihren Ursprung hatte sie wohl einerseits darin, dass Texte neu aus dem Griechischen übertragen wurden, und andererseits darin, dass bereits vorhandene Übersetzungen als unzulänglich oder durch Textverderbnis entstellt erkannt und mithilfe von griechischen Vergleichstexten überarbeitet wurden. Ein derartiges Redigieren ist für eine Literatur, die aus kulturellen und – mehr noch – religiösen Gründen sprachlich und inhaltlich so stark von einer fremden abhängt wie die kirchenslawische von der griechischen, keinesfalls ungewöhnlich. Wir erfahren bereits aus dem Traktat des Mönchs Chrabrъ, wenn dies denn der Name des Verfassers war, dass man die Übersetzungen der kyrillomethodianischen Ära als mangelhaft ansah und deshalb einer Revision unterzog. Der Text lautet in der Rekonstruktion William Veders (1999: 144ff.):

Ašte li kъto rečetъ. jako něstъ oustroilъ dobrě. da po nemъ sę postrajajotъ i ješte. otvěť rečetъ simъ. i grъčъsky [sc. kъnigy]² takožde mъnogašъdy sotъ postrajali. akъla i sъmъmachъ. i po tomъ i ini mъnodzi. oudoběje bo jestъ poslěžde potvoriti. neže pъvoje sъtvoriti [...] // („Wenn jemand sagt, er (Konstantin-Kyryll) habe sie nicht gut geschaffen, sodass sie nach ihm immer noch verändert werden, so sagen wir ihnen zur Antwort: Auch die griechischen hat man oftmals verändert, (und zwar) Aquila und Symmachos und danach viele andere; es ist ja leichter, etwas im Nachhinein zu ändern als es von Anfang an zu erschaffen.“)

¹ Ein auch nur annähernd vollständiger Überblick zur Literatur zu diesem Thema kann in diesem Rahmen nicht gegeben werden. Eine konzise Zusammenfassung bietet A. Totomanovas Beitrag *Ezikovo-pravopisna reforma na Patriarch Evtimij* in Petkanova (2003: 157f.). Genaueres etwa bei Miltenova (2009: 517f. und 569ff.).

² Veder (1999: 143f.) weist zu Recht darauf hin, dass der nom. pl. fem. *kъnigy* voraussetzt, und nicht etwa das Neutrum *pismena*; beides kann wie griechisch γράμματα und lateinisch *litterae* gleichermaßen 'Buchstaben' wie 'Schriften' bedeuten.

Aus dieser Stelle geht nicht unmittelbar hervor, ob die Überarbeitungen den Schriftzeichen oder den Schriften, d. h. den Texten, galten. Der Verfasser widerlegt aber den implizierten Vorwurf der Minderwertigkeit der slawischen Schriftlichkeit mit dem Hinweis, entsprechende Verbesserungen seien auch im Griechischen vorgenommen worden, und zwar unter anderem durch Aquila und Symmachos. Diese sind aber eben nicht als Schriftreformer, sondern als Übersetzer des Alten Testaments bekannt. Somit ist es das Wahrscheinlichste, dass hier auf eine Revision altkirchenslawischer Übersetzungen angespielt wird. Es war wohl die erste überhaupt. Chrabъs Traktat verteidigt das slawische Schrifttum gegenüber dem Griechischen und kann so nur zu einer Zeit geschrieben worden sein, in der die slawische Schriftsprache im bulgarischen Reich noch nicht vollständig etabliert war, d. h. unter Boris, Vladimir-Rasate oder allenfalls noch in den ersten Regierungsjahren Symeons.

Eine Reform des Schrifttums wie die mittelbulgarische stellte also keineswegs ein Novum dar. Warum diese aus der Sicht ihrer Protagonisten notwendig war, erklärt Grigorij Camblak in seiner Lobrede auf ihren mutmaßlichen Vollender, den Tärnover Patriarchen Evtimij (Rusev et al., 1971: 166ff.):

Нѡ прѣвѣи прѣвѣдите, или за еѣе еллинскаго ѣзыка ѣе и оуѣенїа не вѣ конецѣ вѣд(ѣ)ти, или и за еѣе своего ѣзыка дебелѡсти слѡужити, еѣе издаше кнїгы, неслѡзни вѣ рѣчехъ явише сѣ и разѡумѣнїи грѣѣьскыихъ писанїи несѣгласны, дебелствѡмъ ѣе свѣзаны и не гладкы кѣ теѣенїюу глаголюмѡу. И тѣкмо ѡт еѣе именовати сѣ бл(аго)ѣ(ѣ)стивыихъ кнїгы вѣрноє имѣчѡ, мноѡ ѣе вѣдѣ вѣ нихъ крѣае сѣ и истиннымъ догматѡмъ сѣпротивленїе. // („Die Bücher, die die ersten Übersetzer herausgaben, erwiesen sich, sei es, weil diese die griechische Sprache und Gelehrsamkeit nicht vollständig beherrschten, oder auch weil sie der Grobheit ihrer eigenen Sprache gehorchten, als einfach im Ausdruck und als dem Gehalt der griechischen Schriften unangemessen, als grob zusammengefügt und nicht glatt in Bezug auf den Redefluss. Und allein weil man sie (Werke) der frommen (Männer) nannte, wurden die Bücher für etwas Vertrauenswürdiges gehalten; doch es verbarg sich viel Schädliches in ihnen, auch solches, das im Widerspruch zu den wahren Dogmen stand.“)

Somit war der wesentliche Grund für die Reform – und das ist beileibe nicht überraschend – die Furcht vor Irrlehren, die sich aus der mangelnden Übereinstimmung mit den griechischen Vorlagen ergeben mussten. Hinzu kommt aber ein weiteres Kriterium, und zwar ein ästhetisches, nämlich die sprachliche, wie Camblak es ausdrückt, *debelostь*³ früherer Übersetzungen. Das Ziel der mittelbulgarischen Reformer um Evtimij war es entsprechend der Darstellung Camblaks also offensichtlich nicht bloß, die vom Standpunkt der Rechtgläubigkeit geforderte Übereinstimmung mit den griechischen Originaltexten

³ Goldblatt (1977: 45) meint, hier sei von zwei „causes of textual corruption“ die Rede, nämlich von „insufficient knowledge of the Greek language and doctrine“ und „improper choice of linguistic norm.“ Diese „adoption of a ‘coarse’ linguistic model (debelostь svoego ѣзыка), which did not correspond to that found in the ‘very old Bulgarian books,’ inevitably led to textual and spiritual error.“ Diese Interpretation würde ich nicht teilen. Camblak spricht nicht von Textverderbnis in der slawischen Überlieferung, sondern von Mängeln in eben den älteren, oder vielmehr sogar ältesten Übersetzungen; diese kommen nicht nur aufgrund der Unkenntnis des Griechischen sowie der christlichen Doktrin, sondern auch wegen der *debelostь* des Slawischen den Vorlagen in Hinblick auf Inhalt und Sprache nicht gleich.

herzustellen; vielmehr sollte das Sprachniveau insgesamt gehoben werden, und dazu bedurfte es nicht allein einer auf den Inhalt bezogenen, sondern eben auch einer sprachlichen Reform. Als deren tragende Elemente würde ich *Gräzisierung*, *Archaisierung* und *Standardisierung* nennen.

Bekanntlich versuchte man in mittelbulgarischer Zeit, die Orthographie des Kirchenslawischen zu normalisieren; dieser ist auch ein wesentlicher Teil des einzigen theoretischen Traktats gewidmet, den uns die literarische Schule von Tárnovo, oder besser gesagt, einer ihrer Epigonen, hinterlassen hat, nämlich Konstantin von Kosteneč' *skazánie izbjavljénno ó písmene/ch/*. Wir können jedoch davon ausgehen, dass die *hommes des lettres* der mittelbulgarischen Zeit auch über andere Bereiche der Sprache reflektierten, um sich so ihrem Ziel, einem dem Griechischen in Form, Inhalt und Würde gleichkommenden Slawischen, anzunähern. Beispielsweise ist der Kasusgebrauch bei Evtimij, wie sich anhand der weitgehenden Vermeidung adnominaler Dative zeigen lässt, noch deutlich gräzisierender als jener unwesentlich älterer Übersetzungen.⁴ Dies zeugt von einem bewussten Umgang mit Sprache, wobei wir die dahinterstehenden theoretischen Überlegungen, das, wenn man so will, zugrundeliegende Sprachdenken jedoch nirgendwo explizit festgehalten finden. Wir müssen dementsprechend versuchen, es aus den erhaltenen Texten zu erschließen.

Dabei wäre es nicht allein in Hinblick auf den theoretischen Unterbau der Reformbewegung, sondern auch auf die zeitliche und räumliche Einordnung von Texten, Redaktionen und Handschriften wünschenswert, verschiedene Strömungen innerhalb dieser zu definieren. Das ist allerdings alles andere als einfach, zumal wir ja mangels aussagekräftiger historischer Quellen noch nicht einmal den Beginn der mittelbulgarischen Schriftreform mit ausreichender Genauigkeit festlegen können; es bleibt bislang unklar, ob er im 14. oder bereits im 13. Jh. anzusetzen ist.

Aufgrund des Kolophons einer im sinaitischen Katharinenkloster aufbewahrten Oktōēchos-Handschrift, des codex Sinaiticus slavicus 19,⁵ ging man gemeinhin davon aus, dass zumindest die Hauptphase der Reform in der Regierungszeit des als bibliophil bekannten Zaren Ioann Alexander (reg. 1331–1371) oder wenig davor einsetzte, und dass sie ihren Ausgang von der Athanasios-Laura am Athos nahm.⁶ Der besagte Kolophon wurde von einem Mönch namens Methodie geschrieben. Dieser nennt den Verfasser seiner mittelbaren Vorlage, einen Starec Ioann; Ioann habe eine Reihe von Büchern „übertragen“ – gemeint ist offensichtlich teilweise auch redigiert –, und zwar das Tetraevangelium, den Praxapostolos, Werke des Johannes Klimakos, des Abbas Dorotheos, des Antiochos von Mar Saba usw.; diese habe er den Kirchen des bulgarischen Landes übergeben. Wo er das getan hat, erfahren wir, streng

⁴ Vgl. hierzu meinen Beitrag zum von Jasmina Grković-Major, Björn Hansen und Barbara Sonnenhauser herausgegebenen Band „Diachronic Slavonic Syntax: The interplay between internal development, language contact and metalinguistic factors“ (in Vorbereitung).

⁵ Faksimile in Jovčeva (2004: 232f.).

⁶ Vgl. zum Folgenden auch Fuchsbauer (2010: 273ff.).

genommen, nicht; der Kolophon besagt lediglich, dass Ioann seine *Vorlagen* aus der Athanasios-Laura hatte.⁷

Man war der Meinung, dass die Tätigkeit Ioanns in die Zeit unmittelbar vor der Abfassungszeit des Codex fällt, also in die erste Hälfte des 14. Jh.. Nun wies aber Spasova (2007: 222) wohl zu Recht darauf hin, dass es sich beim Kolophon im Sin. sl. 19 nicht um einen, sondern eigentlich um zwei handelt, von denen nur der zweite von Methodie stammt; den ersten habe er bereits seiner Vorlage entnommen. Deshalb müsse auch kein enger zeitlicher Zusammenhang zwischen dem Alter des Codex Sin. sl. 19 und dem Wirken Ioanns gegeben sein. Letzteres meint Spasova bereits mit der Regierungszeit des Zaren Ioann Asen II. (reg. 1218–1241), also in der ersten Hälfte des 13. Jh. ansetzen zu können. Genauer gesagt, kann die erste Subskription erst in einer Abschrift, die ein mit Ioann bekannter Kopist anfertigte, erfolgt sein – für Ioann wird immer die dritte Person verwendet. Wir haben es also mit mindestens drei Handschriften zu tun, und zwar erstens mit der Ioanns, zweitens mit der des mit ihm bekannten Schreibers, und drittens mit der des Methodie; von diesen ist lediglich die letzte auf uns gekommen. Zwischen ihr und der zweiten können eine unbestimmte Anzahl weiterer Kopien gewesen sein. Es lässt sich aber keinesfalls erschließen, wann erstere geschaffen wurde. Spasovas Annahme, der Beginn der Bücherrevision sei bereits im 13. Jh. anzusetzen, ist demnach weder zu bekräftigen noch zu widerlegen.

Auch der cod. Sin. sl. 23, ein Fastentriodion, enthält einen Kolophon, der uns über die athonitische Übersetzertätigkeit Aufschluss gibt. Er stammt von einem im Zographou-Kloster wirkenden Schreiber namens Gavriil. Den Beginn des Kolophons würde ich dahingehend deuten, dass ein Starec Iosif ein „Original“⁸ des Fastentriodions übertrug, und zwar physisch – aus der Athanasios-Laura – und

⁷ Der Beginn des Kolophons lautet folgendermaßen: „ $\eta\omega\nu\langle o\rangle\mu o/v/ \omega\chi\tau\omega\iota\chi o/v/ \kappa\omicron\pi\epsilon\varsigma \sigma\tau\acute{\epsilon} b(o)g\omega\mu\tau\acute{\epsilon} + \tau\rho\omicron\upsilon\delta\iota \acute{\zeta}\epsilon \iota \iota\zeta\lambda\omega\acute{\zeta}\eta\nu\acute{\iota}\alpha$ [bis hierher Vjaz’] $\uprho\acute{\rho}\epsilon\pi\omicron\delta\omicron\beta\nu\alpha\alpha g\omicron \acute{\omega}/\tau\acute{\epsilon}/\kappa\alpha \nu\acute{\alpha}\sigma\eta g\omicron \iota\acute{\omega}(anna) \sigma\tau\acute{\alpha}\rho/\kappa\alpha \uprho\acute{\rho}\epsilon\lambda\omicron\acute{\zeta}\iota\nu/\acute{\zeta}\alpha g\omicron \iota\acute{\zeta} \gamma\rho\acute{\upsilon}\tau\sigma\kappa\upsilon/\chi\eta/ \nu\acute{\epsilon} b\lambda\gamma\alpha\rho/\sigma\kappa\upsilon\iota \nu\acute{\alpha}\sigma\eta \acute{\epsilon}\zeta\upsilon\kappa\tau\acute{\epsilon} \cdot \text{Si}\acute{\alpha} \iota\acute{\zeta}b\rho\nu\acute{\epsilon} \acute{\zeta}\omicron/v/ \nu\acute{\epsilon} s(\nu\acute{\epsilon})\tau\omicron\iota \gamma\omicron\rho\acute{\epsilon} \acute{\alpha}\theta\omicron\nu/\sigma\tau\acute{\epsilon}\iota \cdot \iota \lambda\alpha\r\nu\acute{\epsilon} b(o)g\omicron\nu\omicron\sigma\nu\alpha\alpha g\omicron \acute{\omega}/\tau\acute{\epsilon}/\kappa\alpha \nu\acute{\alpha}\sigma\eta/g/\langle o\rangle \acute{\alpha}\theta\alpha\nu\alpha\sigma\iota\acute{\alpha} : \cdot \iota\acute{\zeta}b\omicron\nu\acute{\iota}\epsilon \iota \acute{\zeta}\eta\lambda\alpha\nu\acute{\iota}\epsilon\mu\acute{\epsilon}\tau\alpha, \iota b\langle a\rangle g\langle o\rangle/d/\langle a\rangle\tau\iota\epsilon s(\nu\acute{\epsilon})\tau\alpha g\omicron d(u)\chi\alpha \cdot \uprho\acute{\rho}\epsilon\lambda\omicron\acute{\zeta}\iota \iota \iota\pi\sigma\iota\alpha \kappa\nu\acute{\iota}\gamma\upsilon, \iota\chi\acute{\zeta}\eta \iota\mu\epsilon\nu\alpha z/d/e [...]$ “ // („Ende des neuen Oktōēchos mit Gott. (Es sind) die Werke und Arbeiten unseres ehrwürdigen Vaters, des Starec Ioann, der sie aus den griechischen in unsere bulgarische Sprache übertragen hat. Nachdem er diese am Heiligen Berg Athos, in der Laura unseres gottesträgenden Vaters Athanasios ausgewählt hatte, übertrug und schrieb er mit Liebe und Eifer und der Gnade des Heiligen Geistes die Bücher, deren Namen hier (folgen) [...]).“). Es ist nicht klar, worauf sich *Siā* bezieht – vermutlich auf die danach aufgezählten Werke insgesamt. Syntaktisch gehört das Kolon *Siā ... aθanasia* jedenfalls trotz der Interpunktion zum Folgenden. Das beweist das Partizip *izbrav’šo/v/*, das nicht mit dem vorangehenden Genitiv „ $\uprho\acute{\rho}\epsilon\pi\omicron\delta\omicron\beta\nu\alpha\alpha g\omicron \acute{\omega}/\tau\acute{\epsilon}/\kappa\alpha \nu\acute{\alpha}\sigma\eta g\omicron \iota\acute{\omega}(anna) \sigma\tau\acute{\alpha}\rho/\kappa\alpha \uprho\acute{\rho}\epsilon\lambda\omicron\acute{\zeta}\iota\nu/\acute{\zeta}\alpha g\omicron \iota\acute{\zeta} \gamma\rho\acute{\upsilon}\tau\sigma\kappa\upsilon/\chi\eta/ \nu\acute{\epsilon} b\lambda\gamma\alpha\rho/\sigma\kappa\upsilon\iota \nu\acute{\alpha}\sigma\eta \acute{\epsilon}\zeta\upsilon\kappa\tau\acute{\epsilon} \cdot \text{Si}\acute{\alpha} \iota\acute{\zeta}b\rho\nu\acute{\epsilon} \acute{\zeta}\omicron/v/ \nu\acute{\epsilon} s(\nu\acute{\epsilon})\tau\omicron\iota \gamma\omicron\rho\acute{\epsilon} \acute{\alpha}\theta\omicron\nu/\sigma\tau\acute{\epsilon}\iota \cdot \iota \lambda\alpha\r\nu\acute{\epsilon} b(o)g\omicron\nu\omicron\sigma\nu\alpha\alpha g\omicron \acute{\omega}/\tau\acute{\epsilon}/\kappa\alpha \nu\acute{\alpha}\sigma\eta/g/\langle o\rangle \acute{\alpha}\theta\alpha\nu\alpha\sigma\iota\acute{\alpha} : \cdot \iota\acute{\zeta}b\omicron\nu\acute{\iota}\epsilon \iota \acute{\zeta}\eta\lambda\alpha\nu\acute{\iota}\epsilon\mu\acute{\epsilon}\tau\alpha, \iota b\langle a\rangle g\langle o\rangle/d/\langle a\rangle\tau\iota\epsilon s(\nu\acute{\epsilon})\tau\alpha g\omicron d(u)\chi\alpha \cdot \uprho\acute{\rho}\epsilon\lambda\omicron\acute{\zeta}\iota \iota \iota\pi\sigma\iota\alpha \kappa\nu\acute{\iota}\gamma\upsilon, \iota\chi\acute{\zeta}\eta \iota\mu\epsilon\nu\alpha z/d/e [...]$ “ kongruiert. Es handelt sich vielmehr um einen *dativus absolutus*, der im mittelbulgarischen Kirchenslavisch durchaus auch bei Subjektgleichheit anstatt eines *participium coniunctum* auftreten kann. Da ein Präteritalpartizip verwendet wird, ging die dadurch bezeichnete Nebenhandlung, das Auswählen der Werke in der Athanasios-Laura, der Haupthandlung, dem Übertragen, zweifellos voraus; wo dieses erfolgte, bleibt unklar. Es ist also letztlich ungewiss, wo Ioann wirkte, in der Laura oder an einem anderen Ort. Da diese ausdrücklich als Herkunftsort der Vorlagen bezeichnet wird, würde ich entgegen der gängigen Meinung annehmen, dass sie eben nicht die Wirkungsstätte Ioanns war.

⁸ Das scheint *izvodb* hier zu meinen. Miklosich gibt in seinem *Lexicon* die Bedeutungen *origo*, *archetypon*, $\acute{\alpha}\nu\tau\iota\gamma\rho\alpha\phi\omicron\nu$, *exemplar* an (außerdem $\delta\iota\omicron\rho\upsilon\acute{\xi}$, *canalis*).

sprachlich – aus dem Griechischen ins Slawische, bzw., wie es hier bereits heißt, ins Bulgarische.⁹ Gavriil erwähnt einen zweiten *izvodъ*, den ein gewisser *Zakchej Filosofъ* aus Konstantinopel gebracht habe.¹⁰ Dieser Zakchej wird in einer von Gavriil stammenden Randeintragung zum ebenfalls von ihm geschriebenen Blumentriodion cod. Sin. sl. 24 (fol. 26v) auch *Zagorjaninъ* genannt; er war also offensichtlich Bulgare (vgl. Popov 1978: 406). Der *izvodъ* des Zakchej enthielt dabei ein Synaxar, das aller Wahrscheinlichkeit nach von Nikēphóros Kállistos Xanthópoulos verfasst wurde. *Terminus ante quem* für die Entstehung des griechischen Originals ist das Alter des frühesten Zeugen, eines Bodleianus (Auct E.5.14), der auf 1305–1309 datiert wird (Taseva 2010: 31). Was die slawische Version angeht, so ist letztlich nicht völlig klar, ob Zakchej nun der Übersetzer war, oder ob er lediglich eine Handschrift mit dem griechischen Text aus Konstantinopel brachte. Ersteres ist wohl das Wahrscheinlichere.

Die slawische Übersetzung des Synaxars – ich will im Folgenden davon ausgehen, dass sie von Zakchej Zagorjanin geschaffen wurde – zeigt in Bezug auf die Behandlung von aus dem griechischen entlehnten Wortgut eine Besonderheit, die Lora Taseva als erste systematisch beschrieben hat. Anstatt etablierter griechischer Fremdwörter, darunter gerade auch *termini technici* aus dem liturgischen und religiösen Bereich, werden oftmals slawische Wörter gebraucht (Taseva 2004: 50). Als Beispiel mag das Wort εὐαγγέλιον dienen. Im Glossar zu Tasevas Edition der athonitischen Synaxar-Übersetzung (Taseva 2010) ist kein Beleg für das seit dem Altkirchenslawischen gängige Lehnwort *evangelie* verzeichnet; der Übersetzer verwendet stattdessen die Lehnübersetzung *blagovĕstie*. Diese ist zwar für das Altkirchenslawische nicht belegt, allerdings kennt der Suprasliensis *blagovĕštenie* in der Bedeutung „Evangelium“ (vgl. Miklosich’ *Lexicon palaeoslovenico-graeco-latinum* 1862–1865). Weitere typische Beispiele für Zakchejs Lehnübersetzungen sind *prъvosvĕštenikъ* für ἀρχιερέυς, *posĕtitelъ* für ἐπίσκοπος und *otъsenačĕlbnikъ* für πατριάρχης. Alle diese Lehnübersetzungen sind in Miklosich’ *Lexicon* verzeichnet, jedoch bedauerlicherweise ohne dass Belegstellen angegeben würden. *Blagovĕštenie*¹¹ und *otъsenačĕlbnikъ*¹² sind in dieser Bedeutung für ostslawische Dokumente belegt, allerdings deutlich später.

⁹ Ein Facsimile findet sich in Popov (1978: 405). Die relevante Stelle lautet: „sī izvodъ izvede starecъ s(vĕ)tyi gl(agol)emyj iŕôsifъ izъ s(vĕ)toĕ lavrŕ izъ grъčъskago ōzyka vъ našъ blъgarъskyi: i ĕšte i glagolemyi zachej filosofъ i tъ izvede drougyi iz’vo/d/ izъ c(a)rigrada vъ nemže sĕ ōbrĕtochochŕ (sic!) sъbrania ĕže drъžŕ/t/ ō/t/ farisea do vъsĕ/ch/ s(vĕ)t(yi)chъ“. // („Das Original übertrug ein heiliger Starec namens Iosif aus der heiligen Laura aus der griechischen Sprache in unsere bulgarische. Und überdies übertrug auch der sogenannte Zakchej Filosof ein anderes Original aus Konstantinopel; in diesem befindet sich das Synaxar, welches (die an den jeweiligen Tagen vorzutragenden Texte) vom (Sonntag) des Pharisäers bis Allerheiligen enthält“). Es wäre wohl möglich, *izъ s(vĕ)toĕ lavrŕ* auf Iosif zu beziehen; dafür, dass hier vom Transport einer Handschrift aus der Athanasios-Laura die Rede ist, spricht allerdings die parallele Formulierung im folgenden Satz „izvede drougyi iz’vo/d/ izъ c(a)rigrada“. Es wird hier wohl tatsächlich die Herkunft der *izvodi* hervorgehoben.

¹⁰ Siehe die vorangehende Fußnote.

¹¹ In Sreznevskijs *Materialy* sind Belege für Tetraevangelien von 1383 und 1389 verzeichnet.

¹² Der *Slovar’ russkogo jazyka XI-XVII vv.* kennt einen Beleg aus 1676.

Ich würde es also für durchaus wahrscheinlich erachten, dass es sich um von Zakchej geschaffene Neologismen handelt.

Dabei erscheint die Bildungsweise von *posëtitelb* als einigermaßen unglücklich. Hier liegt offensichtlich keine morphemgetreue Lehnübersetzung vor. Vielmehr ging der Übersetzer vom griechischen Verbum ἐπισκοπέω aus, das die Nebenbedeutung „besuchen“ hat; dies setzte er – formal korrekt – mit slawisch *posëtiti* gleich und bildete davon ein *nomen agentis*. So machte er den Bischof, eigentlich ein „Auf-seher“, zu einem *visitor*. Es nimmt nicht wirklich wunder, dass sich diese Lehnbildung nicht durchgesetzt hat. Dennoch bezeichnet erstaunlicherweise auch Grigorij Camblak in seiner Vita des Stefan Dečanski den heiligen Nikolaus, Bischof von Myra, als *posëtitelb*.¹³

Bemerkenswert ist aber auch, dass die Lehnbildungen in der athonitischen Synaxar-Übersetzung Anstoß erregten. Das von Zakchej übertragene Synaxar wurde, wie Taseva (2010: 87f.) annimmt, im dritten Viertel des 14. Jh. in Tärnovo einer Überarbeitung unterzogen, von der insbesondere auch die Lexik betroffen war. Gerade die eben erwähnten Lehnübersetzungen wurden durch die gängigen Lehnwörter ersetzt, also *prvovsëštenikb* durch *archirei*, *posëtitelb* durch *episkorb* und *otscenačëlbnikb* durch *patriarchb*.

Lehnübersetzungen wie die in der Synaxar-Übersetzung finden wir auch beim zuvor erwähnten Schreiber des cod. Sin. sl. 19, Methodie; dieser verwendet für den Terminus ὀκτώηχος die Lehnübersetzung *osmoglasnikb*¹⁴, wohingegen der von ihm kopierte erste, ältere Kolophon mit der Nennung des Starec Ioann das Lehnwort *ohtwoichb* hat.¹⁵ Auch Gavriil gebraucht in seinem Kolophon zum cod. Sin. sl. 23 eine Lehnübersetzung für griechisch συναζάριον, nämlich *sëbranie*.

Es hat also den Anschein, dass in der ersten Hälfte des 14. Jh. eine Gruppe von *literati* Fremdwörter bewusst – aber natürlich nicht konsequent¹⁶ – vermied, und stattdessen Lehnbildungen bevorzugte. Zu diesen gehörten offenkundig Gavriil, Methodie und Zakchej; der einzige Hinweis für eine genaue räumliche Einordnung ist die Erwähnung des Zographou-Klosters im Kolophon Gavriils. Die Einstellung dieser Übersetzer, Redaktoren, Kopisten zu Fremdwörtern dürfte sich von der ihrer Vorgänger wie dem Starec Ioann deutlich unterscheiden haben; aber auch ihre Nachfolger in Tärnovo, etwa der Bearbeiter des Synaxars, teilten sie nicht. Der wohl relativ enge zeitliche Zusammenhang und

¹³ Vgl. Davidov et al. (1983: 104): „Таже і с(ъ)кoнb вьздвiже oт oснoвaнii добромoу posëtitelju velikomou Nikolaou ...“ // („Danach errichtete er von Grund auf eine Kirche für den guten Bischof, den großen Nikolaus“). Die Herausgeber merken an, dass Kukulevič-Sakcinski und Šafárik entgegen der handschriftlichen Bezeugung zu *pospëšitelju* („Helfer“) emendieren wollten; sie selbst übersetzen das Wort mit *pokrovitel* („Beschützer“). Ich halte es für weitaus plausibler, dass hier tatsächlich „Bischof“ gemeint ist – es ist wenig wahrscheinlich, dass ein oder mehrere Schreiber *pospëšitelju* absichtlich oder aus Versehen durch das ungewöhnliche *posëtitelju* ersetzen.

¹⁴ Im *Slovar' russkogo jazyka XI-XVII vv.* findet sich ein Beleg für den in Moskau gedruckten Oktōēchos von 1578, in Miklosich' *Lexicon einer für das Leksikonb trejazyčnyj* des Fedor Polikarpov-Orlov von 1704.

¹⁵ Vgl. Jovčeva (2004: 232f.); auch der Sin. sl. 20 hat *osmoglasnikb* (ibid. S. 231).

¹⁶ Wörter wie *apostolb*, *angelb* sind laut dem Glossar zu Tasevas Textausgabe (Taseva 2010: 545f.) im Sin. sl. 23 vorhanden.

die einheitliche Vorgehensweise lassen vermuten, dass es sich bei diesem Purismus nicht bloß um individuelle Vorlieben einzelner Schreiber handelt.

Wie verhalten sich andere Schreiber und Redaktoren der Zeit zu Lehnwörtern? Was das Londoner Evangeliar angeht, so stellt Talev (1973: 355f.) fest, dass aus diesem Text griechische und hebräische Lehnwörter systematisch beseitigt wurden. Er merkt zu Recht an, dass es ein solches Ersetzen von Lehnwörtern durch Lehnbildungen bereits im Altkirchenslawischen gegeben habe, im Londoner Evangelium sei es aber „almost completed“ (S. 357) – das mag allerdings übertrieben sein. Die Beispiele, die Talev bringt, lassen nur selten die morphemgetreue Nachbildung des griechischen Wortes, die von den zuvor genannten athonitischen Kopisten bevorzugt wurde, erkennen. Hier finden wir zum einen gängige slawische Wörter, die anstelle von griechischen Entlehnungen verwendet werden, wie etwa *choulq* (fol. 92r) statt *vlasfmiję* (in den Vergleichshandschriften, die Talev verwendet) oder *masla* (fol. 72v) statt *olěi*. Bemerkenswerter sind aber die verdeutlichenden Übersetzungen, die in mittelbulgarischen Übertragungen aus dem Griechischen, wo sonst streng auf die numerische Übereinstimmung mit dem Original geachtet wird, nicht übermäßig häufig auftreten;¹⁷ hier haben wir jedoch *съ voněmi blagočhannymi* (fol. 267v) statt *съ aromaty* entsprechend μετὰ τῶν ἀρωμάτων (Joh. 19,40) und *starěišiny žrbьsky* (fol. 197r) statt *archierei* entsprechend οἱ ἀρχιερεῖς (Luk. 20,19).

Nun könnten die griechischen Lehnwörter bereits in der Vorlage des Londoner Evangeliums durch slawische Wörter ersetzt worden sein, und auf einige mag das auch zutreffen. Allerdings bezeichnet der Mönch Simon, der 1356 auf Geheiß des Zaren Ioann Alexander am Hof zu Tărnovo dieses Tetraevangelium abschrieb, es im Kolophon als *čtvorobl(a)gověstnikъ* (fol. 274r) bzw. *čtvorobl(a)gověstnik(ъ)* (fol. 275r). Offensichtlich zog also auch er selbst slawische Lehnbildungen griechischen Lehnwörtern vor. Wir finden den „athonitischen Purismus“ also auch in der Hauptstadt, und das nicht allzu lange bevor beim Überarbeiten der Übersetzung Zakchejs die Lehnbildungen im Synaxar wiederum durch die bekannten Lehnwörter ersetzt wurden.

Nehmen wir nun zum Vergleich die slawische *Dioptra*, ein umfangreiches byz. Lehrgedicht, das jedenfalls vor den 1360er Jahren ins mittelbulgarische Kirchenslawisch übertragen wurde.¹⁸ Hier wird etwa Triodion nicht durch das vermutlich bereits etablierte Lehnwort *triodb*, für das Miklosich' Lexicon Belege aus dem 13. und 14. Jh. kennt, wiedergegeben, sondern ganz im Stil der Puristen durch die Lehnübersetzung *trьpěsьcbь*. Im Unterschied zur Synaxar-Übersetzung Zakchejs werden für ἐπίσκοπος und πατριάρχης aber die entsprechenden Lehnwörter verwendet. Für εὐαγγέλιον und seine Derivate finden sich sowohl die Lehnbildungen mit *blagověst-*, als auch die Lehnwörter mit *evangel-*, wobei letztere allerdings mehr als doppelt so häufig vorkommen als erstere. Im Falle von ἀρχιερεύς bedient sich

¹⁷ Vgl. hierzu Fuchsbauer (2010: 195f.) und die weitere hier zitierte Literatur.

¹⁸ Vgl. Fuchsbauer (2010: 260ff.) sowie Miklas/Fuchsbauer (2013: 39 und 296).

der Übersetzer zwar nicht des Lehnworts *archirei*, aber auch nicht einer morphemgetreuen Lehnübersetzung wie Zakchejs *прѣвосвѣštenikъ*; vielmehr gebraucht er eine freiere Lehnschöpfung, und zwar *svęštennonačelnikъ*.

Ob die slawische Dioptra den ausgeprägten Purismus Zakchejs oder Simons teilt, kann ich nicht mit Bestimmtheit sagen. Falls dem nicht so ist, spräche dies dagegen, sie in deren unmittelbares Umfeld einzuordnen. Dennoch gehen wir aufgrund der Qualität der Gesamtübersetzung der *Dioptra*, die sich etwa auch von jener der im hesychastischen Umfeld des Parorie-Klosters erstellten Auszugsübersetzung stark abhebt (vgl. Fuchsbauer 2012), davon aus, dass das Werk in einem der Zentren der mittelbulgarischen Schriftkultur ins Slawische übertragen wurde, d. h. vermutlich am Athos oder in der Umgebung Tarnovos. Es wäre jedenfalls ratsam zu überprüfen, inwiefern sich das Auftreten des lexikalischen Purismus mit den bestimmter orthographischer Regelungen deckt. So ließe sich feststellen, ob wir es tatsächlich mit einer einigermaßen klar definierten Schule zu tun haben.

Insgesamt werden in der *Dioptra*, die allein schon wegen ihres Umfangs und der Vielfalt der hier verarbeiteten Texte reiches lexikalisches Material bietet, Lehnübersetzungen dem Anschein nach Lehnwörtern vorgezogen (vgl. hierzu Fuchsbauer 2013). Wo Lehnwörter verwendet werden, verfügen diese üblicherweise über eine sehr spezifische Bedeutung, wie z. B. *vysěnbъ* zu griechisch βύσσινος („linnen“, von βύσσιος = Flachs) oder *psifida* zu griechisch ψηφίς („kleiner Stein“, gemeint sind Mosaiksteinchen). Lehnübersetzungen sind in der *Dioptra* zahlreich, und viele davon formen ihr griechisches Vorbild morphemgetreu nach, wie etwa *blagopodatъnbъ* griechisch εὐμετάδοτος („freigebig“) oder *pamętozlobovati* griechisch μνησικακεῖν („sich an erlittene Übel erinnern“). Der Übersetzer bedient sich gängiger Wortbildungsmuster. Auffällig sind jene Fälle, bei denen ein und derselbe griechische Begriff mit unterschiedlichen Mitteln nachgebildet wird, wie z. B. γεωργία („Ackerbau“) durch *zemedělanie*, *zemedělstvo* und *zemoděanie* oder ἀμάθεια („Unwissenheit“) durch *nenavyknovenie*, *nenavyčenie* und *nenaučenie*. Diese Inkonsistenz beweist, dass der Übersetzer, wo er es für angebracht hielt, Lehnübersetzungen *ad hoc* schuf.

Es bleibt für mich rätselhaft, was die Neigung zum Ersetzen etablierter Lehnwörter durch Lehnbildungen hervorgerufen hat, scheint doch diese puristische Tendenz, wenn wir sie so nennen wollen, in eklatantem Widerspruch zur sonstigen Gräzisierung des Kirchenslawischen in der mittelbulgarischen Zeit zu stehen. Man wird annehmen können, dass Sprachpurismus entweder fremde Einflüsse auf die eigene Sprache eindämmen will oder dass er auf eine Steigerung der Verständlichkeit abzielt. Beides scheint mir im Falle des mittelbulg. Purismus nicht zuzutreffen, eben wegen des überwältigenden Einflusses des Griechischen und weil, wie etwa die sprachlich äußerst komplexen Werke des Patriarchen Evtimij beweisen, breite Verständlichkeit offensichtlich nicht das vorrangige Ziel der mittelbulgarischen Literaten war. Dabei war auch den Byzantinern Sprachpurismus nicht fremd.

Hochsprachliche Texte vermeiden im Unterschied zu den volkssprachlicheren fremdes Lehngut sehr konsequent; stattdessen werden gern Archaismen verwendet (vgl. etwa Ševčenko 1981: 291). Den mittelbulgarischen Sprachreformern ging es hingegen wohl um eine Nachbildung der griechischen Modelle mit den Mitteln der eigenen Sprache. Das ist, was die syntaktische Imitation der Vorlagen betrifft, evident (vgl. Fuchsbauer 2010: 170ff. und die weitere hier zitierte Literatur). Der oben behandelte Purismus weist jedoch darauf hin, dass es entsprechende Tendenzen auch bezüglich des Wortschatzes gab.

Literatur

- DAVIDOV, ANGEL; DANČEV, GEORGI; DONČEVA-PANAJOTOVA, NEVJANA; KOVAČEVA, PENKA; GENČEVA, TODORKA (eds.) (1983): *Žitie na Stefan Dečanski ot Grigorij Camblak*. Sofia: Izdatelstvo na BAN.
- FUCHSBAUER, JÜRGEN (2010): *Die Übertragung der Dioptra ins Slavische. Ein Beispiel mittelkirchenslavischer Übersetzungstechnik, dargestellt anhand des vierten Buches des Werkes*. Wien: Univ.-Diss.
- FUCHSBAUER, JÜRGEN (2012): Identifying Listening Errors in Slavonic Translations? (On Some Peculiar Mistakes in the Partial Translation of the *Dioptra*). In: Nichoritis, Konstantinos (ed.): *Kyrillos kai Methodios. Parakatathikes politismou. Praktika diethnous synedriou „I politismiki klironomia tou ergou ton agion Kyrillou kai Methodiou os paragontas enotitas me tous laous tis N.A. Europis“ (Amyntaio 21–22 Maiou 2010)*. Thessaloniki: Epikentro, 155–164.
- FUCHSBAUER, JÜRGEN (2013): The *Dioptra* as a source of Miklosich's *Lexicon Palaeoslovenico-Graeco-Latinum*. Grković-Mejdžor, Jasmina; Loma, Aleksandar (eds.): *Miklosichiana bicentennialia. Zbornik u čast dvestote godišnjice rođenja Franca Miklošiča* (= Posebna izdanja SANU knj. DCLXXXV / Odeljenje jezika i književnosti 58). Beograd: SANU, Staroslovenski odbor, 33–64.
- GOLDBLATT, HARVEY (1977): *Orthography and orthodoxy: Constantine Kostenečki's Treatise on the Letters*. Yale: Univ.-Diss.
- JOVČEVA, MARIJA (2004): Novoizvodnijat slavjanski oktoich po naj-rannija prepis v kodeksite 19 i 20 ot manastira „Sv. Ekaterina“ v Sinaj. In: Taseva, Lora et al. (eds.): *Prevodite prez XIV stoletie na Balkanite. Dokladi ot meždunarodnata konferencija Sofija, 26–28 juni 2003*. Sofija: Izdatelska kašta „GoreksPres“, 205–234.
- MIKLAS, HEINZ; FUCHSBAUER, JÜRGEN (2013): *Die kirchenslavische Übersetzung der Dioptra des Philippos Monotropos, Band 1: Überlieferung. Text der Programmata und des ersten Buches*. Wien: Holzhausen.
- MIKLOSICH, FRANZ VON (1862–1865): *Lexicon palaeoslovenico-graeco-latinum emendatum auctum*. Wien: Braumüller.
- MILTENOVA, ANISAVA (ed.) (2009): *Istorija na bälgarskata srednovekovna literatura*. Sofija: Iztok-Zapad.
- PETKANOVA, DONKA (ed.) (2003): Ezikovo-pravopisna reforma na Patriarch Evtimij. In: *Starobälgarska literatura. Enciklopedičen rečnik*. Veliko Tärnovo: Abagar.
- POPOV, GEORGI (1978): Novootkrito svedenie za prevodačeska dejnost na bälgarski knižovnici ot Sveta gora prez pärvata polovina na XIV v. In: *Bälgarski ezik XXVIII*, 5, 402–410.

- RUSEV, PEN'Ō ET AL. (1971): *Pochvalno slovo za Evtimij ot Grigorij Camblak*. Sofija: Izdatelstvo na BAN.
- ŠEVČENKO, IHOR (1981): Levels of style in Byzantine prose. In: *Jahrbuch der österreichischen Byzantinistik* 31/1 (= Akten des XVI. Internationalen Byzantinistenkongresses I/1, Wien, 4.–9. Oktober 1981), 289–312.
- SLOVAR' RUSSKOGO JAZYKA XI–XVII vv.*, 1. Band 1975. Moskva: Izdatel'stvo «Nauka».
- SPASOVA, MARIJA (2007): Pripiskata za stareca Joan i problemāt za atonskata redakcija na starobālgarskite knigi. In: Dančev, Georgi (ed.): *Sv. Evtimij, Patriarch Tārnovski, i negovata duhovna misija v Evropa. Osmi meždunaroden simpozium, Tārnovska knižovna škola, Veliko Tārnovo, 14.–16. oktombri 2004 g.* Sofija: Universitetsko Izdatelstvo „Sv.Sv. Kiril i Metodij“, 211–230.
- SREZNEVSKIJ, IZMAIL (1893-1903): *Materialy dlja slovarja drevne-russkogo jazyka po pis'mennym pamjatnikam* (1. Bd. 1893, 2. Bd. 1902, 3. Bd. 1903). Sankt Peterburg.
- TALEV, ILIJA (1973): *Some problems of the second South Slavic influence in Russia*. München: Sagner.
- TASEVA, LORA (2000): Ezikāt na prevodača Zakhej: Meždu knižovnoto nasledstvo i formalističnite tendencii na epochata. In: *Slavia* 69, 189–210.
- TASEVA, LORA (2004): Leksikalnata variativnost v južnoslavjanskite prevodi na triodnite sinaksari. In: *Palaeobulgarica* XXVIII, 4, 43–69.
- TASEVA, LORA (2010): *Triodnite sinaksari v srednovekovnata slavjanska knižnina. Tekstologično izsledvane. Izdanie na Zakheevija prevod. Slovoukazateli*. Freiburg i. Br.: Monumenta linguae slavicae dialecti veteris. Fontes et dissertationes LIV.
- VEDER, WILLIAM (1999): *Utrum in alterum abiturum erat? A study of the beginnings of text transmission in Church Slavic. The prologue to the Gospel Homiliary by Constantine of Preslav, the text On The Script and the treatise On The Letters by anonymous authors*. Bloomington/Indiana: Slavica.

Abstract

Das Makedonische und das Türkische blicken sprachgeschichtlich auf einen über 600 Jahre langen Kontakt zurück. Die Zahl der in den diachronen Korpora registrierten Turzismen in der makedonischen Sprache beläuft sich auf mehr als 3000 Elemente. In diesem Artikel wird versucht mittels einer korpusbasierten Untersuchung die gegenwärtige Verbreitung der Turzismen in der makedonischen Pressesprache näher zu bestimmen. Als Datenkorpora dienen die elektronischen Archive zweier makedonischer Tageszeitungen *Дневник* <www.dnevnik.com.mk> und *Утрински весник* <www.utrinski.com.mk>. Nach der Überprüfung von 3155 Turzismen aus der diachronen Liste wurden im heutigen Makedonischen 1083 türkische Lehnwörter festgestellt, die aber nur rund 34% der kompletten Liste ausmachen. Nur die Hälfte der belegten Elemente weist Frequenzen von über 10 Belegen auf. Trotz des quantitativen Rückganges des türkischen Lehnwortschatzes weist das moderne Makedonische türkische Lehnwörter in allen Wortarten außer Zahlwörtern auf. In der Pressesprache kommen nicht nur Inhalts-, sondern auch Funktionswörter (grammatische Marker und Derivationsaffixe) vor. Die Anwendung der gewonnenen Daten in Hinblick auf Entlehnungsskalen nach Matras und Thomason sowie die Feststellung von diversen semantischen Feldern geben Aufschluss über die Qualität und die Tiefe des Kontaktes zwischen dem Osmanisch-Türkischen und dem Makedonischen, wobei auch makedonische Daten die in den Skalen postulierten Implikationsbeziehungen weitgehend bestätigen.

0. Introduction

Among numerous groups of loanwords in the languages of South Eastern Europe Turkisms play the most prominent, but at the same time also a very controversial role. Here we use the term Turkism as a reference to linguistic elements which emerged as a result of the linguistic processes of language replication from Ottoman Turkish. On the one hand, they appear to be numerically the most significant layer of loanwords. It is also clear that the length and intensity of language contact caused a favoured usage of Turkish elements in many cases. On the other hand, the Ottoman Turkish language lost its prestige with the establishment of national states in South Eastern Europe. Various extra-linguistic factors increasingly have taken centre stage here at least since the 19th century. A major problem regarding the current status of the Turkish elements is still the lack of empirical studies.

The question of the Ottoman influence on Balkan languages attracted the attention of researchers already in the 19th century after the publication of pioneering works of Franz Miklosich (1813–1891), who was the first to analyse Ottoman-Turkish elements in the Slavic languages from the perspective of diachronic linguistics. His successors (including Skok, Škaljić, Grannes, Boretzky, Jašar-Nasteva) focused mainly on individual Balkan languages and also worked mostly in the field of diachronic linguistics. Based on their research, several diachronic lists of Turkisms were compiled for certain Balkan

languages. These lists usually include both the broadest and the largest selection of possible text corpora, thus covering the whole period of language contacts with the Ottoman Turkish language from the early 15th century till modern times. In this way, such collections of Turkisms reflect the entire history of language contact without reference to the current status of Turkish loanwords. Therefore, the items that these lists provide have a very limited usage range from the perspective of synchronic linguistics. Another point of criticism would be the fact that the data collected is often listed without further theoretical analysis, which significantly hinders further interpretations of the intensity of language contact.

With regard to modern Macedonian Koneski and Ago (cf. Ago 2006: 63) emphasize that the commonly cited number of Turkish loanwords (over 3,000 items) actually contains both still active elements and words that are no longer in use (cf. Ago 2006: 55 and Leontić 2012: 48). The theoretical analysis with regard to the current language contact research (e.g. theory of borrowing scales) has still largely failed to materialize. The present paper primarily represents an empirical investigation of Ottoman Turkish loanwords in the modern Macedonian language with a special emphasis on the language of the press (newspaper language). The empirical data have been collected in two Macedonian newspapers. In this paper we will deal first of all with the current situation of Turkish loanwords that are firmly established in the Macedonian language and with their brief analysis according to the theory of borrowing scales. One of the main objectives of this empirical study will be to create and digitize the most comprehensive diachronic list of Turkisms. These data will be used as an initial basis for further investigation of Turkish replications in the Macedonian language for the purpose of their verification in the corpus of modern language and finally for the creation of a new catalogue of currently used Turkisms.

This paper is primarily designed as an empirical study. In its first part, a short outline of the historical conditions of the Turkish-Macedonian language contact will be presented. Here we shall also discuss the Macedonian language policy after 1944 and the actual practice regarding the use of Turkisms. In the second, empirical, part, we will make an attempt to determine the current status of Turkisms in the Macedonian newspaper language on the basis of a corpus-based analysis. The obtained data will be interpreted according to the theory of borrowing scales. The relevance and necessity of this empirical study derive primarily from the fact that we still do not have a reliable database of the firmly established Turkish loanwords in the modern Macedonian language. This makes accurate statements about the intensity of language contact between Macedonian and Ottoman Turkish nearly impossible.

1. Historical overview of Turkish-Macedonian language contacts and Macedonian language policy

The first contacts with the Ottomans in Macedonia took place in the second half of the 14th century. The incorporation of the Macedonian territories into the emerging Ottoman Empire started in 1382 with the

conquest of Eastern Macedonia after the battle of Maritsa in 1371, which paved the way for the further Turkish expansion on the Balkans. Over the next decades, the territories of present-day Macedonia completely came under Ottoman rule. The settlement of Turks in the cities and the rural migration of nomadic Yörüks¹ started shortly thereafter. In the second half of the 15th century there was a significant Turkish population in some larger cities (especially Skopje, Bitola, Serres, and later Ohrid and Štip). There was also an intensive settlement of nomadic Yörüks in rural areas, especially in Western Macedonia. Some estimates suggest that in the 16th century the number of Yörüks alone rose to about 140,000 (cf. Čepreganov 2008: 140). The number of Turks was constantly rising during the 16th and 17th centuries. The shares of the Slavic and Turkish populations in urban centres changed dramatically (cf. Apostolski 1969: 235). It has been assumed that in the 16th century around a quarter or even a third of the Muslim population in the larger cities were converted Slavs, while in rural areas the proportion of the converts amounted to only about 3% (cf. Čepreganov 2008: 143). The converted groups of population had close social and cultural ties with the Turkish population. These contacts encouraged a rapid and massive language acquisition of Ottoman Turkish, which led to an early development and significant spread of the Turkish-Slavic bilingualism.

The first Ottoman Turkish loanwords began to penetrate into the Macedonian vocabulary already in the 15th century, e.g.: *амир* ('emir'), *бег* ('bey', 'beg'), *махала* ('city district'), *мусромански* ('Muslim'), *чауш* (Title in the Ottoman army; guard), *дуќан* ('shop', 'store'), *пазар* ('market', 'bazaar'), *султан* ('sultan'), *јорган* ('blanket'), *белез* ('characteristic'), *чаршија* (a quarter or a street in the city with traditional shops) (cf. Jašar-Nasteva 2001: 20). The early spread of the Turkish-Slavic bilingualism is indirectly confirmed by the presence of many Ottoman Turkish loanwords in the so-called Damaskins, texts of a religious and didactic nature based on everyday speech. In the 16th and 17th centuries the position and influence of the Church Slavonic language on the Damaskin literature was still very strong. Thus, there is little evidence of the Ottoman Turkish influence in 16th and 17th century texts, only a few Turkisms can be found there. However, in the 18th century anonymous authors largely ignored the Church Slavonic tradition, writing texts much closer to everyday language. At this point of time, the widespread penetration and incorporation of Turkisms into Macedonian Slavic texts became evident (cf. Makarijoska 2009: 195). The folk poetry also shows clear evidence that at least a receptive knowledge of Turkish was common among the different groups of population (cf. Jašar-Nasteva 2001: 28).

However, Ottoman linguistic influences in Macedonia reached their apogee during the second half of the 19th century. According to Koneski the main reason for the increasing penetration of Turkisms into the Macedonian language was the migration of the Macedonian Slavs to the cities, where Turkish played a leading role as the language of trade and cross-border communications (cf. Koneski 1986: 218f.).

¹ The term Yörük applies to a variety of Turkic-speaking pastoral nomads in both Medieval Rumelia and Anatolia.

Ottoman Turkish was the language of interethnic contact and enjoyed the highest prestige among different groups of population regardless of their religion and ethnicity. In fact, many Turkish words became part of everyday language and were actively used by 19th century authors from Macedonia. Marko Cepenkov (1829–1920) was one of the most famous collectors of fairy tales, who collected a lot of quotations, proverbs and sayings in his work. In his collections of fairy tales and everyday stories there are numerous code-switching situations in which Slavic speakers used the Turkish language depending on the respective communicative situation (cf. Friedman 2005b: 27). Thus the work of Cepenkov provides an extensive language corpus for diachronic studies of Turkish loanwords (see Ago 2006).

The authors who tried to write solely in Slavic were compelled to explain Slavic words by providing their Turkish equivalents. For instance, Joakim Krčovski (1750–1820) had to gloss Slavic *зордоцм* with the Turkish word *fodulluk* and Kuzman Šapkarev (1834–1909) *чносоѳ* with *mertebe* and *наџиџа* with *milet* (Koneski 1986: 219). Almost the entire Turkish-speaking population was concentrated in the cities. Correspondingly, in the countryside the number of bilinguals was traditionally much smaller. There is, on the whole, comparatively little evidence of the Turkish influence in rural area. This is reflected linguistically in the rareness of Turkish elements pertaining to village life (cf. Jašar-Nasteva 2001: 28).

However, already in the 2nd half of the 19th century the attitude of the Macedonian speaking intellectuals towards Ottoman Turkish loanwords changed significantly. Numerous attempts were made to ‘clean’ the language of Turkish elements. The speech of Grigor Prličev (1830–1893) at the primary school in Ohrid, which he held in June 1866, has been preserved in its original handwritten version. Prličev published this speech two months later in the newspaper “Врѳмя”. Koneski’s analysis of the two texts showed that Prličev has replaced most Turkisms with Slavic equivalents in the printed version (cf. Koneski 1966b: 206ff.). The language of Krste Petkov Misirkov (1874–1926), who published his famous work *За македонџките работи* in 1903, is also characterized by the absence of Turkish loanwords (Dimitrovski 1966: 162). This language policy remained largely unchanged during the 20th century.

It is important to note that Turkish language elements came into everyday speech firstly directly from Ottoman Turkish, which was the language of the State and government spoken by the urban elites in the larger centres, and secondly, via direct language contacts of the bilingual Slavic population with the Turkish-speaking one, who spoke different dialects. The third way was relevant only in South-Western Macedonia, where the nomadic Yörüks migrated also in rural areas (cf. Reinkowski 2002: 101). The intensive contacts between the Turkish and Slavic population resulted at least in the 19th century in a widespread bilingualism, which served as a basis for the acquisition of numerous loanwords (cf. Friedman 1995b: 58). According to Kazazis (1972: 87) the Macedonian language and the Macedonian speaking

population received the Ottoman Turkish language and cultural heritage in a similar manner as other Balkan languages:

It was largely the same Turkish elements that were borrowed into each language and even the fate of individual Turkism has been by and large the same in all languages: some Turkisms have disappeared, some have become well anchored and virtually indispensable elements of a given language, others have acquired a pejorative, ironical, and vulgar connotation and so on.

At the same time, all Balkan languages reveal many individual differences. Macedonian plays a very special role here, since it is a comparatively recently standardized language and the directed language policy and language cultivation had a prominent role in all standardization processes during the 20th century:

But there are also many differences from one language to another: Turkism which may be part and parcel of the literary variety of one language may have been considerably lowered in the stylistic scale of another or become historical words and so forth. (ibid.)

Macedonia experienced a new wave of considerable debates about the role of Turkisms after 1944. While some advocated the complete adoption of Turkisms into the emerging standard, others favoured their replacement by Slavic words. The initially prevailing opinion that Turkisms are part of the vernacular language and that Turkisms could even help to distinguish Macedonian from other Slavic languages was very soon harshly criticized by language planners. Blaže Koneski judged a translation of Molière's "La Tartuffe", which according to him included too many Turkisms, as follows: "Toa znači da go sniziš istančeniot poetski jazik na Moliera [...] do nivoto na našeto, kasabsko, čaršiško muabetenje." // ("This means that the fine poetic language of Molière is downgraded to the level of slangy vernacular speech of Čaršija and Kasaba (in our cities)"). (Friedman 2003: 16). Koneski's criticism played a crucial role for the future use of Turkisms in Macedonian. Later, however, he changed his opinion again, emphasizing that no special resistance should be organized against the use of Turkisms, since their disappearance was already predetermined, because they belonged to the bygone era of Turkish feudalism (cf. Makarijoska 2009: 207). Koneski considered Turkish loanwords to be no longer productive. They could have only very narrow and specific meanings compared to the corresponding Slavic items. These facts created natural restrictions for the use of Turkisms in modern day language (cf. Koneski 1966a: 50). Nevertheless, this revision of Koneski played only a minor role, so in the 1970s a school grammar of the Macedonian language declared that „[...] postoi strog kurs za čistenjeto na našiot literaturnen jazik od turcizmi [...]“ // ("there are large-scale purification efforts against the use of Turkisms in our literary language [...]"). (Kepeski 1972: 137).

In his late works Koneski explained once again his position towards Turkish loanwords underlining that

Sosem jasno deka nikoj ne pomisluval nitu pak se obidival da gi zamenuva onie turcizmi ili greizmi što, označuvajki izvesni predmeti ili odnosi vo našiot bit, se stanati nerazdelni sastavki od rečnikot na našiot jazik. (Koneski 1987: 77). // (“It is absolutely clear that nobody intended or tried to replace those Turkisms or Hellenisms which denote certain objects or relations of our daily life and which have become an inseparable part of the Macedonian lexicon.”)

According to Troebst, and to some extent also to Friedman, Macedonia is the only country in the Balkans where the perception of the Ottoman-Islamic past and its cultural and linguistic heritage remains largely positive (cf. Troebst 2005: 63; Friedman 2005b: 33). The opposition "Christian/Muslim" in the Balkans forms some kind of an inter-ethnic boundary. The Muslim population has often been seen as a remnant of the Ottoman heritage and a hazardous alien element (Turks/Immigrants) in most of largely Christian states. The negative stereotypes about the Ottoman era (“Ottoman yoke”) have become an integral part of the national historical image and hence, this idea has manifested itself in the cultural memory of many people largely due to their school education. These widespread views can also have a lasting influence on language usage. Actual processes in contemporary Croatia could be regarded as an example of this policy (cf. Friedman 2005b: 35). However, Macedonia is a special case in every respect in the Balkans:

In Makedonien beziehen sich zwar Historiker ebenfalls auf in Jahrhunderten osmanischer Herrschaft verursachte regionale Entwicklungsblockaden, doch trifft dies nicht für die Öffentlichkeit und politische Kultur des Landes zu.² (Troebst 2005: 63).

The negative stereotypes were not, and largely still are not, present in the official Macedonian history and its public perception. Conversely, Muslim scholars, such as, for example, Evliya Çelebi/Evlija Čelebi, have often been associated with positive stereotypes such as “wisdom”, “cosmopolitanism” and “open-mindedness”. This had a consequent impact on the Macedonian everyday speech where Turkish loanwords still remain strong (cf. Koneski 1998: 146).

A similar observation has been made by Kramer who analyzed the language of the Bulgarian original of the book of Aleko Konstantinov *Бай Ганьо* and its Macedonian translation: in the Macedonian translation there were significantly fewer explanations of the same Turkish loanwords than in the Bulgarian original (101 explanatory comments in Bulgarian and 40 in Macedonian). The reason for these differences is a better preservation of Turkisms in Macedonian (Kramer 1992: 50):

[...] many of the discrepancies in the number may reflect a tendency for Turkism to play a livelier role in Macedonian [...] contemporary speech and, the consequent necessity to assume a greater knowledge of their meanings on the part of speakers of these languages³
[...]

² In Macedonia historians also refer to regional development hindrances caused by age-long Ottoman rule, but this does not apply to the public and political culture of the country.

³ This refers to Macedonian and Albanian, in contrast to Bulgarian.

However, she also notes that no reliable conclusions could be drawn, because of the small size of this sample. As regards the other styles, if there is no need of proximity to colloquial style, the use of Turkisms remains very restricted. This is exactly the main aim of our empirical study. The attitudes of the language-policy actors towards the Ottoman legacy were and still remain ambivalent. Depending on the period of time and on personal views they may vary between strong puristic tendencies and a certain nostalgia towards the use of Turkish elements.

2. Empirical corpus study

The sources of the data material are the electronic archives of the Macedonian daily newspapers *Дневник* (www.dnevnik.com.mk) and *Утрински весник* (www.utrinski.com.mk)⁴. In Macedonia both of them are regarded as high quality newspapers with a wide range of topics and a certain openness to almost all spheres of life. They are available both in print and online versions and thus address a large and diverse audience using a common and contemporary language. Lexical items appearing in these papers, among them also Turkisms, are easily understood by average readers.

Both newspapers have websites with integrated search engines, which provide reliable search results. Therefore, their archives can be used as an electronic data corpus. However, it is not possible to obtain the total number of searched tokens. Instead, only the number of articles in which the search word appears at least once is shown in the results. Nevertheless, the results of validation of Turkish replication and their frequencies are quantitatively little affected, because the comparison variables remain constant. A major advantage of this language corpus is its accessibility, which allows an easy verification of the empirical data.

The newspaper archives contain all electronic editions of the *Дневник* since March 2006 till March 2014 and of the *Утрински весник* since October 2006 till March 2014. The size of the language corpus can be estimated only very roughly. The number of tokens of the high frequency conjunction *и* ‘and’ was about 240,000 and 175,000 in *Дневник* and *Утрински весник*, respectively. Thereby the total number of articles contained in the archive was around 415,000 in March 2014.

For our purpose we created a comprehensive diachronic list of Turkisms in alphabetical order from *аалару* to *иурун*. It is based on the list of Jašar-Nasteva (Jašar-Nasteva 2001: 248-266), which has

⁴ The Macedonian linguistic community does not have a linguistic corpus which stores the diachronic language and its synchronic material and which could be conducted as a national linguistic corpus of Macedonian, although several attempts to create Macedonian corpora were made in recent years. This problem plaguing empirical research of Macedonian is often solved by creating individual linguistic corpora (cf. Leontić 2011).

been supplemented by additional entries from the *Digitalen rečnik na makedonskiot jazik*⁵. Our complete list thus contains 3,155 lexical entries.

A well-known problem in corpus linguistics is the homonymy of word forms. The diachronic list of Turkisms included a number of words that coincide with proper names (geographical or personal names) and also with other Macedonian word forms. If the number of the searched items in both archives surpassed 100, these words were not included in the final list of Turkisms. Some examples are: *бурма/Бурма* (‘wedding ring’, geographic name), *тула/Тула* (‘brick’, geographic name), *јасмин/Јасмин* (‘jasmine’, first name and surname), *машиа/Машиа* (‘pliers’, first name), *соба/соба* (‘oven’ as Turkish loanword, ‘room’ as Hungarian (through Serbo-Croatian) loanword) or *имам/имам* (1. person singular present of *има*, ‘I have’), *чувал/чувал* (‘sack’, perfect or renarrative of *чува*, ‘to guard, protect’). This relatively small number of words requires a separate analysis.

The list of Jašar-Nasteva also contains some idiomatic expressions which are alphabetically ordered. Since phrases are outside of our interest and the integrated search engines do not allow searching for a phrase, these expressions were not taken into consideration. Even if there are texts with historical and partly dialectal content in both newspaper archives, most of the texts reflect current language use, thus providing access to firmly established borrowed elements. The present study therefore aims to present a snapshot of the present day Macedonian language.

Finally, we would like to note that the term “newspaper language” (“language of the press”) is used as a purely linguistic concept to distinguish language used in print media and online newspapers from language used in other circumstances. This is not to imply that there is a single newspaper language with specific language peculiarities that could apply to the entire press in general (cf. Holland 2007: 31).

2.1. Lexical replications

The results of the empirical data analysis decisively confirm the assumption that there is a sharp decline of actively used Turkish loanwords in Macedonian. Of the 3,155 analysed words only 1,083 or 34.3% were found in the corpora, with almost half of them (49,5%) occurring with very low frequency of only 1 to 10 examples. Moreover, the number of words with up to 5 text examples (tokens) prevails. Tab. 1 illustrates the frequency of the elements with up to 10 examples in the corpus.

Tab. 1: Distribution of low frequency Turkisms by number of text examples

| No. of text examples (tokens) | No. of words (types) |
|----------------------------------|-------------------------|
| 1 | 185 |

⁵ This dictionary of the Macedonian language is based on multi-volume academic works, amongst others: *Rečnik na makedonskiot jazik* od Blaže Koneski, Tom 1-3, Detska radost, 1994 and *Tolkoven rečnik na makedonskiot jazik*, Tom 1-4, izdaden od Institutot za makedonski jazik “Krstе Misirkov”, 2003–2009. For more information, see: <<http://makedonski.info/impresum>> [accessed on 2015-03-07]

| | |
|--------|-----|
| 2 | 87 |
| 3 | 67 |
| 4 | 48 |
| 5 | 32 |
| 6 – 10 | 116 |

In our sample the number of more frequently occurring elements (> 10 examples) is only 17.3% of the total of Turkisms that have been registered in the diachronic word lists of the Macedonian language. The number of high-frequency words (> 1,000 text examples) is limited to 32 elements. There are 889 nouns (82%), 91 adjectives (8.4%) adjectives, 46 verbs (4.3 %), 28 adverbs (2.6%), 23 particles (2.1%) particles of the total number of words in the corpus. In the language corpus there are also function words: 3 conjunctions (0.3%), 2 pronouns (0.2%) and one preposition (0.09%). Only numerals were not found at all. However, nouns such as *дузина* ('dozen') or *череќ* ('quarter') are at least semantically very close to the class of numerals. The 15 most common Turkisms in Macedonian are represented by 9 nouns, 2 particles, 2 conjunctions, one verb, and one adverb.

Tab. 2: The most common Turkisms in Macedonian

| Word | No. of tokens ⁶ | Part of speech | Meaning ⁷ |
|-------------|----------------------------|----------------|---|
| 1. пари | 35,531 | noun | платежни средства од метал или од хартија ('money') |
| 2. барем | 14,653 | particle | изразува извесно ограничување ('at least') |
| 3. пазар | 13,599 | noun | место каде се тргува ('bazaar', 'market') |
| 4. ама | 9,319 | conjunction | но ('but') |
| 5. кафе | 3,642 | noun | кафе ('coffee') |
| 6. боја | 3,412 | noun | боја, фарба ('colour') |
| 7. пиштол | 3,239 | noun | огнено оружје ('pistol') |
| 8. баш | 2,842 | adverb | баш ('exactly', 'just'; 'main'; 'definitive') |
| 9. парче | 2,580 | noun | дел од нешто ('piece') |
| 10. кале | 2,493 | noun | кале; тврдина, град ('castle', 'fortress') |
| 11. макар | 2,349 | particle | барем, најмалку ('at least') |
| 12. темели | 2,266 | verb | заснове ('to base') |
| 13. чаршија | 2,128 | noun | улица или дел од градот со дуќани и работилници ('čaršija') |
| 14. ем | 2,067 | conjunction | и, притоа; и... и ('and') |
| 15. кат | 1,944 | noun | дел од куќата меѓу два тавана; етаж во зграда ('floor') |

2.2. Nouns

Nouns form numerically the largest group of Turkish borrowings which have been preserved in modern Macedonian. On the one hand, this could be explained by the fact that nouns are the largest class of words

⁶ As already mentioned, this figure does not correspond to the number of tokens, but to the number of texts in which the item occurs.

⁷ According to *Digitalen rečnik na makedonskiot jazik*, available on-line on: <www.makedonski.info>; English translations according to *Digitalen rečnik na makedonskiot jazik* <<http://www.idividi.com.mk/recnik>>, *Makedonsko-Angliski rečnik* <<http://www.time.mk/?q=&search=dict>> [all Url's accessed on 2015-03-07].

in the lexicon. On the other hand, these results also confirm the theory of borrowing scales of Thomason, Matras and Moravcsik, according to whom nouns are the first class of words affected by processes of borrowing. The most common borrowings are *пари* ('money'), *пазар* ('market', 'bazaar'), *кафе* ('coffee'), *боја* ('colour') and *пиштол* ('pistol', 'gun'). These words do not have native Slavic counterparts in Macedonian. *Пари* is the most frequent element of the entire list with a total amount of 35,531 text examples. The high frequencies of the first two nouns can partly be explained by their semantics and relevance to the newspaper language. *Пиштол* is a loanword of Italian origin, which was borrowed via Ottoman Turkish. This fact can be clearly seen in its phonetic transformations in Turkish. Considering other loanwords we can recognize Turkish cultural influences on architecture and everyday city life, as is proven by words like *кале* ('fortress'), *чаршија* (borough with traditional shops), *кат* ('floor'), *маало* ('neighbourhood'), *џамија* ('mosque'), *ѓубре* ('garbage'), or even new economic activities such as the widespread cultivation of tobacco (especially in Pelagonija), which was first introduced during the Ottoman rule: *тутун* ('tobacco'). The high-frequent nouns *парче* ('piece') and *топка* ('ball', 'globe') belong to the everyday vocabulary, but do not have native Slavic equivalents.

Tab. 3: Most frequent nouns

| Word | No. of examples | Meaning |
|------------|-----------------|--|
| 1. пари | 35,531 | платежни средства од метал или од хартија ('money') |
| 2. пазар | 13,599 | пазар ('market', 'bazaar') |
| 3. кафе | 3,642 | кафе ('coffee') |
| 4. боја | 3,412 | боја, фарба ('colour') |
| 5. пиштол | 3,239 | огнено оружје ('pistol') |
| 6. парче | 2,580 | дел од нешто ('piece') |
| 7. кале | 2,493 | кале; тврдина, град ('castle', 'fortress') |
| 8. чаршија | 2,128 | улица или дел од градот со дуќани и работилници ('čaršija') |
| 9. кат | 1,944 | дел од куката меѓу два тавана; етаж во зграда ('floor') |
| 10. тутун | 1,707 | тутун ('tobacco') |
| 11. топка | 1,534 | геометриско тело; планетата Земја; предмет со сферична форма, валчест предмет; спортски реквизит ('ball'; 'globe') |
| 12. џеб | 1,512 | дел од облека ('pocket') |
| 13. маало | 1,403 | одделен дел на градска или селска населба ('neighbourhood') |
| 14. џамија | 1,395 | џамија ('mosque') |
| 15. ѓубре | 1,379 | отпадоци, нечистотија, смет ('garbage') |

Two nouns *кафе* and *боја* form to some extent the core of well-established word families in Macedonian, which include multiple words. There is yet another loanword with the meaning colour in Macedonian: *фарба* is a borrowing of German origin. However, only 6 examples of this word have been found in the entire corpus compared to 3,412 of *боја*.

Tab. 4: Derivations of *кафе*

| Word | No. of | Meaning |
|------|--------|---------|
|------|--------|---------|

| | examples | |
|--------------|-----------------|--|
| 1. кафе | 3,642 | кафе ('coffee') |
| 2. кафеав | 54 | што има боја на кафе ('brown') |
| 3. кафеане | - | локал каде што се пие и се јаде ('café') |
| 4. кафеана | 655 | локал каде што се пие и се јаде ('café') |
| 5. кафеанче | 15 | мала кафеана ('small café') |
| 6. кафеански | 85 | што има врска со кафеана ('café', attr.) |

Tab. 5: Derivations of *боја*

| Word | No. of examples | Meaning |
|---------------|------------------------|--|
| 1. боја | 3,412 | боја, фарба ('colour') |
| 2. бојадисува | 18 | направи нешто да добие определена боја ('to paint') |
| 3. бојадиса | 20 | направи нешто да добие определена боја ('to paint') |
| 4. бојација | 3 | занаетчија што се занимава со боење на разни материјали ('dyer') |
| 5. бојосува | 18 | направи нешто да добие определена боја ('to paint') |
| 6. бојоса | 11 | направи нешто да добие определена боја ('to paint') |

Other nouns reflect the intensity of contact with Ottoman Turkish: *домат* ('tomato'), *ислам* ('Islam'), *бакар* ('copper'), *бунар* ('well'), *дузина* ('dozen'), *барут* ('gunpowder'), *бербер* ('hairdresser'), *занает* ('craftsmanship'), *занаетчија* ('craftsman'), *зејтин* ('oil'), *афион* ('poppy'), *кајсија* ('apricot'), *јоргован* ('lilac'), *бакиши* ('tip'); Ottoman titles and other words associated with Ottoman feudalism (*бег*, *беј*, *емир*, *јаничар*, *дервиш*, *ага*, *вилајет*); designations for a variety of foods (*бурек*, *баклава*, *алва*, *боза*, *ѓеврек*, *јогурт*, *кајгана*, *кавурма*) and musical instruments (*гајда*, *зурла*, *дајре*, *дудук*, *кавал*). These nouns belong to a number of different semantic fields, which can be taken as proof for a comprehensive and intensive contact with the Ottoman everyday culture.

In summary, we have selected the 10 most important semantic fields based on the frequency of the borrowed elements. These postulated semantic fields confirm once again the fact that the language contacts between the Slavic and Turkish population mainly took place in the cities. Obviously, only a few of the firmly established more frequent elements are connected with village life. The lexical elements which were closely connected to the life in the Ottoman Empire have become largely obsolete. These include, for example, words linked to the Ottoman administrative system (*нашалак*, *спашлак*, *везир*, *чауш*, *каза*). They appear occasionally only in historical texts.

Tab. 6: Important semantic fields

1. Eating/Drinking:

Кафе ('coffee'), *шеќер* ('sugar'), *шише* ('bottle'), *чај* ('tea'), *ракија* ('rakia'), *домат* ('tomato'), *чинија* ('plate'), *јогурт* ('yogurt'), *чорба* ('soup'), *портокал* ('orange'), *пита* ('pita', 'pie'), *бурек* ('burek', 'börek'),

лимон ('lemon'), *баклава* ('baklava'), *бубрег* ('kidney'), *тава* ('pan'), *зејтин* ('oil'), *сарма* ('sarma', 'roulade'), *софра* ('table'), *мезе* ('side dish'), *тенџере* ('pot'), *локум* ('turkish honey'), *филџан* ('cup'), *менџија* ('dish').

2. Crafts/Commerce:

Пару ('money'), *пазар* ('market', 'bazaar'), *дуќан* ('shop'), *занает* ('craft'), *алат* ('instrument', 'tool'), *тезга* ('stand', 'counter'), *пазарење* ('buying', 'selling'), *карван* ('caravan'), *трамна* ('barter'), *сефте* ('first purchase, sale'), *кусур* ('change'), *муштерија* ('customer'), *вересија* ('credit'), *танија* ('property proof'), *бакиши* ('tip'), *занаетчија* ('craftsman'), *занаетчиство* ('handicraft'), *чирак* ('apprentice').

3. Clothes/Cosmetics/Jewellery:

Џеб ('pocket'), *синџир* ('necklace'), *фустан* ('skirt'), *торба* ('bag'), *панталони* ('trousers'), *чанта* ('briefcase'), *костум* ('suit'), *чорани* ('socks'), *чадор* ('umbrella'), *костим* ('costume'), *ќесе* ('purse'), *копче* ('fastener'), *шал* ('shawl'), *шамиија* ('headscarf'), *ѓердан* ('necklace'), *пудра* ('powder'), *кадифе* ('velvet').

4. Architecture/City life:

Кале ('fortress'), *чаршија* ('čaršija'), *мало* ('neighbourhood'), *комишија* ('neighbour'), *џамија* ('mosque'), *ан* ('inn'), *сокак* ('lane'), *кула* ('tower'), *ќор-сокак* ('dead end'), *кирија* ('rent'), *мегдан* ('place'), *темелник* ('foundation stone'), *амам* ('hamam'), *чеиша* ('fountain'), *мермер* ('marble'), *бунар* ('well'), *копак* (large mansion/residence), *меана* ('tavern').

5. Leisure/Entertainment:

Кафеана ('café'), *муабет* ('talk'), *амам* ('hamam', 'spa'), *кукла* ('doll'), *кавал* ('flute'), *тамбура* ('tambour'), *бербер* ('hairdresser'), *чалгија* ('music instrument'), *лакрдџија* ('talk', 'joke'), *гајда* ('bagpipe'), *комар* ('gamble').

7. Interior/Design:

Кат ('floor'), *килим* ('carpet'), *подрум* ('basement'), *орман* ('cupboard'), *ќебе* ('blanket'), *ошак* ('chimney', 'fireplace'), *сандак* ('box'), *чардак* ('oriel'), *таван* ('ceiling'), *одаја* ('room'), *ќумбе* ('oven'), *софра* ('table').

8. Fauna and flora:

Тутун ('tobacco'), *памук* ('cotton'), *кајсија* ('apricot'), *афион* ('poppy'), *портокал* ('orange'), *сусам* ('sesame'), *лале* ('tulip'), *лимун* ('lemon'), *патлиџан* ('eggplant'), *боранија* ('string beans'), *сунѓер* ('sponge'), *мисирка* ('turkey'), *бадем* ('almond').

9. Weapons/Military:

Пиштол ('pistol'), *куриум* ('bullet'), *барут* ('gunpowder'), *аскер* ('army', 'soldier'), *ендек* ('ditch'), *фишек* ('bullet').

10. Abstract nouns:

Темел ('base'), *гајле* ('worry'), *абер* ('notion'), *ќеф* ('pleasure', 'wish'), *мерак* ('desire'), *раат* ('relaxation'), *дереџе* ('situation'), *простотилак* ('ignorance').

The vast majority of even frequent borrowings has native Slavic counterparts. Their Slavic equivalents usually have much higher frequencies. Cf. Tab. 7 for details.

Tab. 7: Frequencies of Turkisms vs. Slavic equivalents

| Turkism | No. of examples | Slavic equivalent | No. of examples |
|------------|-----------------|-------------------|-----------------|
| 1. темел | 935 | основа | 21,657 |
| 2. мегдан | 748 | плоштад | 3,863 |
| 3. муабет | 659 | разговор | 6,558 |
| 4. коше | 598 | агол | 1,543 |
| 5. дуќан | 532 | продавница | 1,687 |
| 6. алка | 530 | врска | 18,538 |
| 7. гајле | 492 | грижа | 3,089 |
| 8. фајде | 286 | корист | 7,101 |
| 9. нишан | 238 | знак | 8,195 |
| 10. дереце | 233 | состојба | 18,656 |
| | | положба | 3,381 |

Such semantic doublets coexist as distinct, yet near-synonymous lexemes in modern Macedonian. Turkish loanwords usually contain additional components. Typically, they pertain to lower stylistic levels or are semantically restricted, as, for instance, *дуќан* – *мала, занаетчиска продавница често за сопствените производи* ('small shop, often crafts shop'), *гајле* – *грижа* ('worry', informal), *дереце* – *состојба, обично неподнослива, тешка* ('condition', usually intolerable condition'), *мегдан* – *плоштад во населено место* ('square in a populated place, archaic'), *коше* – *агол* ('angle', informal), *фајде* – *корист* ('benefit', colloquial). These Turkisms are frequently found in colloquial language. As far as the written language is concerned they are used in texts with a certain approximation to the vernacular. In some (raw) cases there are also Turkish loanwords that are more frequent than their Slavic equivalents (cf. Tab. 8.).

Tab. 8: Frequencies of Turkisms vs. Slavic equivalents

| Turkism | No. of examples | Slavic equivalent | No. of examples |
|-----------|-----------------|-------------------|-----------------|
| 1. будала | 175 | глупак | 15 |
| 2. алат | 385 | прибор | 215 |
| | | орудие | 190 |
| 3. кодош | 281 | клеветник | 7 |

All borrowed nouns have been fully adapted to the morphological system of Macedonian due to the long period of time since their first occurrence in Macedonian.

2.3. Verbs

The replicated verbs have also been fully integrated into the morphological system of the recipient language so that even the Slavic aspectual opposition emerged, cf. perfective *бендиса* and imperfective

бендисува ('to like'), perfective *бојадиса/бојоса* and imperfective *бојадисува/бојосува* ('to colour'), perfective *ќару* and imperfective *ќарува* ('to benefit'), perfective *накалеми* and imperfective *накалемува* ('to transplant', 'to graft'). The competing Slavic equivalents are usually much more frequent. So there are only four examples of the verb *борчи* ('to owe') in the entire corpus, while its Slavic equivalent *должи* amounts up to 4,991 examples, correspondingly *куртули* ('to save') is attested 112 times, *спаси* 4,696 times; *мајтани* ('to joke') 36 times, *шегува* 371 times; *докусури* ('to destroy') 29 times, *уништи* 2,463 times. The most common verbs are shown in Tab. 9.

Tab. 9: Most frequent verbs

| Word | No. of examples | Meaning |
|--------------|-----------------|--|
| 1. темели | 2,266 | заснове ('to base') |
| 2. текне | 397 | (се) досетува; паѓа на ум ('to remember') |
| 3. ќари | 179 | спечали, заработи, доби ('to profit') |
| 4. пишмани | 128 | колеба, кае, менува мислење ('to regret') |
| 5. куртули | 112 | спаси, најде спас ('to save, 'to rescue') |
| 6. есапи | 96 | оценува, искажува свое мислење; зема предвид; смета ('to calculate'; 'to value') |
| 7. бендиса | 62 | се допадне (некому, сам себеси, еден на друг) ('to like') |
| 8. инаети | 44 | прави, тера инает, се однесува тврдоглаво ('to defy') |
| 9. бендисува | 41 | се допаѓа (некому, сам себеси, еден на друг) ('to like') |
| 10. ќарува | 38 | печали, добива, заработува ('to profit') |
| 11. мајтапи | 36 | се шегува, се подбива, се потсмева ('to joke') |
| 12. докусури | 29 | уништи, упропасти до крај ('to destroy') |
| 13. осакати | 24 | направи да биде сакат ('to injure') |
| 14. бојадиса | 20 | направи нешто да добие определена боја ('to colour') |
| 15. пелтечи | 19 | тепка ('to stutter') |

The verbs *бојадиса/бојоса* and *бојадисува/бојосува* have no Slavic equivalents in Macedonian. Their low frequency in the corpus of newspaper language can be probably explained by their semantics.

2.4. Adjectives and adverbs

Matras (2009: 187) points out that in the linguistic system adjectives and adverbs have a smaller inventory of types than nouns and verbs. Therefore, the occurrence of borrowed adjectives or adverbs indicates a higher intensity of the language contact. Among the replicated adjectives and adverbs there are those that still have not been integrated into the morphological system of the Macedonian language. So they remain invariable following the structural pattern of their origin language Turkish. Examples are, cf. *баишка економисти* – *başka ekonomistler* ('special economists'), *баишка адвокати* – *başka avukatlar* ('special lawyers') or *аурлија работа* – *hayırlı işler* ('good, happy work'), *аурлија назар* – *hayırlı*

pazar(lar) ('good sales'). Such adjectives often function also as adverbs. The analysed corpus contains 119 adjectives and adverbs, which in general have higher frequencies than verbs.

Tab. 10: Most frequent adjectives and adverbs

| Word | No. of examples | Part of speech | Meaning |
|-----------------|-----------------|-----------------|--|
| 1. баш | 2,842 | adj./adv./part. | баш ('exactly', 'just'; 'main'; 'definitive') |
| 2. темелно | 932 | adverb | основно, базично, солидно, во детали ('thorough') |
| 3. ќор | 774 | adjective | слеп ('blind') |
| 4. тутунски | 763 | adjective | што се однесува на тутун ('tobacco', attr.) |
| 5. муслимански | 674 | adjective | со врска на муслиманство ('Moslem', attr.) |
| 6. исламски | 641 | adjective | со врска на ислам ('Moslem', attr.) |
| 7. газе | 482 | adjective | свеж ('fresh') |
| 8. ептен | 448 | adverb | сосем ('entirely') |
| 9. бадијала | 395 | adverb | бесплатно, напразно, попусто, без корист ('free of charge') |
| 10. занаетчиски | 177 | adjective | што се однесува на занаетчија или занает ('craftsman's') |
| 11. башка | 137 | adj./adv. | друг, посебен ('other') |
| 12. алал | 94 | adverb | желба некому нешто да му биде убаво ('bless', 'bravo') |
| 13. подрумски | 90 | adjective | од подрум ('basement', attr.) |
| 14. кафеански | 85 | adjective | што има врска со кафеана ('café', attr.) |
| 15. топтан | 71 | adverb | се вкуп, се заедно, не поединечно, не на парче ('wholesale') |

The high frequency adverb *баш* ('at least') has no corresponding Slavic equivalent. The same applies to the adjectives *тутунски* ('tobacco', attr.), *муслимански* ('Muslim', attr.), *исламски* ('Islamic'), and *занаетчиски* ('handicraft', attr.). The two adverbs *ептен* ('completely', 'totally') and *бадијала* ('free of charge') are less frequent than the autochthonous Slavic words.

2.5. Other parts of speech

In addition to the borrowing of content words such as nouns and verbs there is also evidence of the replication of pure function words in Macedonian. The conjunction with the contrastive meaning *ама* ('but') and its variant *ами* are common in Macedonian especially in colloquial speech, though they are clearly outnumbered by the more formal conjunction *но* (164,028 examples) in our corpus.

Tab. 11: Conjunctions

| Word | No. of examples | Meaning |
|--------|-----------------|--------------------------|
| 1. ама | 9,319 | но ('but') |
| 2. ем | 2,067 | и, притоа; и...и ('and') |
| 3. ами | 535 | но ('but') |

To the class of function words belong also numerous particles that were borrowed from Turkish into Macedonian. The particle *барем* ('at least') is the second most frequent lexeme of the total list of Turkisms.

Tab. 12: Particles

| Word | No. of examples | Meaning |
|------------|-----------------|---|
| 1. барем | 14,653 | најмалку, макар, изразува извесно ограничување ('at least') |
| 2. ајде | 1,411 | ајде ('let's (go)', 'come (on)') |
| 3. макар | 2,349 | барем, најмалку, за искажување целосна решеност за извршување на некое дејство ('at least') |
| 4. демек | 891 | значи, то ест, да речеме, да претпоставиме, божем ('that is') |
| 5. аман | 378 | молам! ('please!') |
| 6. нејсе | 243 | честичка за изразување согласност, помирување со нешто, така нека е, било што било ('whatever') |
| 7. белким | 75 | изразува можност, веројатност, претпоставка, надеж ('maybe') |
| 8. гоа | 59 | божем ('as if') |
| 9. машала | 41 | за изразување восхит, пофалба, браво! ('bravo') |
| 10. аферим | 37 | за изразување восхит, пофалба, браво! ('bravo') |

Among these particles only *ајде/ајдете* has no real equivalents having relatively high frequencies in the newspaper language, although the use of these particles is actually limited to the domain of spoken language only. *Ајде* can be used either as hortative particle or as independent expression and thus as interjection with which the speaker encourages the listener to perform a specific action ('Let's (go)!'). As is well known, the Turkish particle *haydi* has been replicated into Bulgarian, Serbo-Croatian, Romanian, Albanian and Slovene.⁸ Further replications which can function as interjections would be the following: *аман* ('please'), *аманте* ('please'), *аферим* ('bravo'), and *дур бакалам* ('wait', 'let's see'), *зиди* ('look now'). Most of the other particles are used relatively rarely, since they all have Slavic equivalents or equivalents borrowed from other languages that have a significantly higher degree of occurrence on the frequency scale.

Tab. 13: Frequencies of Turkisms vs. their equivalents

| Turkism | No. of instances | Equivalent | No. of examples |
|-----------|------------------|---------------------|------------------|
| 1. барем | 14,653 | најмалку | 18,495 |
| 2. макар | 2,349 | најмалку | 18,495 |
| 3. аман | 378 | молам | 1,239 |
| 4. белким | 75 | веројатно можеби | 17,588 28,145 |
| 5. гоа | 59 | божем | 713 |
| 6. аферим | 37 | браво (loanword) | 842 |

⁸ The Turkic loanword *ајда* and very rarely *зайда* (pl. *ајдате*) is also widespread in the Russian colloquial language.

These particles can be viewed as discourse markers. It is often considered that discourse particles can be easily transferred to other languages having a high degree of borrowability in language contact situations (cf. Riehl 2009: 93 or Blankenhorn 2003: 75). Actually, there is a relatively large number of borrowed discourse markers in Macedonian. However, most of these elements have low frequencies, which could be partially explained by the fact that they are mainly used in the colloquial language. Finally, it should be noted that pronouns and prepositions are also represented in the corpus of Turkisms. The negative pronoun *нич* ('nothing') which indicates the non-existence of people or things is represented with 250 text examples, and the local preposition *каришу* ('opposite') with 40 examples. However, their Slavic counterparts have significant higher frequencies. *Ничто* ('nothing') scored 20,753 examples and *спроти* ('opposite') has been used 3,197 times. Another pronoun that is without any doubt already out of use has been found only once in a literary letter. It is *филанкишиџа*⁹ ('someone'), which appears in the following context: "*преведено од македонски од филанкишиџа*" (translated from Macedonian by someone/person X).

2.5. Grammatical replications

Macedonian possesses not only a significant lexical component of Turkish origin, but there are also a few grammatical phenomena that have developed as a result of a prolonged language contact situation. Among these phenomena are seven nominal suffixes: **-џија**: *боџаџија* (*boyacı*, 'dyer'), *јабанџија* (*yabancı*, 'foreigner'); **-чија**: *кебанчија* (*kebarci*, 'barbecue cook'), *саатчија* (*saatçi*, 'watchmaker'); **-џик**: *капиџик* (*kapıcık*, 'small door'), *арпаџик* (*arpacık*, 'onion'); **-лија**: *амајлија* (*hamaylı*, 'lucky charm'), *каismetлија* (*kismetli*, 'fortunate man'); **-лак**: *аџамилак* (*acetilik*, 'inexperience'), *калабалак* (*kalabalık*, 'crowd'); **-тар**: *бајрактар/барјактар* (*bayraktar*, 'standard bearer'); **-суз**: *баксуз* (*bahtsız*, 'bad luck'). These suffixes except **-суз** have already been mentioned by Miklosich in his 19th century works.

Some of these Turkish suffixes do not only occur together with words of Turkish origin, but remain productive and can also be combined with Slavic Macedonian words. This fact clearly shows that they can be treated as part of grammatical borrowings. Some of these nominal suffixes can be combined with indigenous words: e.g. indigenous stem with the Turkish suffix **-лак** vs. indigenous stem with Slavic suffixes **-ство/-ина**: *домакинлак* vs. *домакинство* ('household'), *ѓаволитилак* vs. *ѓаволитина* ('devilry'); indigenous stem with the Turkish suffix **-џија** vs. indigenous stem with the Slavic suffix Stamm **-ец**: *ловџија* vs. *ловец* ('hunter') or also **-лија** vs. **-ак**: *новајлија* – *новак* ('newcomer'). Moreover these nominal suffixes remain productive even in combination with foreign stems e.g.

⁹ Slavic equivalents are *тој* и *тој*, *некој*.

пубертетлија ('teenager'), *филмаџија* ('film maker', 'cinophile'), *узурпаџија* ('usurper'), *фестивалџија* ('participant of festivals').

Some of the borrowed adjectives behave in Macedonian syntactically exactly the same way as in the donor language. Syntactical processes of putting adjectives and nouns together to form a noun phrase are uncommon in standard Macedonian. It should also be noted that the first element, the adjective, remains uninflected. That means that the modifier does not agree with its head. Some examples of this phenomenon can be found in contemporary Macedonian: *тазе весници* – *taze gazeteler* ('new newspapers'), *тазе мајка* – *taze anne* ('young mother'), *тазе пензионер* – *taze emekli* ('recently retired man'). Further examples of this phenomenon are *ќор фишек* ('empty bullet'), *ќор сокак* ('dead end'), *аирлија работа* ('good work'), *баишка економисти* ('special economists'), *ѓул градина* ('rose garden'). However, we have to emphasize here once again that the majority of adjectives borrowed from Turkish has fully been adapted to the morphological system of Macedonian and is declinable. There are also uninflected adjectives borrowed from other languages, e.g. *фер избори* ('fair elections'), *фер конкуренција* ('fair competition'), *фер судење* ('fair trial'), *фер однос* ('fair conduct').

Another possible Turkism is the so called "admirative", a semantic and morphological category of evidentiality. This grammatical category was described for Bulgarian in 1910 by Conev (see Conev 1910: 1ff.). The admirative is often cited as an example of the Turkish influences on the grammatical structure of Macedonian (cf. Conev 1910: 1ff; Sandfeld 1926: 93; Koneski 1986: 173ff.). However, it is also a topic of controversial linguistic discussions (cf. Levin Steinmann 2000: 152ff or Friedman 2010: 131ff.).

3. Data analysis in the light of the theory of borrowing scales

The results of the present empirical study are in complete agreement with our previous publication on this topic, although we have examined a much larger amount of data (see also Kuperdyayev/Hansen: 2013). On the one hand the data analysis of the entire corpus of determined Turkisms confirms that the vast majority (83%) of them fell into disuse at least in the Macedonian newspaper language. On the other, it is also evident that despite their steady decline Turkisms remain numerically one of the largest layers of loanwords in Macedonian. The total number of detected replications of Ottoman Turkish including low-frequency elements is 1,083. Some of these replications are among the high-frequency lexical items which do not have native Slavic equivalents (e.g. *пари*, *пазар*, *боја*, *баи*, *кат*, *тутун*). There are 32 lexical elements in our corpus with frequencies over 1,000 text examples.

With regard to the theory of borrowing scales it is important to mention that Macedonian shows a high intensity level of language contact with the Ottoman Turkish, described by Thomason (2001: 70 – 71) "[...] as slightly more intense or more intense language contact". Not only are there replications of content words, but also of function words. The occurrence of derivational suffixes, which can also be

connected with non-Turkish stems (e.g. *границитилак, пубертетлија, филмаџија*), indicates even a more intense level of language contact. The following Tab. 14 shows Macedonian data (highlighted in bold) found in the corpus on the borrowing scales according to Thomason.

Tab. 14: Macedonian Turkisms on the Thomason’s borrowing scale (Thomason 2001: 70 – 71)

| Language contacts | Lexicon | Phonology | Morphology | Syntax |
|--|--|--|--|--|
| Casual contact (borrowers need not be fluent in the source language, and/or few bilinguals among borrowing-language speakers) | Only content words; most often nouns, but also verbs, adjectives, and adverbs; only nonbasic vocabulary borrowed | | | |
| Slightly more intense contact (borrowers must be reasonably fluent bilinguals, but they are probably a minority among borrowing-language speakers) | Function words (e.g. conjunctions and adverbial particles like ‘then’) as well as content words; still nonbasic vocabulary | New phonemes realized by new phones, but in loanwords only | | New functions or functional restrictions for previously existing syntactic structures, or increased usage of previously rare word orders |
| More intense contact (more bilinguals, attitudes and other social factors favouring borrowing) | More function words borrowed (pronouns and low numerals, derivational affixes); basic vocabulary | Loss of some native phonemes not present in the source language, addition of new phonemes even in native vocabulary, prosodic features such as stress placement, loss or addition of syllable structure constraints and morphophonemic rules | Borrowed inflectional affixes and categories may be added to native words | Word order (e.g. SVO beginning to replace SOV or vice versa) and the syntax of coordination and subordination |
| Intense contact (very extensive bilingualism among borrowing-language speakers, social factors strongly favouring borrowing) | Continuing heavy lexical borrowing in all sections of the lexicon, heavy structural borrowing | Anything goes, including structural borrowing that results in major typological changes in the borrowing language | | |

By comparing Macedonian data with the frequency based borrowing hierarchies (scales) of Matras (2009: 157) we could detect a surprisingly wide coverage of word types on the scale including nouns (*пари, пазар, боја*), conjunctions (*ама*), verbs (*темели, текне, ќари*), discourse markers (*демек, белким*), adjectives (*ќор, тутунски, муслимански, бакарен*), interjections (*ајде, аман, аферим*), adverbs (*баи, бадијала, ептен, темелно*), one preposition (*ќарши*), two pronouns (*ич, филанкишија*) and derivational affixes (*-лак, -лија, -џија*).¹⁰

Tab. 15: Macedonian Turkisms on Matras’ borrowing scale (relevant issues highlighted in bold)

nouns, conjunctions>verbs>discourse markers>adjectives>interjections>adverbs>other particles, adpositions> numerals >pronouns>derivational affixes> inflectional affixes

¹⁰ See also Kuperdyjev/Hansen (2013) for the full discussion of the interpretation of Macedonian data on selected borrowing scales.

4. Closing remarks

Macedonian and Turkish look back on a half a millennium of language contact. The relevance of this contact situation has encouraged extensive research into Turkish elements in the languages of South Eastern Europe. Since the 19th century the focus of this research has largely been on diachronic oriented tradition. In addition to that, research motivated and promoted by politically oriented language planning played an important role for most of the 20th century. Both research directions focused mainly on material borrowings. However, little attention has been paid to the current language usage of Turkish elements.

In the present paper we made an attempt to determine the current status of Turkish replications on the basis of a corpus analysis of the Macedonian newspaper language. For our purpose we created an extensive list of diachronic Turkish replications with 3,155 entries. This comprehensive digitized collection of Turkisms was verified by means of the corpus of two Macedonian newspapers *Дневник* and *Утрински весник*. Present day Macedonian still contains numerous Turkish loanwords, their number and importance is, however, not comparable to their status in the 19th century: 34.3% or 1,083 lexical entries of 3,155 in the original diachronic list have been identified in the newspapers. Only 17.3% of them occur with the frequency of over 10 text examples. Our empirical study thus confirms the quantitative decline of Turkish elements in Macedonian, although the Turkish loanwords represent numerically the largest layer of borrowings in the Macedonian language. Moreover, Macedonian data confirm the postulated borrowing scales (e.g. of Thomason and Matras). Nouns can be easily borrowed and represent 82% of all the borrowed lexical material. However, even contemporary Macedonian contains replications of all other parts of speech except for numerals, albeit with significantly lower frequencies. At the same time Turkish replications cover a wide range of borrowing scales indicating more intense language contact between Macedonian and Ottoman Turkish that is still evident even after a century of restrictions and attempts to marginalize Turkisms in Macedonian.

5. References

- AGO, AKTAN (2006): *Turskite jazični elementi vo jazikot na Marko Cepenkov*. Skopje: Logos-A.
- APOSTOLSKI, MICHAILO (ed.) (1969): *Istorija na makedonskiot narod. Kniga prva. Od predistoriskoto vreme do krajot na 18 vek*. Skopje: Nova Makedonija.
- BLANKENHORN, RENATE (2003): *Pragmatische Spezifika der Kommunikation von Russlanddeutschen in Sibirien. Entlehnung von Diskursmarkern und Modifikatoren sowie Code-switching* (= Berliner Slawistische Arbeiten 20). Frankfurt am Main: Peter Lang.
- BORETZKY, NORBERT (1975/1976): *Der türkische Einfluß auf das Albanische*. Wiesbaden: Harrassowitz.
- CONEV, BENJO STEFANOV (1910): *Opredeleni i neopredeleni formi v bälgarsija ezik*. In: *Godišnik na Sofijskija universitet*, 1–18.
- ČEPREGANOV, TODOR (2008): *Istorija na makedonskiot narod*. Skopje: Institut za nacionalna istorija.

- DIMITROVSKI, TODOR (1966): Za literaturnata leksika na Misirkov. In: Andonov-Poljanski, Hristo et al. (eds.): *Krste Misirkov: Naučen sobir posveten na 40-godišnjata od smrtta. Kniga 6*. Skopje: Institut za makedonski jazik „Krste Misirkov“.
- FIELD, FREDRIC W. (2002): *Linguistic borrowing in bilingual contexts*. Amsterdam – Philadelphia: John Benjamins.
- FRIEDMAN, VICTOR (1995): Persistence and change in Ottoman patterns of code-switching in the republic of Macedonia: nostalgia, duress and language shift in contemporary Southeastern Europe. In: *Code-switching and language contact*. Ljowert/Leeuwarden: Fryske Academy, 58–67.
- FRIEDMAN, VICTOR (2001): *Macedonian*. SEELRC.
- FRIEDMAN, VICTOR (2003): *Turkish in Macedonia and beyond. Studies in contact, typology and other phenomena in the Balkans and the Caucasus*. Wiesbaden: Harrasowitz Verlag.
- FRIEDMAN, VICTOR (2005a): Admirativity: Between modality and evidentiality. In: *Sprachtypologie und Universalienforschung* 58, 1, 26–37.
- FRIEDMAN, VICTOR (2005b): From Orientalism to Democracy and Back Again. Turkish in the Balkans and in Balkan Languages. In: Detrez, Raymond; Plas, Pieter (eds.): *Developing Cultural Identity in the Balkans. Convergence vs. Divergence*. Bruxelles: Lang, 25–43.
- FRIEDMAN, VICTOR (2010): The Age of the Albanian Admirative: A Problem in Historical Semantics. In: Ronald Kim et al. (eds.): *Ex Anatolia Lux: Anatolian and Indo-European Studies in Honor of H. Craig Melchert*. New York: Beech Stave Press, 31–39.
- GRANNES, ALF ET AL. (2002): *A dictionary of Turkisms in Bulgarian*. Oslo: Novus.
- HAZAI, GYÖRGY; KAPPLER, MATTHIAS (1999): Der Einfluss des Türkischen in Südosteuropa. In: Hinrichs, Uwe (ed.): *Handbuch der Südosteuropa-Linguistik*. Wiesbaden: Harrasowitz Verlag, 649–675.
- HINRICHS, UWE; BÜTTNER, UWE (2002): Turkologie und Südosteuropalinguistik. In: Kappler, Matthias (ed.): *Turkish language contacts in South-Eastern Europe*. Istanbul: The Isis Press, 273–288.
- HOLLAND, OLGA (2007): *Anglizismen in der deutsch- und russischsprachigen Presse. Eine kontrastive Untersuchung am Beispiel der “Financial Times Deutschland” und der “Vedomosti”*. Hamburg: Kovač.
- JAŠAR-NASTEVA, OLIVERA (1987): *Turski elementi vo jazikot i stilot na makedonskata narodnja poezija*. Skopje: Makedonska Akademija na naukite i umetnostite.
- JAŠAR-NASTEVA, OLIVERA (2001): *Turskite leksički elementi vo makedonskiot jazik*. Skopje: Institut za makedonski jazik „Krste Misirkov“.
- KAZAZIS, KOSTAS (1972): The Status of Turkisms in the Present-Day Balkan Languages. In: Birnbaum, Henrik; Vryonis, Speros (eds.): *Aspects of the Balkans: Continuity and Change*. The Hague: Mouton, 86–116.
- KEPESKI, KRUME (1972): *Gramatika na makedonskiot literaturni jazik za učilištata od II stepen*. Skopje: Prosvetno delo.
- KONESKI, BLAŽE (1966a): *Gramatika na makedonskiot literaturni jazik. Del I i II*. Skopje: Prosvetno delo.
- KONESKI, BLAŽE (1966b): Eden govor na Grigor Prličev od 1866 godina. In: *Makedonski jazik XVII*, 206–216.
- KONESKI, BLAŽE (1986): *Istorija na makedonskiot jazik*. Skopje: Kultura.

- KONESKI, BLAŽE (1987): *Gramatika na makedonskiot literature jazik*. Skopje: Kultura.
- KONESKI, BLAŽE (1998): Sostojbi i promeni vo makedonskiot jazik od 1945 godina. Zboroobrazuvanje i leksika. In: Minova-Ġurkova, Liljana (ed.): *Makedonski jazik*. Opole: Uniwersytet Opolski, 129–159.
- KRAMER, CHRISTINA (1992): The Use of Turkisms in Balkan Languages as Reflected in Konstantinov's Novel *Baj Ganjo*. In: *Zeitschrift für Balkanologie*, 28, 1–2, 44–60.
- KUPERDYAEV, ALEXANDER; HANSEN, BJÖRN (2013): Neue Perspektiven zu den türkischen Elementen in den südosteuropäischen Sprachen. Aktuelle Turzismen des Makedonischen im Lichte von Entlehnungsskalen. In: Grković-Mejdžor, Jasmina; Loma, Aleksandar (eds.): *Miklosichiana bicentennialia. Zbornik u čast dvestote godišnjice rođenja Franca Miklošiča* (= Posebna izdanja SANU knj. DCLXXV / Odeljenje jezika i književnosti 58). Beograd: SANU, Staroslovenski odbor, 381–406.
- LEVIN-STEINMANN, ANKE (2000): Die Legende vom Renarrativ (am Beispiel des Albanischen, Bulgarischen, Makedonischen und Türkischen). In: Hinrichs, Uwe; Büttner, Uwe (eds.): *Die Südosteuropa-Wissenschaften im neuen Jahrhundert. Akten der Tagung vom 16.-19.10.1999 an der Universität Leipzig*. Wiesbaden 2000, 151–162.
- LEONTIĆ, MARIJA (2012): Blaže Koneski i negoviot odnos kon turcizmite vo sovremeniot makedonski jazik. In: *Lingvistika: Međunaroden seminar za makedonski jazik, literatura i kultura: XXXVIII naučna konferencija*. Skopje: Univerzitet „Sv. Kiril i Metodij“, 39–51.
- LEONTIĆ, MARIJA (2011): Makedonskiot jazičen korpus i mestoto na turskite elementi vo nego. In: *Naučna konferencija „Naukata za makedonskiot jazik – sostojbi i perspektivi“*. Skopje: Makedonskata akademija na naukite i umetnostite. (unpublished paper)
- MAKARIJOSKA, LILJANA (2006): Odnosot kon zaemkite vo makedonskite tekstovi od XIX vek. In: *Lingvistika: Međunaroden seminar za makedonski jazik, literatura i kultura: XXXII naučna konferencija*. Skopje: Univerzitet „Sv. Kiril i Metodij“, 279–286.
- MAKARIJOSKA, LILJANA (2009): Kontinuetot na turskoto vlijanie vrz makedonskiot leksički sistem. In: *Makedonistika: Kniga 10*. Skopje: Institut za makedonski jazik „Krstе Misirkov“, 189–238.
- MATRAS, YARON (2009): *Language contact*. Cambridge: Cambridge University Press.
- MIKLOSICH, FRANZ (1884): *Die türkischen Elemente in den südost- und osteuropäischen Sprachen*. (Griechisch, Albanisch, Rumunisch, Bulgarisch, Serbisch, Kleinrussisch, Grossrussisch, Polnisch). Wien.
- MIKLOSICH, FRANZ VON (1889): *Über die Einwirkungen des Türkischen auf die Grammatik der südosteuropäischen Sprachen*. Wien: Tempsky.
- MORAVCSIK, EDITH A. (1978): Universals of language contact. In: Greenberg, Joseph H. (ed.): *Universals of Human language. Vol. 1*. Stanford: Stanford University Press, 93–122.
- REINKOWSKI, MAURUS (2002): Kulturerbe oder Erblast? Zum Status der Turzismen in den Sprachen Südosteuropas, insbesondere des Bosnischen. In: *Mediterranean language review* 14, 98–112.
- RIEHL, CLAUDIA MARIA (2009): *Sprachkontaktforschung: Eine Einführung*. Tübingen: Narr.
- SAKEL, JEANETTE (2007): Types of loans: Matter and pattern. In: Matras, Yaron; Sakel, Jeanette (eds.): *Grammatical Borrowing in Cross-linguistic Perspective*, 15–31. Berlin: Mouton de Gruyter.
- SANFELD, KRISTIAN (1926): *Balkanfilologien. En oversigt over dens resultater og problemer*. København: Bianco Lunos.
- SKOK, PETAR (1937/38): Prilozi proučavanju turcizama u srpskohrvatskom jeziku. In: *Slavia* XV, 166–190, 336–366, 481–505.

- ŠKALJIĆ, ABDULAH (1966): *Turcizmi u srpskohrvatskom jeziku*. Sarajevo: Svjetlost.
- THOMASON, SARAH GREY; KAUFMAN, TERRENCE (1988): *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, Los Angeles: University of California Press.
- THOMASON, SARAH GREY (2001): *Language contact*. Edinburgh: University Press.
- TROEBST, STEFAN (2005) Vertraute Fremdheit – Das Osmanische Reich in der makedonischen Geschichtskultur. In: *Südosteuropa Mitteilungen* 6, 62–70.
- WEINREICH, URIEL (1953/1979⁹): *Languages in contact*. The Hague: Mouton Publishers.
- WIEMER, BJÖRN; HANSEN, BJÖRN (2012): Assessing the range of contact-induced grammaticalization in Slavonic. In: Wiemer, Björn; Wälchli, Bernhard; Hansen, Björn (eds.): *Grammatical replication and borrowability in language contact*. Berlin: Mouton de Gruyter, 67–155.
- WHITNEY, WILLIAM DWIGHT (1881): On mixture in language. In: *Transactions of the American Philosophical Association* 12, 1–26.

Barbara Sonnenhauser (Zürich)

Borrowing in context: a pragmatic perspective on Turkisms in pre-standardised Balkan Slavic*

Abstract

Der vorliegende Beitrag beschäftigt sich mit Turzismen im vorstandardisierten Balkanslavischen. Textdokumente der Zeit zwischen Ende 17. bis Mitte des 19. Jh., die zum einen zunehmend volkssprachlich, zum anderen noch kaum puristischen und normativen Tendenzen unterworfen sind, erlauben einen vergleichsweise unmittelbaren Blick auf die Verwendung von Turzismen. Dies wiederum ermöglicht Rückschlüsse auf pragmatische Aspekte von Entlehnung und somit auf die situative und kommunikative Einbettung des literarischen Schaffens dieser Periode. Vor diesem Hintergrund wird die Rolle von Turzismen nicht nur in der sprachlichen, sondern auch in der literarischen Entwicklung der späteren ost-südslavischen Standardsprachen gezeigt.

1. Introduction

Borrowings indicate the nature of contact between donor language and recipient language, such as the hierarchical relationship between both languages, the interactional spheres of their speakers and the intensity of contact. The latter is associated with the linguistic levels at which borrowings occur, with contact-induced changes on the morphosyntactic level presupposing longer lasting and more intensive contact than influences on the lexical level. While this aspect is among the major topics in research on borrowing, the present paper focuses on yet another facet, namely the pragmatic side of borrowings. It will be illustrated how far different usage types of borrowings in written documents allow drawing conclusions on the situational embedding of literary production. Based on this kind of evidence, more information can be gained about the communicative setting of a pre-standardised language for which first-hand observation is no more possible. This will be shown here on the example of Turkisms in Balkan Slavic text documents dating to the 17th to 19th century. Not yet being subject to prescriptive tendencies, the usage of Turkisms in these documents promises – at least indirect – insight into the linguistic situation of those times with respect to their entrenchment, the authors' linguistic awareness and the language background shared by the authors/compilers and their intended audience.

Importantly, this study neither aims at representativeness nor statistic balancing for the texts and examples chosen. It rather aims to provide first insight into pragmatic aspects of borrowing in older stages of a language, which may serve as a basis for follow-up investigations on possible correlations between, e.g., genre and usage of Turkisms, individual preferences of authors, semantic changes of Turkish elements, etc.

* The research for this paper has been supported by the Austrian Science Fund FWF (project 'The emergence of narrativity in Early Neo-Balkan Slavic', grant number M1536 G23). This support is gratefully appreciated. I would also like to thank the editors and reviewers for their careful comments on an earlier version of this paper.

2. Background

Turkisms result from the intensive contact between the Balkan people and the Turkish speaking population in the administrative and political contexts of the Ottoman Empire. Especially in the cities, Turkish was in wide use, both as a first and second (or third) language, serving as ‘interethnic means of communication’ (Kappler 2002: 821). This intensive contact is visible in the types of borrowings which include not only content words, but also function words, derivational morphemes and pragmatic markers.¹

In describing the history of the South and Balkan Slavic literary languages, the assumption that “[t]he handling of Turkish loanwords is a fundamental issue” (Henninger 1990: 14) is widespread. It predominantly refers to the expulsion of Turkisms and their replacement by ‘genuine’ Slavic elements starting from the mid 19th century. However, it is not only with respect to these purist tendencies, that Turkisms constitute an important part of the literary development in Balkan Slavic. It is also the very usage of Turkisms in written documents itself. Supporting the comprehensibility of these works, they are an indispensable part of the vernacularisation process of the written language and, consequently, its functional elaboration. Turkisms also contribute to the functional-stylistic differentiation of Balkan Slavic, in that their usage supports the emergence of new, increasingly mundane genres. The contribution of Turkisms to the elaboration of the later literary languages is also an indirect one, since it was these elements that made the purists feel the need for elaborating their language to enhance it for all functional styles.

The observation that writers between the 17th and 19th century incorporated Turkish elements or even had to rely on them in order to be understood by their audience is the starting point for the present paper. It will be shown that the usage of Turkish elements is not identical in any case, which is easily ignored if the investigation focuses on listing Turkish lexemes and compiling an inventory of Turkisms. The differences observed in the specific way in which they are used provide further insight as regards their function and purpose.

The focus will be on the usage of Turkisms in documents dating from the 17th to the late 19th century. The motivation for this temporal delimitation is twofold: The 17th century marks the beginning of an increasing vernacularisation, while from the mid-19th century onwards, puristic tendencies emerge that are targeted to a large part on the expulsion of Turkisms from written documents. The time span under investigation is characterised not so much by conscious language planning with the aim of establishing a literary language, as by the striving for comprehensibility.

The prevalent view in the literature assumes Turkisms in texts dating to the end of the 18th/beginning of the 19th century to be used in a chaotic way, presuming that the literary knowledge of

¹ For an overview of the history of borrowings from Turkish see, e.g., Mirčev (1963: 73–84) and Stamenov (2014), for Turkish elements in Bulgarian derivational morphology Selimski (2012).

the writers was only low (e.g. Petkanova-Toteva 1965: 223). However, Turkish borrowings may – at least to some extent – indeed have been used in an intentional way. Given that they were part of the linguistic variety used for specific genres, and given that they were also used as explanations to expressions the compilers and authors thought to be unknown to their intended audience, the writers of that time employed Turkisms in a conscious and deliberate way. This in turn suggests a high degree of reflection on language, texts and interlocutors – which is hardly imaginable for unskilled writers. Judging older stages of linguistic and literary development from the standpoint of the contemporary prescriptive standard and contemporary literary conventions thus does not do justice to a time when a standard or any kind of inter-individual norm had not been developed yet.

3. Turkisms in Balkan Slavic

Lexical Turkisms, which are the main focus of the present paper, can be defined as “loanword[s] that [have] entered Bulgarian through the medium of the Turkish of the Ottoman Empire and latter-day Turkey”,² including “words that are themselves loanwords in Turkish” (Grannes et al. 2002: vii), such as Persian or Arabic elements which are widespread in Ottoman, the official language of the state.³

The influence of Ottoman Turkish on the Balkan and South Slavic languages has received considerable attention in the literature. Earlier works, starting with Miklosich (1884), focused on lexical elements, thus presenting lists of words with the aim to provide an inventory of Turkisms. More recent analyses consider them in larger contexts, i.e. within that of language planning and linguistic ideologies (Friedman 1997) or as examples of the mechanisms and regularities involved in language contact (Kuperdyaev/Hansen 2013). From a Balkan linguistic point of view the question is discussed whether there is a common lexicon of Turkisms that characterises the Balkan languages. Furthermore, Turkish lexical elements are examined with respect to possible stylistic effects, predominantly in view of their more recent usage in the mass media (e.g. Mevsim/Čakárova 2007).

3.1. Simple language

The increasing appearance of Turkisms in literary texts from the late 16th century onwards is embedded in the larger context of vernacularisation and the changing function of literary production. Given the political circumstances, religion was an important factor supporting the preservation of a common identity within the Ottoman administrative structures, and literature the main means to distribute faith among the people (see Ilievski 2005 for the relation between literary tradition and political circumstances). This meant a change in the communicative setting for literary production, which in turn had a bearing on the

² See Stachowski (2014: 1199–1200) for a short list of earlier Turkic influences on Balkan Slavic. Hinrichs (2004) argues for a more intensive Turkish influence already before the Ottoman Empire.

³ See Fijuzi (2004) for Persian elements that entered Bulgarian via Ottoman Turkish.

composition of texts, most prominently concerning the replacement of Church Slavonic elements, structural simplification and reference to the living situation of the addressees (for more details cf. Sonnenhauser forthcoming). Large parts of the intended audience being illiterate, the texts also had to be designed to be read out (Punčo emphasises that his *sbornik* is to be printed such that priests can read it out to the people; see Punčo 1v–2), which contributed to their more oral character. The increasing simplification, such as in (2), where the participles *sědeštu* and *stoještixъ* appearing in (1) have been replaced by finite constructions with *sědi* and *stoęťь*, is thus less an index of the decreasing literacy of the scribes and authors, but primarily testifies of an intentional revision and elaboration of the original texts.

- (1) Къ утръны же съномъ объеть бывъ, мнѣше се caricu někoju na přěsvětlěmъ *sědeštu* přěstolě zřěti, i množьstvo mnogo světlыхъ воинъ окръсть тоје *stoještixъ* (Novaković, 56; 16th century)
 ‘But towards morning, being embraced by sleep, it seemed to him that he saw a tsarina *sitting* on a very bright throne and a crowd of bright worriers *standing* around her.’
- (2) I na съмнуване заспа. i vidě edna žena kato edna carica, i *sědi* na stol zlatъ světlіkavъ. I okolo neja *stoęťь* mnogo voine xubavi světlіkavy. (Tixonravovski, 56v, 96; 17th century)
 ‘And at the break of dawn he fell asleep. And he saw a woman like a tsarina, and she *is sitting* on a golden, bright chair and a lot of worriers, fine and brave, *are standing* around her.’

That the usage of these simpler structures and elements is not just a slip of the pen or an accidental immersion of the vernacular into written documents is evinced by the introductory words to the respective texts. In many cases these parts announce the usage of a ‘simple’ language in order to be understood by the ‘ordinary people’, as at the beginning of the Nedelnik in (3):

- (3) Prepisannyj ot Sloveskago i ot Grečeskag Glubočajšago ęzyka na Bolgarskij prostyj ęzykъ [...] krazuměnjju prostomu narodu. (Nedelnik)
 ‘Transcribed from the Slavic [= Church Slavonic, BS] and the most profound Greek language into the simple Bulgarian language to be understood by the ordinary people.’

In the introduction to PE, explicit reference is made to the fact that most people are illiterate, which is why they do not know how to become heirs of the Kingdom of Heaven. To change this situation, this book is written in simple Bulgarian:

- (4) Obače въ dnešnoto vreme [...] Bulgarskij narodъ mnogo čelověcy ima neučeni, i otъ svętoto pisanie ne razuměvatъ, [...]: a koga go ne znaeme i ne razuměvame kakъ šte bōdime naslědnici na carstvoto nebesnoe [...] Za radi tova sę potrudixъ i přěpisaxъ taę svęštenia kniga na prosto-Vъlgarskij jazykъ, za da by razumělъ sěkij Vъlgarinъ štoto sę izkasva otъ neę, i sěkij da sę nastavi na istinnuj i pravuj pōť. (PE: IV–V)
 ‘But in this day and age, there are many Bulgarians that are illiterate and do not understand anything about the Holy Scripture; [...] and when we do not know and understand it, how

will we become heirs of the Kingdom of Heaven? [...]. Therefore I made the effort and transcribed this holy book into simple Bulgarian, such that every Bulgarian will understand what is said by it and choose the true and right way.’

Pop Punčo emphasises at the beginning of his *sbornik* that he translated from Greek such that also the lesser educated may read and understand:

- (5) [...] произведохъ i o/t/ грѣѣ/s/ki i izvodeъ paki na prostago jazika prepisaxъ i domislixъ азъ okajani i mnogo rešni eren punčo da pročetuť nerazumejušti da razumejuť się prosta slovesa [...] (Punčo: 1v)
‘[...] I translated from the Greek source into the simple language, I rewrote and thought, me the very poor and sinful monk Punčo, such that those, that do not understand, will understand these simple words [...].’

The need for a simple language is pointed out also in the early textbooks, such as Petăr Beron’s reading primer, which is meant to include entertaining instead of religious texts. Popularisation is thus achieved not only by using a simple language, but also by choosing more accessible topics.

- (6) Sobraxъ i několko basy i različny poučenię [...] i gledaxъ da gi napiša kolkoto moga prosto, štoto vseki da gi razibra a koito rči mi se viděxa man mačny gudixъ gi prigradeni i turski. Vsekiј kojto vidi тази книжка naděę se da se zaraduva, a naj mnogo učiteli-tě, zašto tii, mi se struva, [...] štexa do [sic] ostaveť psaltiri-tě i časoslovi-tě, ot koito děca-ta ne razbiratъ ništo. (RB 3)
‘I have chosen some fables and various advices [...] and tried to write them as simple as possible, such that everyone will understand them, and those words that seemed difficult to me, I have added in Turkish. Everyone who sees this book will hopefully rejoice, most of all the teachers, because they, it seems to me, [...] wanted to put Psalters and books of hours, of which the children do not understand anything.’

It is this background of the quest after intelligibility against which the first uses of Turkisms in written documents have to be regarded.

3.2. Linguistic levels

Turkish influence can be observed on all linguistic levels. Among the most cited examples of contact influence in the verbal system (e.g. Neweklowsky 2012) is the so called ‘renarrative’ category, which is assumed to have developed under Turkish influence. Turkish has a specific verbal category used to express events that have not been witnessed personally. Gołąb (1960) regards the Bulgarian *l*-periphrasis with the renarrative interpretation as a structural copy of the Turkish model. This copy was possible “only because of the structural identity of the corresponding constructions” (Gołąb 1960: 38). An important component was the keeping and the new interpretation of the aorist vs. perfect opposition and the

morphological structure of the Turkish form which, according to Gołąb's analysis, has its parallel in the Slavic construction of *l*-participle and auxiliary, as illustrated in (7) for Macedonian:⁴

- (7) a. öğrendim (tr.) učiv (mac.)
 learn.pst.1sg learn.aorist.1sg
 öğrenmişim (tr.) učil sum (mac.)
 learn.pst.1sg learn.perfect.1sg
- b. öğren-miş-im (tr.)
 learn-*miş*.participle-1sg/copula
 učil-l sum (mac.)
 learn-*l*.participle 1sg/copula

Fielder (1999) shows how both categories are used for comparable discourse purposes, regarding these parallels as an instance of 'conceptual convergence' between Turkish and Balkan Slavic. As example for grammatical calques on the nominal level, Gołąb (1960: 41) cites possessive constructions of the type *na car'u majkamu* 'the mother of the Tsar'. In this construction, the Balkan Slavic possessive clitic *mu* is used as flecional morpheme corresponding to the Turkish possessive marking *-in* (cf. *o fukaranın kadın-i* 'this poor man's wife').

Particularly prone to contact influences is the lexical level, where Turkish influence can be traced in derivational morphology (for an overview see, e.g., Stachowski 2014: 1206-1207), and structural calquing as in (8): *često često* corresponds to the Turkish reduplication *sık sık* 'frequently; lit.: often often', which indicates intensity:

- (8) i Christosъ dzelo mnogo *često često* dumaľ [tr. *sık sık* 'frequently'] (RP 252)
 'and Christ said very much *often often*'

The most obvious manifestation of Turkish influence are lexical borrowings. Among the first to establish a list of Turkish lexical elements in Balkan Slavic was Miklosich (1884). His glossary captures the historical-linguistic situation of that time as evinced in the works he has searched. Taking Miklosich as a starting point, Kuperdyjev/Hansen (2013) provide a survey of the situation in contemporary Standard Macedonian (see also Kuperdyjev this volume).

In the pre-standardised text under investigation here, the picture is similar. Content words belong to abstract vocabulary as in (9) and to every-day vocabulary, as in (10). The example of *kefove*, (9c), is

⁴ That is, Gołąb (1959) relates the retention of the synthetic aorist in Balkan Slavic, as opposed to the other Slavic languages (with the exception of Serbian), to the influence of Turkish. The importance of this retention for the functional development of the former perfect forms in Balkan Slavic is pointed out also in Meermann/Sonnenhauser (forthcoming).

typical of morphological integration, with the suffix *-ove* as the Balkan Slavic plural morpheme for one-syllable nouns.⁵

- (9) a. seki imasi ot Boga *izmetъ* [tr. *hizmet* ‘position, function’] (RP 2)
 ‘everyone is assigned his *position* by God’
 b. u toj časъ imъ došelъ *gazerъ* ot Boga [tr. *gazap* ‘rage’] (ČB 5)
 ‘at this time they were struck by divine *rage*’
 c. kefove [tr. *keyif* ‘joy, pleasure’] (RP 144)
 ‘[kinds of] amusement’
- (10) a. maxana ‘tavern’ [tr. *meyhane* ‘pub’] (Svištovski)
 b. deva ‘remedy’ [tr. *deva* ‘medicine, remedy’] (Svištovski)
 c. mangъr ‘copper coin’ [tr. *Mangır* ‘money, copper coin’] (Svištovski)
 d. ališъ verišъ ‘shopping, trade’ [tr. *alışveriş* ‘trade, shopping’] (KO 104)

Examples of function words are given in (11). They, too, can be incorporated into the morphological system: *nesaltъ*, (11c) has the Slavic negation affix *ne-*.

- (11) a. Pak Iisus *činki* pozna mislite xmi, reče [tr. *çünkü* ‘because’] (TE 106)⁶
 ‘and Jesus, *because* he knew their thoughts, said’
 b. ias karštavam fav oda, *andžek* meždu vas [...] [tr. *ancak* ‘only, solely; but’] (TE 96)
 ‘I baptise in the room, *solely* among you’
 c. nesaltъ [tr. *salt* ‘solely, simply’] (RP 77)
 ‘not only’

In (12a) *kerečeš’* is transformed from a nominal element (tr. *gerek* ‘necessity; [it is] necessary’) in Turkish into a verbal finite form in Balkan Slavic by the second person singular ending *-eš*; in (12b) *kerekъ* is used as modal construction alongside with the Slavic modal construction *trebe*:

- (12) a. kerečeš’ (KO 65)
 ‘you have to’
 b. Каѡъ ќе дѣте да заучи книга [...] *trebe* Попъ да му čati molitvy na glava [...] *kerekъ* seme koga ќе вакаѡъ на нива, *kerekъ* gumno koga ќе nasadišъ (KO 36)
 ‘When as a child is to learn a book, a priest *is necessary* in order to say him a prayer [...], seed *is necessary* when it is time for the field, a granary *is necessary* when you grow something’

Lexical borrowings are commonly grouped into ‘core’ and ‘cultural’ (see, e.g., Haspelmath 2009; Zenner/Speelman/Geeraerts 2014 offer a critical view on this distinction). This distinction concerns different contact situations and can hence be assumed to point to different groups of language users and result from different ways of mediation.

⁵ Underlining marks Slavic morphology.

⁶ This is an example of meaning change, since in contemporary Bulgarian, its meaning is ‘nali, nima, kato če li’ i.e. ‘isn’t it, really, indeed’. A systematic investigation of such meaning changes would be desirable, as would be an overview of the indexing of Turkisms in contemporary dictionaries as ‘dialectal’ or ‘archaic’.

Core borrowings „duplicate meanings for which a native word already exists“ (Haspelmath 2009: 46). The question why elements are borrowed even though native words are available is usually answered with reference to the prestige of the donor language the speakers want to be associated with (Haspelmath 2009: 48). During the time span under consideration here, Ottoman Turkish was not only “a language with which a good number of non-Turks were more or less familiar” (Kazazis 1972: 91), but also had a considerable degree of prestige as the “language of the state and the town” (Friedman 2006: 670). Both facts manifest themselves in the large number of Slavic-Turkish synonymous variants. An impression of such lexical doublets can be received from Neofit Rilski’s grammar (Gramatika: 5–13), where many of them are listed, as in the excerpt in (13), which gives different variants for ‘wall’:

- (13) Na drugo mĕsto paki govoratъ *dzidъ* [...], na drugo *stĕna*, po Slavĕnski, a na drugo *duvarъ* [tr. *duvar*] (Gramatika: 6)
 ‘At another place people say *dzidъ*, at another *stĕna*, in Slavic, and at another *duvarъ*’

However, doublets are given not only for Slavic-Turkish, but also for different Slavic variants, see (14):

- (14) i v glagolyte podobnĕ [...] *govorimъ, xoratimъ, xortuvamъ, sboruvamъ*, [...], *besedimъ, sborimъ, psuvamъ* i *poporžamъ* [...] (Gramatika: 8)
 ‘and it is similar with verbs: *govorimъ, xoratimъ, xortuvamъ, sboruvamъ*, [...], *besedimъ, sborimъ, psuvamъ* and *poporžamъ* [all meaning ‘say, tell, speak’]’

Obviously being treated on a par with the general variety indicates the deep entrenchment of Turkisms in the Balkan Slavic lexicon at that time. Neofit Rilski goes on to point out that this variability is not a wealth of language, but harmful (Gramatika: 10). This harm, apparently, does not refer to Turkisms alone, but to the general variability which he considers an obstacle in constructing an overarching norm and standard.

The second main type of lexical borrowings is used to name new concepts adopted from the culture of the donor language. These concepts “are the more likely to be expressed by loanwords, the more widely the donor language is known” (Haspelmath 2009: 48) – and, one may add, the more the respective concepts are relevant for the recipient language and its speakers. In the present case these are predominantly ‘administrative’ Turkisms (see also Kazazis 1972: 92) related to the structures of the Ottoman Empire. Most of these borrowings are found in newspapers that have emerged in the early 19th century onwards (e.g. *Ljuboslovie, Edirne, Carigradski Vestnik*). The examples in (15) show that such Turkisms are used in an ‘integrated’ way or as an explanation in brackets (see section 4) and may also be adapted to the morphological system of the Slavic recipient language. In (15a) it is the derivational morpheme for adjectives *-sk-* amended with the flexional suffix, in (15b) it is not only the derivational

morpheme *-stv-* with the flexional suffix, but also a semantic calquing with the usage of *veliko* ‘Grand’, and in (15c) it is the plural morpheme *-i* with the postponed definite article.

- (15) a. *vilaetsky* [tr. *vilayet* ‘vilayet, province’] (E 1869 24)
 ‘concerning the Vilayet’
 b. *veliko vezirstvo* [tr. *vezir-i azam* ‘supreme vezier’] (CV 1850/86, 223)
 ‘Grad Vizierate’
 c. *Diplomate-te čuždi (elčii-te) bexa došli* [tr. *elçi* ‘ambassador, envoy’] (LS 1844/5,75)
 ‘foreign diplomats had come’

Turkisms like these have most probably not been in everyday usage but constituted part of a special language. However, terminology is borrowed not only from Turkish, but also from other languages such as Greek. In (16) Graecisms serve as *explanans* for new Slavic coinages, (16a), or as *explanandum* with the Slavic explanations given in parentheses (16b):

- (16) a. *bytiopisanieto (istoriēta); zemlemērieto (Geometriēta)* (KN 1)
 ‘descripition of life (*history*); measuring of the earth (*geometry*)’
 b. *Thermometrъ (mēra za goreštinata) značeše 8 stepeni podъ zero* (LS 1844/5, 75)
 ‘the *thermometer* (measure for heat) shows 8 degree below zero’

Interestingly, more ‘true-to-life’ sciences such as medicine and administration, i.e. domains which inevitably affected the personal spheres of people, are explained in terms of Turkisms. This indicates the different domains of contact with different contact languages:

- (17) a. *za da stani soveršenъ ili filosofъ (vъ Filosofičeskata nauka) ili estestvoslovъ (vъ Fizikata) ili Doktorъ (vъ xekimadžiiskata nauka)* [tr. *hekim* ‘physician, doctor of medicine’] (KN 11)
 ‘in order to become a perfect philosopher (*in the philosophical science*) or natural scientist (*in physics*) or doctor (*in the medical science*)’
 b. *naguba mnogo něčta vъ smětvy-ty si (xesapy-ty)* [tr. *hesap* ‘calculation, account, bill’] (CV 1850/85, 222)
 ‘he lost many things in his *calculations*’

While the administrative Turkisms in (15) can be classified as specific terminology, others are more part of everyday life, even though they are borrowed from an administrative context as well. Examples are given in (18), all taken from the Svištovski damaskin:

- (18) a. *saraj* ‘palace’ b. *bej* ‘bey’
 c. *kadija* ‘qadi’ d. *spaxija* ‘sipahi’

As Reinkowski (2002: 101) points out, Turkisms entered into Balkan Slavic in different ways, mainly via the language of the state and government as mediated by the ruling elite in the urban centres and via the

spoken language of the Turkish people. That is, not only different borrowing domains (basically core and cultural) have to be distinguished, but also different functional domains.

Gołąb (1960: 30) points out that the Turkisms can be found also in religious vocabulary, which is usually characterised by Old Church Slavonic. Some examples are given in (19):

- (19) a. kurbanъ ‘sacrifice’ [tr. *kurban*] (KO 71)
 b. šaitlikъ ‘being a witness’ [tr. *şahitlik*] (KO 70)

An interesting case – not only lexically, but also syntactically – is *kurtulisam* ‘save’ in (20). As borrowing, it is used transitively, even though it contains the Turkish subject-demoting (reflexive/passive) suffix *-il-* (*kurtulmak* ‘to be saved’, ‘save oneself’ vs. *kurtarmak* ‘to save s.o.’), see (20a). In (20b) it is passivised by adding the Slavic subject-demoting particle *se*:

- (20) a. *ni kurtulisъ otvěčna muka* [tr. *kurtulmak* ‘be saved’] (RP 248)
 ‘he *saves us* from eternal torment’
 b. *taq duša ne može dase kurtulisa* ot strašni gjumrucy diavolski [tr. *kurtulmak* ‘be saved’] (SU 25)
 ‘this soul cannot *be saved* from the horrible tolls of the devil’

For borrowings into the domain of religion neither the need for denominations for new concepts, nor the prestige of the donor language seems to be a plausible explanation. The more obvious motivation can be sought in ‘comprehensibility’.

3.3. Development

The usage of Turkisms in written documents changes over time, the decisive poles being the vernacularisation starting by the end of the 16th century and the onset of puristic tendencies by the mid 19th century. Turkish borrowings were most numerous at the beginning of the 19th century, when they intruded also into the basic vocabulary. The situation changed drastically in the course of the next one hundred years, beginning with the Bulgarian independence of the Ottoman Empire in 1878. Afterwards, “[...] nito edin nov turcizām ne navliza v bālgarskija ezik [...]” (“not one single Turkism entered the Bulgarian language”) (Stamenov 2014: 126), which obviously relates to the reassessment and reinterpretation of the Ottoman heritage. This development is accompanied – and partly triggered – by changes in ideological judgments concerning Turkisms.

Turkisms are in regular usage in written texts from the earliest *damaskini* onwards, with their number increasing in the younger *damaskini* (cf. Miletičъ 1908: XIV and Miletičъ 1923: 69–71 for an exemplary comparison) and texts adhering to that tradition. In these early documents, they appear as integral part of the language used. With language becoming a means of defining national identity, the

attitude of the scribes gradually changes from the mid 19th century onwards. The reassessment of Turkisms in the course of the standardisation efforts is visible in explicit statements included in the respective works. In his grammar, Neofit Rilski refers to the Romance speaking people who enhanced their language within a very short time using lots of foreign words – which he says is not necessary for Bulgarian (Gramatika: 13). And even though some Turkish words might not be replaceable by native words, this is not an obstacle for the improvement of the language – as can be seen from the comparison with Serbian, which still has some Turkish words (Gramatika: 14). These statements do not explicitly condemn Turkisms but exhibit a primarily pragmatic view on language usage. At the end of his grammar, Neofit Rilski provides a list of lexemes as a dictionary giving Slavic equivalents to – mainly – Turkish and Greek lexemes. The underlying intention is to familiarise readers with the Slavic words:

Rěči Turski i několko Grečeski, koi to sa vo upotreblenie na segašno to vreme vь sička ta Bulgaria, istolkuvany po vozmožnosti sã ravnoznamenovatelny Slavenksi, ili Russijski. (Gramatika: 204)

‘Turkish words and some Greek, which are currently in use all over Bulgaria, explained as far as possible by equivalent Slavic or Russian words.’

The pedagogic – virtually prescriptive – intention underlying such a list of words is visible in Kipilovski’s *Kratkoe načertanie na vseobštata istorië* (“Short outline of world history”) (KN, 1836), which is a translation from Russian and is meant not only to provide a history book, but also to elaborate Bulgarian:

ęzykъ upotrebixъ takavzi kakvoto da soči kamъ obrabotěni na ęzykatъ ny, zato i griexъ mnogo rěči, otъ prebogata Majka na našatъ ęzykъ Slavenskiatъ ęzykъ, kojto vь krajnatъ na Istorięta vь Leksikonъ raspoložixъ i gi otkryxъ sasъ prostoljudny i turski rěči, za da ta ulěsnę da možišъ soveršenno da razuměišъ sičko štoto čitešъ, [...] (KN 15–16)

‘I have used such language which indicates the elaborateness of our language, and I used many words from the rich Mother of our language, the Slavonic language, which I have arranged at the end of the *Istoriya* [title of the book, BS] in a lexicon and elucidated them with words known to the ordinary people and Turkish words, such that it is easier for you to completely understand everything what you read, [...]’

Kipilovski’s intention to contribute to the enlargement of the Bulgarian lexicon becomes even more evident in the dictionary he provides at the end of his book, where the puristic intention becomes visible:

Rěčnikъ ili leksikonъ na Slavenski rěči koito treba da sę prisvoętъ vь našatъ ęzykъ, a gi otkryvatъ i sasъ turski za da ulesnę pročitateľatъ da pročita polnorazumno, štoto čite, i da gi zapomni ta da gi upotrebleva i vь govordatъ si, za da sę očisti ęzykatъ ny otъ čuždy rěči. (KN 187)

‘Dictionary or lexicon of the Slavic words which have to be integrated into our language, and I elucidate them with Turkish such that it is easier for the reader to read with full comprehension what he reads, and such that he understands and is able to use them when he speaks, such that our language is cleaned of foreign words.’

These examples are representative of the tendency that in the course of the late 19th century, “Turkish loanwords were discarded in various ways” (Henninger 1990: 3). Giving a Church Slavonic word and putting the Turkish borrowing into brackets afterwards (see section 4) is interpreted by Henninger (1990: 3) as one “means of indicating that the loanword was needed in order to explain the meaning of the new Slavonic word or usage, but also that it was from now on meant to be put aside“. Friedman (1994) makes the same observation, albeit without the prescriptive interpretation. He simply notes that in the 19th century writers “had to gloss many of their Slavic words with Turkish equivalents, which were more familiar to ordinary people” whereas “modern-day collections of folk tales [...] must have glossaries of Turkisms” (1994: 528). That is, the situation at the end of the 20th century is opposite to that at the end of the 19th century.

A turning point towards the opposite direction was the year 1989, which marked changes in the language of the media, including an again increasing usage of Turkisms. Friedman (2003: 62) regards this as “an attempt to ‘democratise’ via colloquialisations”, whereby “the use of Turkisms in formal contexts becomes a marker of ‘democracy’” (ibid.). Krāsteva (2003: 8) takes this development as evidence for the fact that journalism had freed itself of the “kazionnata svrāxnornativnost na blizkoto minalo, kojato ne dopuskaše upotreba na turcizmi vāv vestnicite, v radio-televizionnite predavanija” (“the official over-normativity of the close past, which did not permit the usage of Turkisms in newspapers, radio and television broadcasts”) (2003: 7). Reinkowski (2002: 111) observes that Turkisms appear in genres of ‘subjective political character’ but lack in economic coverage. Mainly because of their high degree of ‘expressivity’ and their usage to convey specific ‘emotional effects’, Turkisms are regarded as ‘anti-aesthetic’ in present day normative stylistics (Krāsteva 2003: 8f.). However, as Krāsteva (ibid.) rightly points out, this misses the stylistic effects associated along with their uses. Both tendencies – colloquialisation/democratisation and genre-sensitivity – have their parallels in the vernacularisation process setting in with the 17th century (this can also be observed on other linguistic levels, cf. Sonnenhauser 2014, Sonnenhauser forthcoming).

These trends show that not only the frequency and distribution of Turkisms altered, but also their assessment. What can be observed is a change in the ideologies, i.e. in the “sets of beliefs about language articulated by users as a rationalisation or justification of perceived language structure and use“ (Silverstein 1979: 193). Friedman (1997) elaborates nine ‘ideological equations’ which he assumes as “characterizing a number of historical developments in the Balkans” (1997: 24). Such equations can be established for the usage and handling of Turkisms as well.

For the scribes and authors between the 17th and early 19th century, the usage of Turkisms contributes to the comprehensibility of their texts, which was in turn regarded desirable in order to establish a common identity by the distribution of religious ideas. There had not yet been the tight relation

between linguistic and notional/political questions that underlies the literacy of the Bulgarian renaissance. Against this background, their ideologies can be formulated as ‘vernacular language = common language = identity’ and ‘common language = comprehensibility = good’. The writers of the renaissance regarded the vernacular in the opposite way, namely as contaminating language. Their intention was to lift the language to a higher level of elaboration, for which purpose it had to emancipate itself from foreign influences. This was considered important not only on the linguistic, but also on the political level. The means to achieve the goal of creating a shared identity thus changed from common contents (religious topics communicated by the vernacular) to a common language with as little ‘foreign’ elements as possible. Neofit Rilski’s thoughts on the difference between the spoken and the written languages are a characteristic example of this point of view (Gramatika: 13–14). The underlying ideology can be paraphrased as ‘vernacular = non-native = non-identity’ and ‘non-native = chaos = disunity’.

Such ‘ideological underpinnings’ (Friedman 1997) can also be observed in linguistic descriptions of Turkisms. Most analyses of the pre-renaissance period focus on the alleged poor elaboration of the literary language which necessitated the employment of non-Slavic lexical elements, the usage of which is typically regarded as ‘chaotic’ (cf. above). According to Petkanova-Toteva (1965: 223), the writers of the late 18th/early 19th century did not think about the language in their anthologies, in particular about the usage of Turkisms, because they did not have a higher literary education. Analyses like these follow the ideology ‘contact = impure = bad = illegitimate’ (Friedman 1997: 28–29).

Actually, it is unclear, what ‘chaotic’ means; using specific linguistic means in order to be understood hardly seems to be chaotic or unintentional. As has been shown, the writers state their intention to write in a simple language in the introductory parts to their works. Intending to use a simple language, they knew that there was also a ‘complex’ language and they must have understood that complex language, otherwise they could not have used these texts as models for their translations and editions. This in turn contradicts their presumed illiteracy. Moreover, such ‘bashing the scribe’ (Fleischman 2000: 47) misses regularities concerning the types of usage. As will be shown in the following section, these different types of usage allow drawing conclusions on the communicative embedding of the texts – which in turn contributes to accounting for tendencies and peculiarities of literary development.

4. Usage types

The listing of lexemes and their origin excludes the usage context (such as author, text, recipient) from the analysis. If the focus of investigation is broadened, differences in usage can be observed which allow for further insight into the changing communicative settings prevailing in the transitional period between the 17th and 19th centuries. In addition, they provide insight into the different motivations for the usage of

Turkisms, the linguistic awareness of the authors and their consciousness as regards the appropriateness of Turkisms in different genres.

4.1. Turkisms as indices

While different linguistic types of Turkisms indicate the domains and intensity of contact, different types of usage permit to draw conclusions on the linguistic and literary development, including the changing roles of the interlocutors and their intentions. In this sense, they can be considered as indices referring to the communicative situation.

As suggested by the discussion in section 3.2, different motivations underlie the usage of Turkisms: different needs (as evinced by core, cultural and terminological borrowings), different intentions (such as intelligibility, prescription, rhetoric purposes) and different attitudes (mainly pragmatic or puristic). These motivations, in turn, are connected to two basic distinctions in usage: non-intentional and intentional. Whereas non-intentional uses point towards the degree of entrenchment of Turkisms, i.e. the degree to which they were regarded as native – not foreign – elements, intentional uses indicate the degree of linguistic awareness on part of the language users. Whether a Turkism is used intentionally or non-intentionally can be deduced from two kinds of evidence: from possible commentaries by the writers themselves, and from the way the respective elements are used.

The two main types of usage can be observed: integrated/object-linguistic usage (corresponding mainly to unintentional uses), and non-integrated/meta-linguistic usage (corresponding to intentional uses). ‘Meta-linguistic’ Turkisms are part of a commenting or explanatory structure, and therein serve either as *explanans* or as *explanandum*. Such explanatory structures deliver additional information. The additional character of the information is indicated by means of punctuation, mainly brackets, or by paraphrasing elements such as *sireč* or *sanki* ‘that is’.⁷ The schemas in (21) sketch these different types of explanatory structures consisting of a Turkism (T) and a ‘native’ element (n) as encountered in the documents investigated here:

- (21) a. integratedT.....
 b. non-integratedn(T).....T(n).....
n *sireč/sanki* T....T *sireč/sanki* n....

A third type of usage can be evinced from lists of words given in the appendices to some works. They indicate a growing pedagogical, puristic and normative attitude and the striving for the elaboration of the vocabulary.

⁷ In this sense, such explanatory structures constitute ‘functional performance structures’ (see Noel Aziz Hanna/Sonnenhauser 2013 for more details).

4.2. Communicative aspects

The different usage types mentioned above and their appearance in specific contexts and genres can tell us something about the communicative aspects of Turkisms. This pertains in particular to the recipients' (presumed) familiarity with the language used in the documents, and the authors' linguistic awareness and reflection upon language and its communicative functions.

4.2.1. Familiarity

Petkanova-Toteva (1965: 222–223) observes for the *damaskini* and later texts written in that tradition that Turkish lexical elements mainly belong to the domain of every-day live. Every-day lexemes are probably the easiest to borrow and the fastest to be integrated; they might not have been noticed as 'foreign' elements and they may not have been regarded as something that is needed in order to fill in specific lexical lacunae. Even puristically oriented authors use Turkisms (e.g. Kipilovski in his KN). As Kazazis (1972: 96) emphasises, “[t]here are many lexical items whose Turkish origin may or may not be transparent, depending on such things as how educated a given native speaker is, whether or not he has ever dabbled in etymology, whether he knows Turkish and how well, and so on.” This corresponds to Joseph's (2013) observation that “different speakers might make different connections in their ‘synchronic etymologizing’, i.e. in the connections that they make among forms as they set about constructing their own mental grammars” (2013: 618). This again underlines that it is not justified to assess and judge older stages of a language from a contemporary – moreover, a linguist's – perspective, applying standards that did not play a role at the time under consideration and were not even formulated as a goal to be implemented.

Petkanova-Toteva (1965: 222) concedes that many of the every-day borrowings became tightly integrated in the language and are hard to replace by Slavic words, even in the contemporary standard languages. These kinds of Turkish elements are used mostly in an integrated way, as in the following cases:⁸

- (22) a. dojde *zamanъ* umre siromaotъ [tr. *zaman* ‘time’] (KO 75)
‘the *time* came and the poor man died’
b. imam nešto *ridža* [tr. *rica* ‘request’] (KG 32)
‘I have some *request*’
c. nee nename car *baška* ot kesara [tr. *baška* ‘another, different’] (TE 118)
‘we do not have a tsar different from Ceasar’

⁸ Not all of them are still in frequent usage today; see, e.g., Ilčev (1998).

Such integrated uses of Turkish elements point towards the degree of vernacularisation of the written language. These integrated Turkisms also indicate the deep degree of – at least individual – entrenchment since they are not marked or commented as ‘foreign’.

4.2.2. Awareness

Non-integrated Turkisms appear as part of what can be called ‘explanatory structures’. They appear both as *explanans* or *explanandum*, with the *explanans* being given either in parentheses or as part of a paraphrase introduced with *sirečъ* ‘that is’ or *sanki* ‘that is [tr.]’, cf. (23):

- (23) a. Мiхailъ побѣздava *dušmani*, *sirečъ* diavoli [tr. *düşman* ‘enemy’] (RP 3)
 ‘Mikhail often times defeated enemies, that is, devils’
 b. четverougolna *sanki* sozъ četiri *kjušeta* [tr. *köşe* ‘corner’] (RP 236)
 ‘a quadrangle, so to speak with four corners’
 c. korablъ *sirečъ gemija* [tr. *gemi* ‘ship’] (RP 142)
 ‘a ship [slav.] that is a *ship* [tr.]’

However, not only the Turkisms can be found in explanatory function, there are also Slavic elements that seem to explain Turkish ones. Moreover, also different variants of Slavic are used in these explanatory contexts, mainly East Slavic or Russian Church Slavonic and their Balkan Slavic equivalents. As with Turkisms, the equivalents are given either in brackets, (24), or as paraphrases, (25). Interestingly, the latter can be introduced by Slavic *sirečъ* or Turkish *sanki* (see also (23) above):

- (24) a. v carstvoto nebesno (*raj*) (PE 19)
 ‘in the Kingdom of Heaven (*paradise*)’
 b. a otečъ (*bašta*) velikago Aleksandra (LS 1844/5, 66)
 ‘but the father (*father*) of Alexander the Great’
 c. xotjanie (*štenie*) (PE 3)
 ‘desire (*desire*)’
 (25) a. *božestvennaja služba* *sirečъ* liturkia (KO 64)
 ‘*divine service* that is liturgy’
 b. da slušatъ čitanie Bžestvennago pisanija *sanki knjige* (RP 251)
 ‘in order to listen to the reading of the divine writings so to speak *books*’

The usage of Turkisms in explanatory structures alongside with other borrowed and native lexical elements once again points towards a more general integration of Turkisms into the domain of ‘comprehensibility’ within which they are on an equal footing with genuine Balkan Slavic expressions.

On a more general level, these explanatory structures not only indicate the striving for comprehensibility, but also the interaction of the author with his text. This meta-communicative usage is one possibility of explicitly relating message and code. It thus exemplifies Jakobson’s (1971) type M/C, i.e. ‘message referring to code’, which he regards as playing „a vital role in the acquisition and use of

language“ (1971: 338). Explicitly relating the two layers of communication, this type introduces the narrator’s voice. Thereby, this meta-communicative usage of Turkisms points towards a change in the conception of the roles of scribes which now emerge as authors and as such interacts with the language they use and with their presumed recipients. This accounts for the fact that Turkisms are used in genres which allow for more creative freedom, i.e. genres which enable – or even require – an author or narrator to be visible in the texts, such as the instructions and explanations to biblical texts (for more details see Sonnenhauser 2014, Sonnenhauser forthcoming). Based on this it seems justified to interpret these meta-linguistic uses also in terms of the elaboration of Balkan Slavic both on the linguistic level (mainly elaboration of the lexicon) and the literary level (changing role of authors and innovations in genres).

4.2.3. Stylisation

There is yet another usage type of Turkisms which even more distinctly shows the intentionality involved in using Turkish borrowings. This usage type can be exemplified by Aleko Konstantinov’s employment of Turkisms in order to characterise Baj Ganju as a backwoodsman. As Kramer (1992: 45) points out, “much of the humor in the novel depends on stylistic contrast in language use”, that is, “Konstantinov recognised the stylistic marking of Turkisms and at times self-consciously drew attention to their use”. This is revealing in two respects: it indicates that by the time Baj Ganju was composed (1895), Turkisms were already marked⁹ (at least for the literary avant-garde; otherwise they would not have been obliged to tell the ‘ordinary’ people that using Turkish elements is ‘bad’ and ‘illiterate’) and shows that this markedness was made use of in order to achieve specific stylistic intentions. These stylistic effects presuppose the availability of a Balkan Slavic alternative that could have been used instead and that is known by the recipients. Turkisms that are either necessary, since they fill a terminological lacunae, or that are regarded as part of the native language, i.e. not identified as ‘foreign’, are more than unlikely to yield such effects. This is a further argument showing that analyses of the early vernacular documents as being chaotic, illiterate, stylistically low etc. miss the point (see also below).

What has been described for Baj Ganju can already be observed several decades earlier. One example is Joakim Krčovski’s *Različna poučitelna nastavljenija* (RP, 1819). This work includes one part that describes a question-answer-sequence between some clergyman and an ordinary man asking questions about religious contents, e.g. the procedures and contents of the liturgy. The questions are of the type illustrated in (26):

(26) a. Vorproszъ. čto znamenuetъ po sredě crkve amvonъ (RP 235)

⁹ Henninger (1990: 7, fn. 3) regards this „stylistic debasement“ as an unintended side-effect of the efforts to explain away Turkisms. Those that remained in usage despite the puristic tendencies became to be marked as stylistically lower.

- ‘Question: what does the pulpet in the middle of the church mean?’
- b. Voprosъ. Zašto katose počine štač liturgia duma govori Popъ, bl̑gosloveno c̑tvo, oča, i sna, i s̑ago dxa (RP 247)
‘Question. Why does the priest, when the holy liturgy is over, say say [doubling] blessed (is) the kingdom of the Father, the Son and the Holy Ghost’
- c. Voprosъ: čto estъ s̑aa liturgia
‘Question: what is the holy liturgy’ (RP 243)

Obviously, this section is addressed to people who do not know every detail of religious service and the theological background. A characteristic feature of this very specific question-answer chapter, a feature that is not observed in any of the other texts investigated here,¹⁰ is the frequent usage of synonymous doublings. These are mostly pairs of Slavic-Turkish or Turkish-Slavic, which are either coordinated by means of *i* ‘and’ or immediately adjoined (27):

- (27) a. tova pokazue platnoto i čaršafъ [tr. *çarşaf* ‘bed sheet’] (RP 236)
‘this shows the cloth and *bed sheet*’
- b. toj e stolpъ direkъ [tr. *direk* ‘mast, pole’] (RP 236)
‘this is a pile *pile*’

This usage can be interpreted as serving a comparable effect as Konstantinov’s utilisation of Turkisms. However, while Konstantinov could simply use them, since Slavic equivalents are readily available, Krčovski had to rely on Slavic-Turkish doublings in order to achieve the stylistic effect of characterising individuals as being lesser educated. These doublings are also observed by Henninger (1990: 13/fn19) who generalises that “in these redundant expressions the Slavonic word normally preceded the Turkish one, which might be interpreted as a kind of unconscious bracketing”. At least in Krčovskij’s text, this generalisation does not hold. The doubling pattern is encountered not only for Turkisms, but also for Slavic-Slavic pairs as in (28a) and even Slavic-Greek as in (28b):

- (28) a. edno telo snaga (RP 238)
‘one body body’
- b. čaša i diskoskъ (RP 238)
‘cup and bowl’

Apart from pairs, triples can be found as well, (29a,b). Moreover, variation concerns also structural patterns, as can be seen from the different (Slavic) expression of possession in (29c).

- (29) a. s̑stennikъ govori, glagoletъ duma; vъ vospominanie Gđa i Bğa, i špca našego Iisa Xřta (RP 239)
‘the priest says, says says; in commemoration of the Lord and God, and our saviour Jesus Christ’

¹⁰ Some instances can be found in the other texts included in RP (not, however, in Krčovski’s other works, such as the *Čudesna presvetye bogorodicy*, ČB, or the *Slovo iskazanoe zaradi umiranje*, SU), but only in the *talkovanije* parts, not in the religious texts themselves.

- b. da dumatъ na narodo, sanki na xorata na čĭvĕcyte (RP 252)
‘that he tells the people, that is the people the people’
- c. na ljudi svoją na svoite čĭvĕcy (RP 249)
‘to people-his to his people’

These doublings do not seem to have a (primary) prescriptive/puristic function, since this pattern cannot be observed in the other texts included in RP. Instead, seems likely that they are meant to imitate or allude to the variety of synonyms available (see Neofit Rilski’s grammar; section 3.2) in order to characterise the interlocutors as being very rural (the one asking) or as adopting to the needs of a rural audience (the one answering).

5. To conclude: the semiotics of borrowing

As has been shown in this paper, simply speaking of ‘Turkisms in pre-standardised Balkan Slavic’ neglects, on the one hand, important differences among authors and genres, and hides, on the other hand, general tendencies that begin to show despite these differences and even in the absence of over-individual norms. These tendencies range from the usage and assessment of Turkisms in order to support comprehensibility in the earliest vernacular documents to their explicit expulsion in order to create a common, ‘native’ literary language in the renaissance documents. Approaching the earlier documents from a post-renaissance perspective, judging the usage of Turkisms as chaotic and the language of these documents as indicating illiteracy on part of the authors and scribes does not do justice to their intentions. Moreover, instead of focusing on Turkisms alone it is more revealing to regard them in a common context with other ‘foreign’ elements, such as Graecisms as well as Church and East Slavonisms.

Importantly, Turkisms are revealing also beyond the aspects of borrowing, language contact and vernacularisation. They are indicative of the striving for larger comprehensibility of the written language, which marks a decisive change in the overall conception of literature. One additional indication of this crucial change is their usage in explanatory structures, which shows the growing linguistic awareness of the authors both as concerns the language used in their documents and as concerns their assumptions concerning the linguistic background of their intended audience.

In particular, it has been shown that there are different types of Turkisms, not only as regards the linguistic level or – for lexical Turkisms – domain of vocabulary. Turkisms may be used as filler for lexical gaps, if there is no Slavic equivalent or be used as integral part of the vernacular, i.e. not be recognised as ‘foreign’. These integrated uses are to be differentiated from non-integrated uses. One main type of them is instantiated by Turkisms figuring in meta-linguistic and meta-communicative structures serving an explanatory function. These structures explicitly relate the two narrative layers – that of what is narrated and that of the narrator, i.e. object and subject of narration. Thereby, the narrator’s voice becomes visible in the text, which is at the same time characteristic of a new narrative quality.

Marking Turkisms as elements to be replaced shows the increasing puristic tendencies, by which the usage of a ‘genuine’ and ‘native’ Slavic element has become more important than the general comprehensibility of the language used in written texts. This is the prerequisite for the third usage type of Turkisms observed in the texts under investigation here: their usage as stylistic device in order to characterise actors in the text in a specific manner.

These different usage types serve different semiotic functions. Integrated Turkisms are used as ‘ordinary’ symbols, while non-integrated Turkisms and the meta-linguistic structure are basically indexical, in that they relate two narrative layers. Turkisms used to characterise communicative situations and their participants as, e.g., backwards or illiterate, assume a secondary symbolic function, which exceeds the primary symbolic-denotational function of the linguistic sign as such. The functional and semiotic aspects of these usage types are summarised in Tab. 1:

Tab. 1: usage types and semiotic dimension

| Level Usage type | Primary: integrated | Secondary: non-integrated | Tertiary: non-integrated |
|---------------------|------------------------------------|---|-----------------------------|
| Function | a) filler of gaps, b) nativised | meta-communicative structure: a) explanation, b) replacement | stylisation |
| Semiotic | symbolic | indexical | secondary symbolic |

From the pragmatic analysis suggested in this paper, which embeds Turkisms into their communicative context, it becomes evident that for the evolution of the later standard languages, Turkisms play a far more important role than suggested by primarily lexicological analyses.

5. References

- FIELDER, GRACE (1999): Development of narrative strategies in nineteenth century East Balkan Slavic prose. In: Dingley, John; Ferder, Leon (eds.): *In the realm of Slavic philology: to honor the teachings and scholarship of Dean S. Worth from his UCLA students*. Bloomington: Slavica, 87–195.
- FIJUZI, XADŽAR (2004): *Persijski dumi v bālgarskija ezik*. Sofija: Avangard Prima.
- FLEISCHMANN, SUZANNE (2000): Methodologies and ideologies in historical linguistics: on working with older languages. In: Herring, Susan C. et al. (eds.): *Textual parameters in older languages*. Amsterdam: Benjamins, 33–58.
- FRIEDMAN, VICTOR (1994): Turkisms in a comparative Balkan context. In: *Septième congrès International d'Études du Sud-Est Européen: rapports*. Athens: Greek National Committee for Southeast European Studies, 521–543.
- FRIEDMAN, VICTOR (1997): One grammar, three lexicons: ideological overtones and underpinnings in the Balkan Sprachbund. In: Singer, Kora et al. (eds.): *CLS 33: Papers from the panels on linguistic ideologies in contact; universal grammar, parameters and typology; the perception of speech and other acoustic signals*. Chicago: Chicago Linguistics Society, 23–44.

- FRIEDMAN, VICTOR (2003): From orientalism to democracy and back again: Turkisms in the Balkan languages. In: Sedakova, Irina A.; Civ'jan, Tat'jana V. (eds.): *V poiskax "oriental'nogo" na Balkanax. Antičnost'. Srednevekov'e. Novoe vremja*. Moskva: Institut slavjanovedenija RAN, 61–64.
- FRIEDMAN, VICTOR (2006): Balkans as a linguistic area. In: Brown, Keith (ed.): *Encyclopedia of language and linguistics. Volume 1*. Oxford: Elsevier, 657–672.
- GOŁĄB, ZBIGNIEW (1960): The influence of Turkish upon the Macedonian Slavonic dialects. In: *Folia Orientalia* 1, 26–45.
- GRANNES, ALF ET AL. (2002): *A dictionary of Turkisms in Bulgarian*. Oslo: Instituttet for sammenlignende kulturforskning.
- HASPELMATH, MARTIN (2009): Lexical borrowing: concepts and issues. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 35–54
- HENNINGER, THOMAS (1990): The Bulgarian national revival: enforced elimination of some turkisms from the lexis. In: *Canadian Slavonic Papers* 32, 1, 1–17.
- HINRICHS, UWE (2004): Ist das Bulgarische kreolisiertes Altbulgarisch? In: Hinrichs, Uwe (ed.). *Die europäischen Sprachen auf dem Wege zum analytischen Sprachtyp*. Wiesbaden: Harrassowitz, 231–242.
- ILČEV, STEFAN (1998): *Rečnik na redki, ostareli i dialektni dumi v literaturata ni ot XIX i XX vek*. Sofija: Bălgarska Akademija na Naukite.
- ILIEVSKI, PETAR H. (2005): *Tradicija i inovacii vo makedonskite crkovnoslovenski kniževni spomenici od turskiot period*. Skopje: Makedonska akademija na naukite i umetnostite.
- JAKOBSON, ROMAN (1971): Shifters, verbal categories, and the Russian verb. In: *Selected Writings, Vol. II*. The Hague: de Gruyter, 130–147.
- JOSEPH, BRIAN (2013): Multiple sources and multiple causes multiply explored. In: *Studies in language* 37, 3, 675–691.
- KAPPLER, MATTHIAS (2002): Türkisch (in Südosteuropa). In: *Enzyklopädie des Europäischen Ostens*. Klagenfurt: Wieser, 817–834.
- KAZAZIS, KOSTAS (1972): The status of Turkisms in present-day Balkan languages. In: Birnbaum, Henrik; Vryonis, Speros (eds.): *Aspects of the Balkans*. The Hague: Mouton, 87–116.
- KRAMER, CHRISTINA (1992): The use of Turkisms in Balkan languages as reflected in Konstantinov's novel Baj Ganjo. In: *Zeitschrift für Balkanologie* 28, 44–60.
- KRÄSTEVA, VESELA (2003): *Tălkoven rečnik na turcizmite v bălgarskija ezik. S ilustrativen material ot literaturata, folklorata, presata, radioto i televizijata*. Sofija: Skorpio.
- KUPERDYAEV, ALEXANDER; HANSEN, BJÖRN (2013): Neue Perspektiven zu den türkischen Elementen in den südosteuropäischen Sprachen. Aktuelle Turzismen des Makedonischen im Lichte von Entlehnungsskalen. In: Grković-Mejdžor, Jasmina; Loma, Aleksandar (eds.): *Miklosichiana bicentennialia: zbornik u čast dvestote godišnjice rođenja Franca Miklošiča*. Beograd, 381–406.
- MEERMANN, ANASTASIA; SONNENHAUSER, BARBARA (forthcoming): Das Perfekt im Serbischen zwischen Slavischer und Balkanslavischer Entwicklung. In: Sonnenhauser, Barbara; Bazhutkina, Alena (eds.): *Beiträge zur Slavistik XXII*. München: Sagner.
- MEVSIM, XJUSEIN; ČĀKAROVA, KRASIMIRA (2007): Otnovo za sădbata na turcizmite v săvremennija bălgarski ezik (izolacija i/ili rehabilitacija). In: *LiterNet* 10/95 <http://litenet.bg/publish18/h_mevsim/turcizmi.htm> [aufgerufen am 14.3.2014]

- MIKLOSICH, FRANZ (1884): *Die türkischen Elemente in den südost- und osteuropäischen Sprachen.* (Griechisch, Albanisch, Rumunisch, Bulgarisch, Serbisch, Kleinrussisch, Grossrussisch, Polnisch). Wien.
- MILETIČŤ, LJUBOMIR (1908): *Koprištenski damaskinъ. Novobălgarski pametnikъ otъ XVII. vĕkъ.* Sofija: Ministerstvo na Narodnata Prosvĕta.
- MILETIČŤ, LJUBOMIR (1923): *Svištovski damaskinъ. Novobălgarski pametnikъ otъ XVIII. vĕkъ.* Sofija: Bălgarska Akademija na Naukite.
- MIRČEV, KIRIL (1963): *Istoričeska gramatika na bălgarskija ezik.* Sofija: Nauka i izkustvo.
- NEWEKLOWSKY, GERHARD (2012): Greek and Turkish influence upon the Balkan Slavic verb. In: Kahl, Thede et al. (eds.): *Balkanismen heute.* Wien: LIT, 195–208.
- NOEL AZIZ HANNA, PATRIZIA; SONNENHAUSER, BARBARA (2013): Verschriftlichung zwischen System und Rede – zur orthographischen Kodierung von Performanzstrukturen. In: *Sprachwissenschaft* 38, 311–341.
- PETKANOVA-TOTEVA, DONKA (1965): *Damaskinite v bălgarskata literatura.* Sofija: Bălgarska Akademija na Naukite
- REINKOWSKI, MAURUS (2002): Kulturerbe oder Erblast? Zum Status der Turzismen in den Sprachen Südosteuropas, insbesondere des Bosnischen. In: *Mediterranean Language Review* 14, 98–112.
- SELSKI, LJUDVIG (2012): Turski elementi v bălgarskoto slovoobrazuvane. In: *iReteslaw – online repository of texts in the field of Slavic studies* <<http://www.ireteslaw.ispan.waw.pl/handle/123456789/67>> [aufgerufen am 03.06.2015]
- SILVERSTEIN, MICHAEL (1979): Language structure and linguistic ideology. In: Clyne, Paul R et al. (eds.): *The elements: A parasection on linguistic units and levels.* Chicago: Chicago Linguistics Society, 193–247.
- SONNENHAUSER, BARBARA (2014): Erzählen im Bulgarischen zwischen Texttyp und Genre. Pressebericht und Evangelium, *Krimi i pravo* und *Tălkovanie*. In: *Die Welt der Slaven* 59/2, 355–381.
- SONNENHAUSER, BARBARA (forthcoming): Sprachliche Strukturen, narrative Strategien. Zum funktionalen Sprachwandel im vorstandardisierten Balkanslavischen am Beispiel der Vita der Petka Tărnovska und des *Sbornik* von Pop Punčo. In: *Zeitschrift für slavische Philologie*.
- STACHOWSKI, STANISŁAW (2014): Türkischer Einfluss auf den slavischen Wortschatz. In: Kempgen, Sebastian et al. (eds.): *Die slavischen Sprachen.* Berlin: de Gruyter, 1198–1210.
- STAMENOV, MAKSIM (2014): Po povod aktualnostta na edna klasifikacija na turcizmite v bălgarskija ezik otpredi sto godini. In: *Rocznik Slawistyczny* 63, 113–128.
- ZENNER, ELINE; SPEELMAN, DIRK; GEERAERTS, DIRK (2014): Core vocabulary, borrowability and entrenchment. A usage-based onomasiological approach. In: *Diachronica* 31, 1, 74–105.

5.1. Sources

- ČB: Joakim Krčovski (1817): *Čudesna presvetyę bogorodicy.* Budimъ.
- CV: *Carigradski vĕstnikъ.*
- E: *Edirne.*
- GRAMMATIKA: Neofit Rilski (1835): *Bolgarska grammatika.* Kraguevac.

- KG: Kiril Pejčinović (1840): *Kniga glagolemaę utęšenje Gręšnymъ*. Solunъ.
- KN: Anastas Kipilovski (1836): *Kratkoe načertanie na vseobštata istorie* (übers. aus d. Russ.). Budimъ.
- KO: Kiril Peičinović (1816): *Kniga się zovomaę Ogledalo*. Budinъ Gradъ.
- LS: *Ljuboslovie*.
- NEDELNIK (1806): Sofronij Vračanskij. *Nedelnik*.
- NOVAKOVIĆ: Novaković, S. (1877): Život sv. Petke od patrijarha bugarskoga Jeftimija. In: *Starine* 9, 48–59. (this is an edition of the version of the *Life of Paraskeva of Epibatai* printed by Božidar Vuković in 1536).
- PE: *Poučitelno Evangelie* 1868.
- PUNČO: *Pop Punčov Sbornik* 1796.
- RB: *Bukvarъ sъ različny poučenę sobrani ot Petra X. Beroviča za Bolgarski-tę učilištę* 1824.
- RP: Joakim Krčovski (1819): *Različna poučitelna nastavljenija*. Budinъ.
- SU: Joakim Krčovski (1814): *Slovo iskazanoę zaradi umiranje*. Budinъ.
- SVIŠTOVSKI: Miletičъ, Ljubomir (1923): *Svištovski damaskinъ. Novobălgarski pametnikъ otъ XVIII. věkъ*. Sofija.
- TE: Miletičъ, Ljubomir (1920): *Dva bălgarski rękopisa sъ gręcko pismo*. 1. *Nedělni poučenija otъ XVIII věkъ*. 2. *Tărlisko evangelie otъ 1861 god*. Sofija.
- TIXONRAVOVSKI: Demina, E.I. (1971): *Tixonravovskij damaskin*. *Bolgarskij pamjatnik XVII v. Issledovanie i tekst*. Sofija.

Ruprecht von Waldenfels (Kraków)
Inner-Slavic contact from a corpus driven perspective

Abstract

Dieser Beitrag behandelt einen korpusbasierten Ansatz mit dem Ziel eines Vergleiches der funktionalen Ähnlichkeit der Derivation in unterschiedlichen slawischen Sprachen. Auf der Basis der Verteilung von verwandten Verbalpräfixen und Nominalsuffixen in einem Paralleltextrkorpus wird gezeigt, dass sich diese beiden Domänen signifikant voneinander unterscheiden und die slawischen Sprachen in jeweils unterschiedliche Gruppen zerfallen, die durch ähnlichen Gebrauch bedingt sind. Die wichtigsten Unterschiede zeigen einen Hang des Slowenischen zusammen mit westslawischen Verbalpräfixen, aber nicht mit Nominalsuffixen aufzutreten. Das Bulgarische hingegen bildet auf der Basis der Nominalsuffixe, nicht jedoch der Verbalpräfixe eine gemeinsame Gruppe mit dem Russischen. Die Unterschiede lassen sich durch die bekannte sprachübergreifende Tendenz der leichteren Entlehnbarkeit von Nomen als von Verben erklären. Darüber hinaus wird die unterschiedliche Rolle des innerlawischen Sprachkontaktes bezüglich dieser beiden Derivationsdomänen manifest. Diese domänenspezifischen Unterschiede sind wichtig für ein umfassendes Verstehen der Geschichte der slawischen Sprachen.

1. Introduction

Investigating the functional similarity of derivational affixes in different languages is difficult due to the mixed nature of their denotational potential. By definition, such formants only partly motivate the semantics of the lexemes they derive (see, e.g., Haspelmath/Sims 2010: 94). Consequently, it is difficult to define and compare their meaning potential. Moreover, their semantics, as much as it can be discerned, is often highly abstract and their functional potential may be captured rather in terms of differing levels of productivity with different subtypes than with an explicit elaboration of senses. For example, a suffix may be freely used for the derivation of nomina agentis, but rather restricted as a formant of nomina instrumenti; for a different suffix the reverse may be true. It follows that a cross-linguistic comparison of derivational morphology faces serious obstacles if suffixes across a larger number of languages are to be compared on the basis of an explicit account of their functional potential.

In the present article I exploit a parallel corpus using an extensionally based, bottom up approach so as to compare the functional scope of cognate derivational morphology across Slavic languages. Theoretically, the approach involves a distributional model of semantics (see e.g. Sahlgren 2008). Instead of proceeding from an analysis of the meaning of cognate derivational affixes, I start from analyzing their distribution in a parallel, i.e., translated text. The basic insight is: If a cognate affix is often used in parallel translations of the same noun into different languages, this affix has a similar functional potential.

If, conversely, a cognate affix is *not* often used in parallel translations, the functional potential differs. A comparison of the context where this affix is or is not used, can then be used to define more precisely differences and similarities in a qualitative way.

In view of the limited space, the present article contains a sketch based on a single parallel text and a method that rests on agglomerated data. Visualizations of the distribution of the affixes in the corpus are systematically compared on the basis of a list of diagnostic questions, and random samples taken from the corpus are used to gauge the error rate of the aggregated data.

The remainder of the article is structured as follows: Section 2 introduces the approach and the data used in greater detail. In section 3 the data are analyzed and discussed; section 4 concludes the paper.

2. Method and Data

2.1. Investigating Functional Convergence on the Basis of a Parallel Corpus

In the present paper I compare cognate verbal prefixes and cognate nominal suffixes across 10 Slavic standard languages or standard variants¹ in 14 versions (7 standard languages are represented in a single version; BCS is represented in two Serbian and one Croatian translation, and there are two translations into Ukrainian and Polish respectively). The parallel text used is Mikhail Bulgakov's novel *Master i Margarita* as included in the ParaSol corpus (Waldenfels 2011).

The basic idea is to compare whether the same or different affixes are used in different languages across a large number of translated segments. The following two examples illustrate this approach (see Waldenfels 2015 for an earlier version of this method):

(1) Bulgakov, *Master i Margarita*

| | | |
|----|---|--|
| ru | S | Вот кого с особенным удовольствием отпущу, - сказал Воланд, [...] |
| uk | E | Ось кого з особливим задоволенням відпущу [...] |
| pl | O | Tego odprawię ze szczególną przyjemnością [...] |
| cz | E | Toho propustím se zvláštním potěšením , [...] |
| sk | O | Tohto prepustím s radost'ou , [...] |
| sl | V | Tu je nekdo, ki ga odpuščam s posebnim zadovoljstvom [...] |
| hr | V | Evo, koga ću sa zadovoljstvom otpustiti [...] |
| sr | V | A ovoga ću sa velikim zadovoljstvom pustiti [...] |
| mk | V | Ете кого ќе го пуштам со особено задоволство [...] |
| bg | S | Виж, него с най-голямо удоволствие ще го пусна да си върви [...] |

¹ Language abbreviations (counter-clockwise from Russian to Bulgarian): ru = Russian, uk = Ukrainian, by = Belarusian, pl = Polish, us = Upper Sorbian, cz = Czech, sk = Slovak, sl = Slovene, hr = Croatian, sr = Serbian, bcs = Bosnian/Croatian/Serbian, mk = Macedonian, bg = Bulgarian. ESl East Slavic, WSl West Slavic, SSl South Slavic.

This example includes 5 different suffix classes concerning cognate formants in different languages. Here and below, these classes are designated in capital letters and abbreviated in the examples to single letters: S = STVIE (-*ство* in Russian and Bulgarian); E = ENIE (-*ення* in Ukrainian, -*ení* in Czech); O = OST (-*ość* in Polish, -*ost'* in Slovak) and V = STVO (-*stvo* in Slovene, Croatian, Serbian, -*ство* in Macedonian). Note that for reasons of readability, I do not use a phoneme-based representation of these suffixes, but the orthographic representation of the morpheme together with the nominative singular ending.

If we were to assume that all examples of abstract nouns in the book were parallel to this one, we would be led to the conclusion that Russian and Bulgarian, Ukrainian and Czech, Polish and Slovak, as well as all the South Slavic languages excluding Bulgarian, form four distinct groups of suffix usage. In fact, of course, corpus uses of these suffixed nouns tend to show very different patterns; consider, for example, the next case (taken for convenience from a different text):

(2) Ostrovskij, *Kak zakaljalas' stal'*

| | | |
|----|---|--|
| ru | E | Ладно, - крикнул, теряя терпение , Тимошенко, [...] |
| by | O | Добра, - крыкнуў, трацячы цярліваць , Тімашэнка. [...] |
| uk | S | Гаразд, - крикнув Тимошенко, якому урвався терпець , [...] |
| pl | O | Dobrze - krzyknął tracąc cierpliwość Timoszenko [...] |
| cz | O | "Dobře, " řekl Timošenko, ztrácеje trpělivost , [...] |
| sk | O | Dobre, - Timošenko už strácal trpezlivosť , [...] |
| sl | E | »Prav,« je kriknil Timošenko, ki ga je bilo potrpljenje minilo. [...] |
| hr | E | Dobro, - viknu Timošenko gubeći strpljenje , [...] |
| sr | E | Dobro, - viknu Timošenko gubeći strpljenje , [...] |
| mk | E | Добро, - губејќи трпение , викна Тимошенко, [...] |
| bg | E | Добре — извика Тимошенко, изгубил търпение , [...] |

Here, Russian and all South Slavic languages show the same suffix class ENIE, while Belarusian and West Slavic exhibit OST. The Ukrainian translation stands alone in using the suffix EC.

Such and other examples contain information along two dimensions: They reveal something about (a) which languages are similar to others in respect to affix use, and (b) which affixes are functionally similar to which other affixes. This latter point can be stated from the perspective of affix classes (i.e., example 1 and 2 contain evidence that STVIE and OST overlap in their use in Slavic) and of language specific affixes (i.e., example 1 and 2 contain evidence that -*ość* in Polish and -*stvo* in Slovene overlap in function). In the present paper, I focus on the first point, that is, the similarity of languages in respect to the use of affix classes (see Waldenfels 2015 for further possibilities).

Taken by itself, the information in individual attestations is essentially anecdotal, and it is difficult to assert its significance. In order to extract robust tendencies it is necessary to aggregate many examples so

that regular patterns can take the upper hand over random factors such as the variation inherent in translation. To this end, the process of comparison is automated.

2.2. Automatization

The automatization of the comparison has three preconditions; two conditions concern the corpus, the remaining one relates to the way the question is formalized:

- a. the corpus is word aligned, so that word forms can be automatically compared
- b. the corpus is lemmatized and POS-tagged, so that the comparison can be based on the citation form of the word forms and restricted to the relevant class
- c. the morphological classification is formally operationalized (see van Zundert et al. 2012:280 for the significance of formalization in digital humanities).

For the purposes of this task, a subset of the ParaSol corpus, which is lemmatized and POS-tagged (Waldenfels 2011), has been word aligned using Uplug (Tiedemann 2003), so that conditions a and b are met.

The formalization condition (c) is accomplished by pattern matching over the above annotation; for example, each Russian word form that is tagged as a noun (that is, fits the regular expression “N.*”) and whose lemma ends on *-ость* (that is, fits the regular expression “*ость\$*”) is considered to be derived with the suffix *{/ost’/}*. As is often the case with automated analysis, this operationalization introduces a certain error rate. For example, forms of the Russian lemma *гость* ‘guest’ will be erroneously labelled as involving the suffix *{/ost’/}*. To deal with complications like these, the prefixes and suffixes for the present article were chosen to involve a relatively low error rate during recognition. Moreover, random samples of the data are inspected in a browser to recognize systematic tendencies that are due to such errors.

2.3. Operationalization and Visualization of the Data

Tab. 1 and 2 show the regular expressions which define the prefix and suffix classes. The operationalizations are mainly based on Kopečný (1973), Vaillant (1974) as well as standard lexicographic sources for the languages involved, and supplemented with corpus searches in the parallel corpus to check for variants and to ensure a small error rate for the regular expressions.

Care was taken to minimize both false negatives and false positives and to derive comparable sets across all languages. This sometimes leads to unintuitive expressions for some languages which are explained by restrictions in other languages; for example, the prefix class O is defined for BCS as involving words that start with ‘o’ but do not continue in ‘d’ or ‘t’ in order to delimit them from words belonging to the class OT which contains words starting with both ‘od’ and ‘ot’. An analogous restriction is

also applied in Russian, even though Russian words with OT are never written with 'од'. The reason is to make sure that the class of erroneously classified verbs is similar in both languages; for example, in this way both BCS *odobriti* and its Russian cognate *одобрить* are not classified as belonging to the O-class. While this is a mistake, it applies to all Slavic languages in the same way, and thus reduces the impact of the error.

Tab. 1: Regular expressions and token counts for nominal suffixes (number of lemma types given in brackets)

| | AR | eK | IK | ICA | NIK | NICA | EC | OST | STVO | NIE | CIJA | KA |
|-----------|-----------|----------------------------------|----------------------------|-----------------------------|------------|-------------|---------|----------|-------------|------------|------------|--------------------------|
| ru | .*ар\$ | .*([чш]ек[[^] чш]ок)\$ | .*[[^] н]ик\$ | .*[[^] н][иы]ца\$ | .*ник\$ | .*ница\$ | .*ец\$ | .*ость\$ | .*ство\$ | .*н[ьи]е\$ | .*ция\$ | .*[[^] иы]ка\$ |
| by | .*ар\$ | .*ак\$ | .*[[^] н]ік\$ | .*[[^] н][іы]ца\$ | .*нік\$ | .*н[іы]ца\$ | .*ец\$ | .*асць\$ | .*[сц]тва\$ | .*нне\$ | .*цыя\$ | .*[[^] іы]ка\$ |
| uk | .*ар\$ | .*ок\$ | .*[[^] н]ик\$ | .*[[^] н]иця\$ | .*ник\$ | .*ниця\$ | .*ець\$ | .*ість\$ | .*[сц]тво\$ | .*ння\$ | .*ція\$ | .*[[^] иы]ка\$ |
| pl | .*arz\$ | .*ek\$ | .*[[^] n]ik\$ | .*[[^] n][iy]ca\$ | .*nik\$ | .*n[iy]ca\$ | .*ec\$ | .*ość\$ | .*[cs]two\$ | .*nie\$ | .*c[ij]a\$ | .*[[^] iy]ka\$ |
| cz | .*[áa]r\$ | .*ek\$ | .*[[^] n][íí]k\$ | .*[[^] n][íí]ce\$ | .*n[íí]k\$ | .*n[íí]ce\$ | .*ec\$ | .*ost\$ | .*stvo\$ | .*ní\$ | .*ce\$ | .*[[^] íí]ka\$ |
| sk | .*[áa]r\$ | .*ok\$ | .*[[^] n][íí]k\$ | .*[[^] n][íí]ca\$ | .*n[íí]k\$ | .*n[íí]ca\$ | .*ec\$ | .*ost\$ | .*stvo\$ | .*nie\$ | .*cia\$ | .*[[^] íí]ka\$ |
| sl | .*ar\$ | .*ek\$ | .*[[^] n]ik\$ | .*[[^] n]ica\$ | .*nik\$ | .*nica\$ | .*ec\$ | .*ost\$ | .*[šs]tvo\$ | .*nje\$ | .*cija\$ | .*[[^] i]ka\$ |
| hr | .*ar\$ | .*ak\$ | .*[[^] n]ik\$ | .*[[^] n]ica\$ | .*nik\$ | .*nica\$ | .*ac\$ | .*ost\$ | .*[šs]tvo\$ | .*nje\$ | .*cija\$ | .*[[^] i]ka\$ |
| sr | .*ar\$ | .*ak\$ | .*[[^] n]ik\$ | .*[[^] n]ica\$ | .*nik\$ | .*nica\$ | .*ac\$ | .*ost\$ | .*[šs]tvo\$ | .*nje\$ | .*cija\$ | .*[[^] i]ka\$ |
| mk | .*ар\$ | .*ок\$ | .*[[^] н]ик\$ | .*[[^] н]ица\$ | .*ник\$ | .*ница\$ | .*ец\$ | .*ост\$ | .*[шс]тво\$ | .*ње\$ | .*ција\$ | .*[[^] и]ка\$ |
| bg | .*ар\$ | .*ък\$ | .*[[^] н]ик\$ | .*[[^] н]ица\$ | .*ник\$ | .*ница\$ | .*ец\$ | .*ост\$ | .*ство\$ | .*ние\$ | .*ция\$ | .*[[^] и]ка\$ |

Tab. 1: Continued

| | AR | eK | IK | ICA | NIK | NICA | EC | OST | STVO | NIE | CIJA | KA |
|------------|--------------|---------------|-------------|-------------|--------------|-------------|--------------|--------------|-------------|---------------|--------------|---------------|
| ru | 84 (11) | 541 (133) | 414 (85) | 161 (27) | 302 (70) | 161 (33) | 355 (50) | 469 (130) | 193 (62) | 1110 (318) | 102 (29) | 1984 (404) |
| by | 1314 (70) | 600 (114) | 130 (37) | 152 (19) | 306 (75) | 121 (22) | 307 (39) | 287 (126) | 120 (27) | 750 (198) | 135 (32) | 2449 (406) |
| uk | 340 (38) | 763 (171) | 261 (64) | 119 (25) | 409 (103) | 120 (35) | 447 (82) | 311 (124) | 101 (29) | 597 (235) | 125 (36) | 2120 (435) |
| uka | 253 (42) | 877 (202) | 283 (81) | 142 (27) | 433 (102) | 127 (41) | 491 (107) | 296 (105) | 95 (30) | 684 (289) | 123 (35) | 2200 (484) |
| pla | 372 (33) | 1138 (225) | 115 (34) | 71 (14) | 375 (65) | 88 (33) | 424 (54) | 670 (149) | 144 (46) | 286 (99) | 333 (73) | 1738 (412) |
| pl | 377 (33) | 1197 (205) | 119 (37) | 88 (15) | 183 (59) | 96 (26) | 392 (48) | 685 (149) | 127 (41) | 1152 (349) | 271 (72) | 1692 (383) |
| cz | 139 (23) | 835 (218) | 357 (61) | 245 (45) | 457 (75) | 111 (34) | 492 (91) | 552 (128) | 34 (7) | 662 (258) | 701 (142) | 2002 (412) |
| sk | 606 (60) | 587 (126) | 376 (97) | 188 (44) | 458 (93) | 96 (32) | 618 (87) | 583 (138) | 114 (32) | 481 (194) | 189 (41) | 2132 (435) |
| sl | 441 | 1229 | 97 | 838 | 865 | 485 | 973 | 588 | 163 | 845 | 89 | 1211 |

| | | | | | | | | | | | | |
|----------------|--------------|---------------|-------------|--------------|-------------|--------------|--------------|--------------|-------------|--------------|-------------|---------------|
| | (55) | (162) | (19) | (135) | (113) | (109) | (186) | (156) | (41) | (217) | (10) | (140) |
| hr | 420 (72) | 937 (162) | 128 (19) | 869 (161) | 666 (93) | 457 (106) | 582 (106) | 456 (118) | 127 (42) | 611 (236) | 129 (34) | 999 (146) |
| sra | 594 (115) | 1086 (233) | 180 (31) | 812 (200) | 602 (94) | 310 (79) | 576 (138) | 525 (138) | 93 (32) | 851 (355) | 107 (29) | 1035 (250) |
| sr | 637 (106) | 1061 (220) | 227 (33) | 806 (171) | 659 (95) | 287 (80) | 652 (117) | 534 (111) | 92 (35) | 734 (275) | 90 (26) | 1090 (233) |
| mk | 506 (79) | 202 (43) | 126 (26) | 261 (58) | 471 (99) | 319 (75) | 410 (62) | 423 (153) | 157 (45) | 688 (274) | 111 (32) | 2401 (387) |
| bg | 289 (46) | 164 (34) | 118 (24) | 327 (62) | 402 (83) | 161 (54) | 394 (53) | 427 (107) | 149 (53) | 648 (177) | 148 (37) | 2030 (366) |
| <i>average</i> | 455 (56) | 801 (161) | 209 (46) | 363 (72) | 471 (87) | 210 (54) | 508 (87) | 486 (131) | 122 (37) | 721 (248) | 190 (45) | 1792 (350) |

Tab. 2: Regular expressions and token counts for verbal prefixes (number of lemma types given in brackets)

| | UV | O | OT | PO | DO |
|-----------|------------------------------------|----------------|----------------------|-----------------|------------|
| ru | ^(y[^sz] в[^зсоы] во[^зс]).*.* | ^o[^тд].* | ^от.* | ^по[^дт].* | ^до.* |
| by | ^(y[^sz] в[^зсаоы] в[ао][^зс]).*.* | ^a[^тд].* | ^ад.* | ^па[^дт].* | ^да.* |
| uk | ^(y[^sz] в[^зсоиі] во[^зс]).*.* | ^o[^тд].* | ^в[іі]д.* | ^по[^дт].* | ^до.* |
| pl | ^(u w[^zsy]).* | ^o[^дт].* | ^od.* | ^po[^дт].* | ^do.* |
| cz | ^(ne)?(u v[^zsy]).* | ^(ne)?o[^дт].* | ^(ne)?od.* | ^(ne)?po[^дт].* | ^(ne)?do.* |
| sk | ^(ne)?(u v[^zsy]).* | ^(ne)?o[^дт].* | ^(ne)?od.* | ^(ne)?po[^дт].* | ^(ne)?do.* |
| sl | ^(u v[^zsi]).* | ^o[^дт].* | ^od.* | ^po[^дт].* | ^do.* |
| hr | ^u[^sz].* | ^o[^дт].* | ^(od ot[^aeiourj]).* | ^po[^дт].* | ^do.* |
| sr | ^u[^sz].* | ^o[^дт].* | ^(od ot[^aeiourj]).* | ^po[^дт].* | ^do.* |
| mk | ^(y[^sz] в[^зсои] во[^зс]).*.* | ^o[^дт].* | ^(од от[^aeиourл]).* | ^по[^дт].* | ^до.* |
| bg | ^(y[^sz] в[^зсои] во[^зс]).*.* | ^o[^тд].* | ^от.* | ^по[^дт].* | ^до.* |

Tab. 2: Continued

| | ZA | NA | PROPERE | RAZ | PRI |
|-----------|-----------|------------|--------------------|------------|------------|
| ru | ^за.* | ^на[^дт].* | ^п(е?ре ро)[^дт].* | ^ра[зс].* | ^при.* |
| by | ^за.* | ^на[^дт].* | ^п(е?ра ра)[^дт].* | ^ра[зс].* | ^пры.* |

| | | | | | |
|-----------|------------|-----------------|---------------------|-------------|----------------|
| uk | ^за.* | ^на[^дт].* | ^п(е?ре по)[^дт].* | ^ро[зс].* | ^при.* |
| pl | ^за.* | ^на[^дт].* | ^prz[eo][^дт].* | ^roz.* | ^przy.* |
| cz | ^(ne)?за.* | ^(ne)?на[^дт].* | ^(ne)?př[eo][^дт].* | ^(ne)?roz.* | ^(ne)?př[íí].* |
| sk | ^(ne)?за.* | ^(ne)?на[^дт].* | ^(ne)?pr[eo][^дт].* | ^(ne)?roz.* | ^(ne)?pr[íí].* |
| sl | ^за.* | ^на[^дт].* | ^pr[eo][^дт].* | ^raz.* | ^pri.* |
| hr | ^за.* | ^на[^дт].* | ^pr[eo][^дт].* | ^ra[zs].* | ^pri.* |
| sr | ^за.* | ^на[^дт].* | ^pr[eo][^дт].* | ^ra[zs].* | ^pri.* |
| mk | ^за.* | ^на[^дт].* | ^пр[ео][^дт].* | ^ра[зс].* | ^при.* |
| bg | ^за.* | ^на[^дт].* | ^пр[ео][^дт].* | ^раз.* | ^при.* |

Tab. 2: Continued

| | UV | O | OT | PO | DO | ZA | NA | PROPERE | RAZ | PRI |
|------------|---------------|---------------|--------------|---------------|--------------|---------------|--------------|----------------|--------------|--------------|
| ru | 1948 (359) | 1141 (303) | 897 (169) | 2041 (325) | 255 (70) | 1158 (356) | 655 (160) | 1018 (251) | 642 (252) | 783 (173) |
| by | 1884 (284) | 569 (203) | 758 (164) | 2198 (330) | 559 (94) | 1454 (382) | 465 (151) | 971 (243) | 615 (211) | 493 (132) |
| uk | 1124 (325) | 465 (157) | 818 (147) | 1784 (359) | 405 (91) | 1545 (469) | 465 (171) | 981 (271) | 526 (200) | 485 (176) |
| uka | 1254 (340) | 469 (177) | 784 (141) | 1613 (373) | 378 (92) | 1416 (441) | 442 (159) | 952 (259) | 544 (199) | 409 (145) |
| pla | 2145 (318) | 743 (188) | 762 (106) | 1962 (275) | 420 (89) | 2027 (389) | 367 (95) | 605 (162) | 536 (163) | 502 (121) |
| pl | 2293 (337) | 817 (206) | 830 (129) | 1990 (275) | 433 (88) | 2371 (434) | 404 (99) | 678 (175) | 560 (179) | 621 (146) |
| cz | 2061 (323) | 762 (156) | 624 (117) | 1256 (219) | 598 (105) | 1710 (349) | 507 (122) | 318 (76) | 545 (154) | 583 (106) |
| sk | 1953 (300) | 851 (188) | 753 (139) | 1787 (256) | 505 (107) | 1581 (393) | 371 (132) | 723 (177) | 537 (152) | 569 (113) |
| sl | 1568 (153) | 795 (150) | 1030 (90) | 2038 (214) | 352 (41) | 1991 (269) | 539 (93) | 480 (118) | 370 (96) | 965 (108) |
| hr | 1221 (168) | 866 (192) | 918 (120) | 1725 (180) | 765 (99) | 1167 (279) | 789 (138) | 888 (183) | 344 (99) | 452 (94) |
| sra | 768 (224) | 819 (248) | 568 (157) | 1367 (264) | 574 (115) | 766 (287) | 654 (192) | 676 (238) | 267 (131) | 348 (120) |
| sr | 696 | 788 | 485 | 1831 | 506 | 760 | 622 | 726 | 189 | 347 |

| | | | | | | | | | | |
|----------------|-----------------------------|----------------------------|----------------------------|-----------------------------|---------------------------|-----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| | (208) | (213) | (135) | (285) | (102) | (244) | (154) | (220) | (88) | (90) |
| mk | 1266 (253) | 668 (167) | 936 (128) | 1731 (260) | 687 (114) | 1073 (303) | 926 (225) | 826 (219) | 601 (126) | 287 (78) |
| bg | 1574 (197) | 1234 (200) | 891 (116) | 1637 (230) | 385 (51) | 1277 (315) | 808 (156) | 765 (192) | 686 (167) | 447 (100) |
| <i>average</i> | <i>1554</i> <i>(271)</i> | <i>785</i> <i>(196)</i> | <i>790</i> <i>(133)</i> | <i>1783</i> <i>(275)</i> | <i>487</i> <i>(90)</i> | <i>1450</i> <i>(351)</i> | <i>572</i> <i>(146)</i> | <i>758</i> <i>(199)</i> | <i>497</i> <i>(158)</i> | <i>521</i> <i>(122)</i> |

Tab. 1 and 2 specify corpus counts of types and tokens for each affix class as defined by the regular expressions. Again, in order to restrict the impact of the error rate, we use only such suffixes that have a minimum mean frequency of at least 30 types and 100 tokens in the text. This gives the final list of prefixes and suffixes considered for this paper. In the following paragraph, I introduce the list of affixes with illustrative examples from each branch of Slavic, and some comments if in order. The following prefixes classes are used:

- UV: The Common Slavic prefixes *u (Kopečný 1974: 261) and *vъ (Kopečný 1974:274–275) have partly merged in several languages; in order to achieve comparability across Slavic, they are conflated into a single group here. Examples: ru *увидеть, вбежать*; pl *uspokoić, wcisnąć*; sl *udariti, vmešati*; bg *вляза, убия*.
- ZA (Kopečný 1974: 283–294): ru *записывать*; cz *zapomenout*; bcs *zaspati*.
- DO (Kopečný 1974: 65–66): ru *добавить*; pl *dodać*; sl *dovoliti*.
- NA (Kopečný 1974: 122–123): ru *напугать*; pl *należeć*; bg *наричам*.
- OT (Kopečný 1974: 154–155): uk *відвезти*; pl *odsunąć*; bcs *otvoriti*.
- O (Kopečný 1974: 138–140): by *азірнуцца*; pl *objąć*; bcs *oživjeti/oživeti*.
- PO (Kopečný 1974: 186–189): uk *поїхати*; cz *počkat*; sl *postaviti*.
- PRI (Kopečný 1974: 213–215): by *прыехаць*; cz *přitisknout*; mk *привлекувам*.
- PROPERE: Common Slavic *pro (Kopečný 1974: 217–219) and *per (Kopečný 1974: 164–167) have merged in parts of West Slavic and Slovene and are conflated here (see above UV). ru *проходить, переїти*; pl *przejsć, przeczytać*; sl *prebrati, prebivati*; bg *прочета, прелетя*.
- RAZ: uk *розтанути*; cz *rozloučit*; hr *razrušiti*.

The following suffix classes were used:

- AR (Vaillant 1974: 318): uk *лікар*; sk *novinár*; mk *болничар*; this is an early loan with generally low frequency (see below discussion of results) .
- CIJA: by *галюцынацыя*; cz *situace*; bcs *kremacija*. CIJA mostly attaches to loan words; it is not inherited, but nevertheless present in all Slavic languages.
- EC (Vaillant 1974: 299–300): ru *красавец*; sk *cudzinec*; bg *старец*.
- eK (Vaillant 1974: 302–204): ru *листок*; pl *kawałek*; bcs *podatak*.
- ICA (Vaillant 1974: 340): uk *продавщиця*; pl *błyskawica*; sl *rokavica*; bcs *radionica*; this suffix is extraordinarily frequent in bcs.
- IK (Vaillant 1974: 539): by *зрузавік*; pl *stolik*; bg *старик*; this suffix is virtually not used in the South Slavic languages, the Bulgarian example is one of few.
- KA (Vaillant 1974: 353–354): ru *кенка*; cz *zápalka*; mk *печка*.
- NICA (Vaillant 1974: 346): uk *робітниця*; sk *kadidelnica*; bcs *prodavnica*; bg *покойниця*.

- NIE (Vaillant 1974: 401): ru *правление*; sk *podozrenie*; sl *navdušenje*; this suffix is originally a combination of *-n- and *-ъje; in some languages, notably Polish, it has become a part of the verbal paradigm.
- NIK (Vaillant 1974: 306): by *субяседник*; sk *návštevník*; bcs *potočník*.
- OST (Vaillant 1974:374–376): uk *радість*; cz *rychlost*; mk *бесмртност*.
- STVO (Vaillant 1974: 408 ff.): uk *вбивство*; sk *svinstvo*; bg *общество*.

Especially the group of nominal suffixes is very heterogeneous, both in respect to when they entered the Slavic languages (for example, CIJA is a late loan), but also in respect to their known productivity. For example, while IK is a widely used diminutive suffix in North Slavic, it is much less frequent in South Slavic, while ICA is very productive in bcs, but much less so in North Slavic or indeed other South Slavic languages. However, note that similarity in non-use, that is, the missing of a certain function, is also potentially revealing, and these suffixes were not removed at this stage. There is, however, the problem that, if a certain affix is infrequent, erroneous classifications can take the upper hand. This is kept in mind during the analysis.

2.4. Qualitative and Quantitative Perspective

The corpus is evaluated based on this operationalization in two different ways. First, the data is investigated from a qualitative perspective. Random samples of instantiations of the variables as found in the corpus are analyzed. The word-aligned forms in these corpus results are color-coded to reflect the types previously defined in the parameter file and allow a quick eye-ball estimation of pattern and possible problems.

Furthermore, word lists of equivalent word forms, together with their classification are aggregated in a list and inspected for error assessment. These two views allow a closer look at the data; they are important for the assessment of obvious patterns in the data and for error recognition. This perspective is used in an initial phase to review and refine the operationalization of the parameters.

In a second, quantitative perspective, the data is aggregated to reveal dominant and recurrent patterns. It is important to realize that the automated approach necessarily leads to errors; note that there will be errors in the word alignment process and recall the erroneous categorization of *гость* described in the last section. However, such errors are assumed to be essentially random, and therefore do not lead to recurrent patterns. Aggregation of many examples both presupposes automatic analysis and offsets its expected shortcomings.

For aggregation, all word aligned word forms in the corpus are classified in respect to the user defined operationalization in terms of word, lemma, tag, and other possible levels, just as it is done for the random samples. This evaluation is restricted to at most four tokens of each lemma to alleviate the influence of high frequency items.

The evaluation of the complete text results in a matrix of aligned word forms that are represented in a table like in figure 1. Each letter stands for a suffix class, while a hyphen means that the word

form in question is not a suffixed noun; cf. the classification in example (1) and (2). In other words, each parallel attestation is represented as a column in the matrix in figure 1.

```
'ru'  -AAAA-----UUUU-----IIII-----UKKK-----UUUUM-----S-----IIII-OOOO-----U-MMMc--
'by'  RAAAA-----U-UUCCU-----RRRR-----IIIIUUUU--UUU--R-----UUUUUU--KR-----IIII-KK-----U-----
'uk'  -AAAA-----UR-----U-----UUUU-----UU-U-UU-----U--Z--KK-----M--OII-I-I-----IU-MMM--
'uk/2' -AAAA-----U--UUUU--U-----UUUU--UUUU-----UU-ZMMKKK-----OI-II-III--IOU--U-MMMO--
'p1'  --AAA-----UUUUU--K-----U--U-----UU-SUIUUU-----K--II-----OO--I-----I--
'p1/2' --AAA-----UUUUU--O--O-----K--U--U--KO--UUZS--UU-----O-----NO-----U-----
'cz'  -----ZZZ-----UUUUUUU-----ZZZM-M--Z--NKKK--KC-----NLUU-U-----K--OIUUU--O--O--I--MM--
'sk'  -AAAA-----R-U-RN-UUUU-----R-----UU-----C-UUU-----UU-LUU--RR-R--SO-UU--C-O--OOU-I--MMMI--
's1'  -ACAA-----U-K-NA-----L--O--K---K---NC--CN-----K-ASSAA-----IIAI-INI-P--O--O-----RRR--
'hr'  -AAAA-----XNXN-----KK-----I-----U-M--Z-AA-----SI-III-PCPM-OOO-I--RRRR-O
'sr'  -A-AA-CC-PA--RA-----C-C-O--CCKK-K-----AA-A--U--XS-----AUA--I-I-I--MMOOO-O-----RRR--
'sr/2' -AA-A--C-AC--RA-----C-R-CC-KKK-----AA-A--U-M--UZU-----A-----IMIII-PMM-O--O-----RRRR--
'mk'  -AAA-----U-----X-----USU-U-----III-I-----OO-O--U-A-RRRR--
'bg'  -AAAA-----U-UUUU-----U-----UUUU-----CMOOOOO-----A-MMM--
```

Figure 1: Matrix representing all aligned word forms in the corpus with a suffixed noun in at least 1 version, based on the operationalization in Tab. 1.

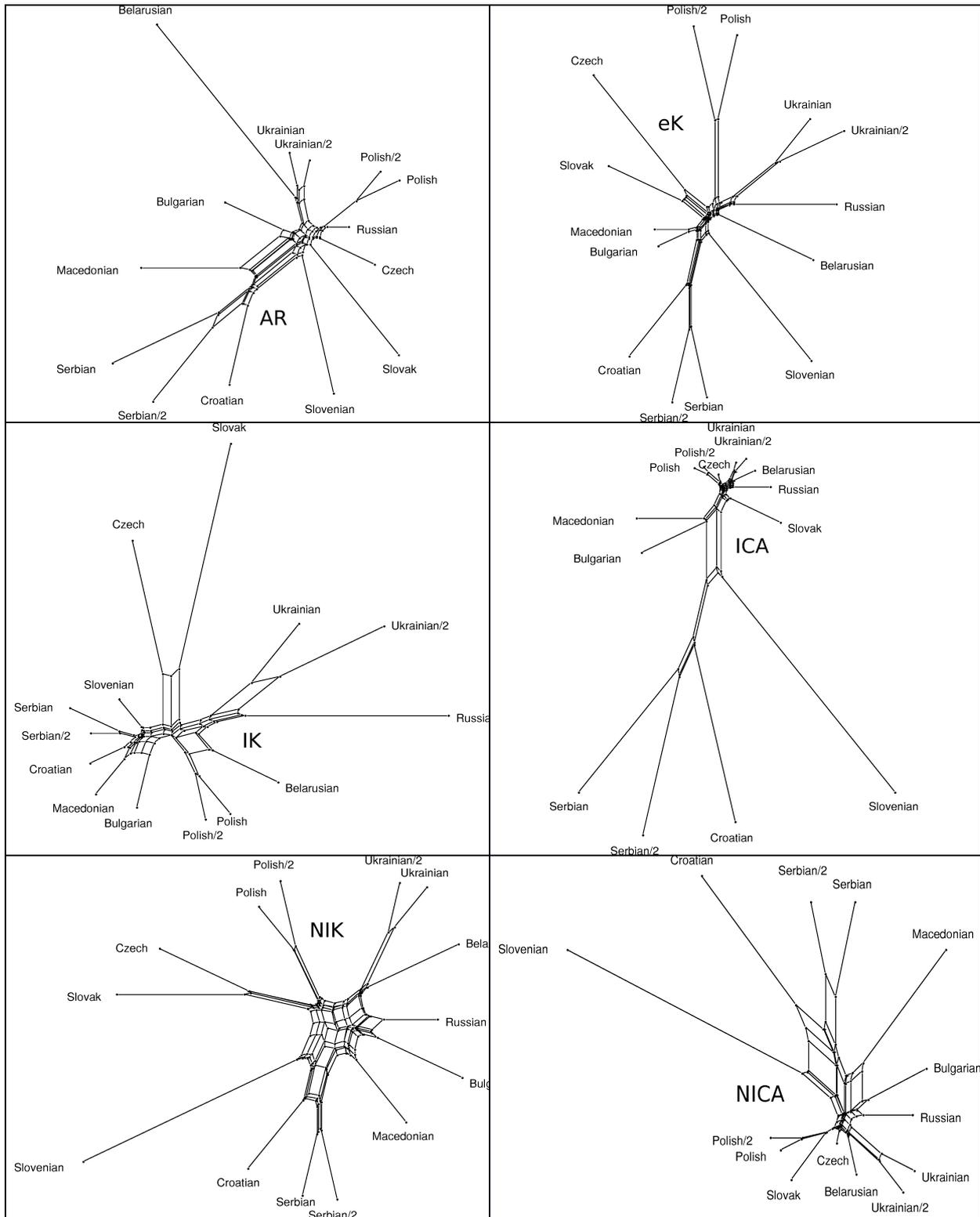
The data can then be aggregated and investigated from different perspectives. The main mechanism is a conversion of the matrix in figure 1 into a distance matrix using hamming distance (cf. Nichols/Warnow 2008). To do this, we compare each string and compute the amount of identical symbols in the same position. This gives us a measure of the similarity of the parallel texts in respect to the variable in question. For instance, if we find that out of ten examples the Polish and the Russian version use a different suffix class in five cases, we arrive at a dissimilarity or distance measure of .5. If two Polish translations of the same text (which are expected to coincide more often) contain different suffixes in, say, two cases, this would give us a distance value of .2.

In other words, the distance *dist* between two translations (the doculect distance) in terms of categorized word forms is

$$\text{dist} (ln1, ln2) = \text{count (different symbols in equivalent word forms in } ln1, ln2) / \text{count (number of equivalent word forms)}.$$

The set of word forms is kept constant across all languages, that is, it involves all word forms that either themselves have a suffix or have at least one equivalent word form with a suffix (in other words, all columns in the table in figure 1). This procedure is used in an analogous way for prefixes.

This computation is done for each pairwise combination of texts, yielding a matrix of distances between the doculects. This distance matrix is then visualized as a NeighborNet (Huson/Bryant (2006), a clustering algorithm that preserves much of the ambiguity typically found in this type of data. Depending on different filters and definitions of the features that are being compared, several perspectives on the data are possible. Figure 2.1 –2.12 give NeighborNets visualizing the similarity of the doculects in respect to each individual suffix type; figure 3.1 –3.10 give analogous graphs for the prefix types. For the purposes of the present paper, we will not go into details in respect to any single affix type; instead, in our analysis, we focus on similarities that set the nominal suffix classes apart from verbal prefix classes.



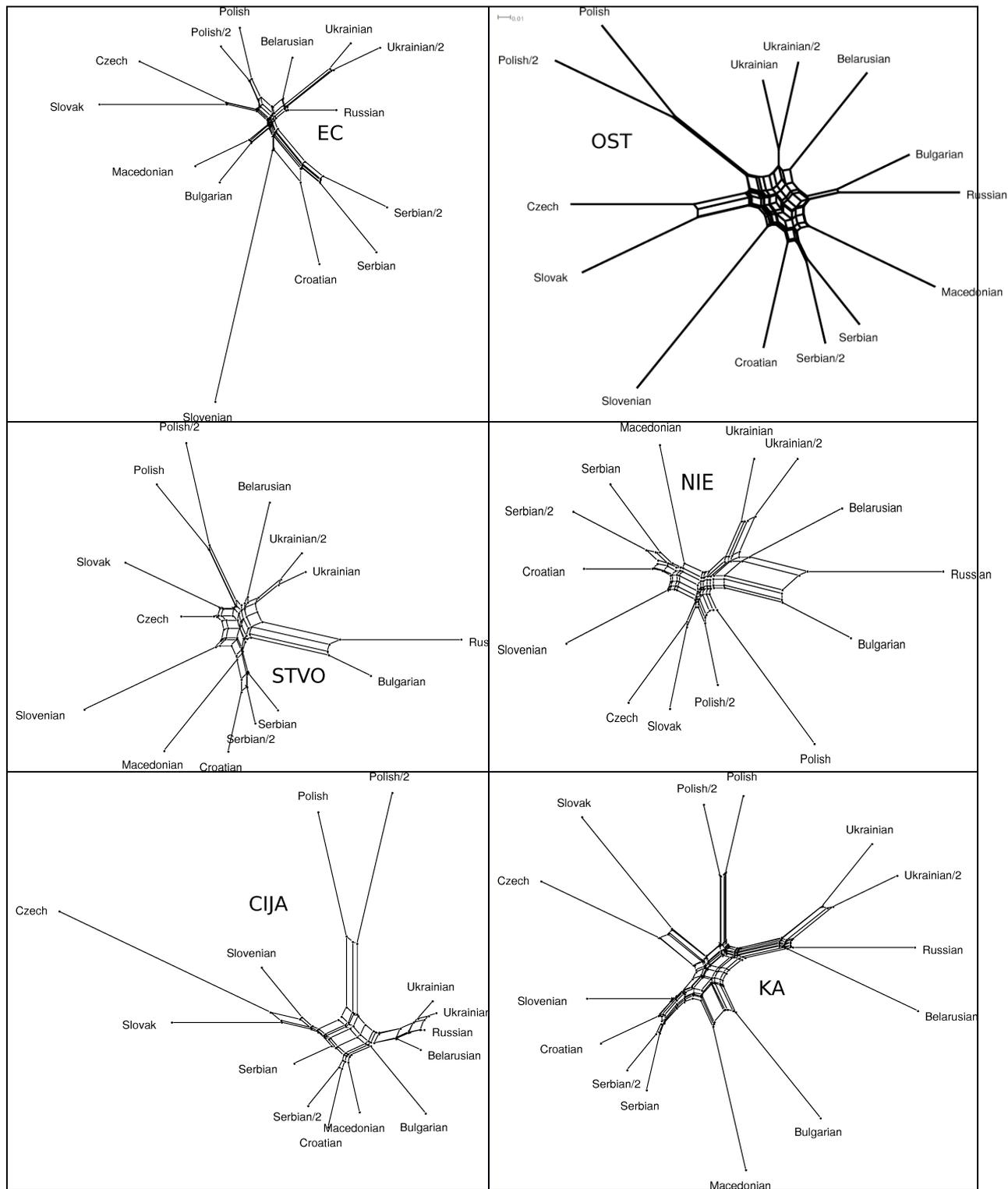
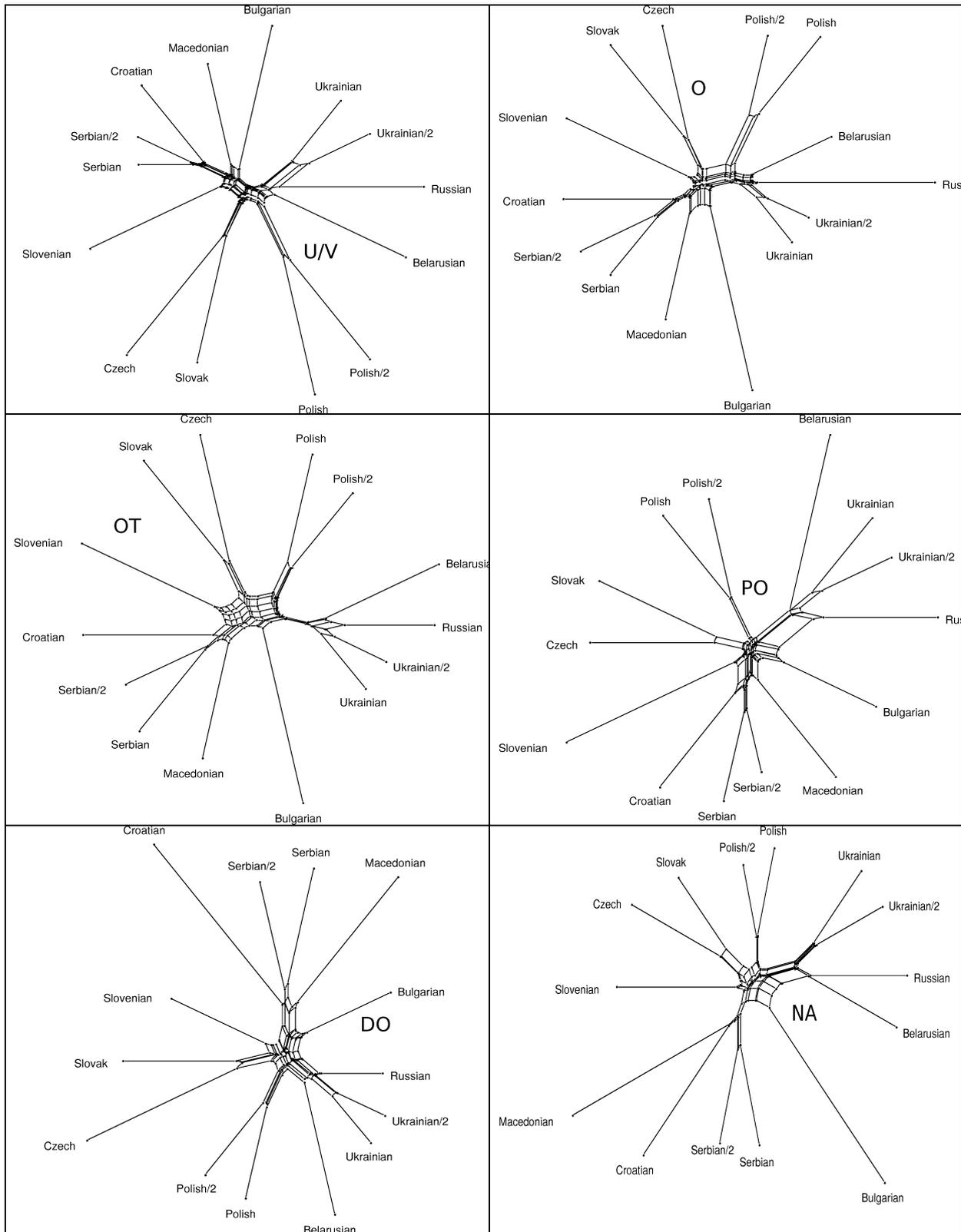


Figure 2.1.–2.12: NeighborNet graphs of suffix types. *Ukrainian/2*, *Polish/2* etc. refer to multiple translations into the same language



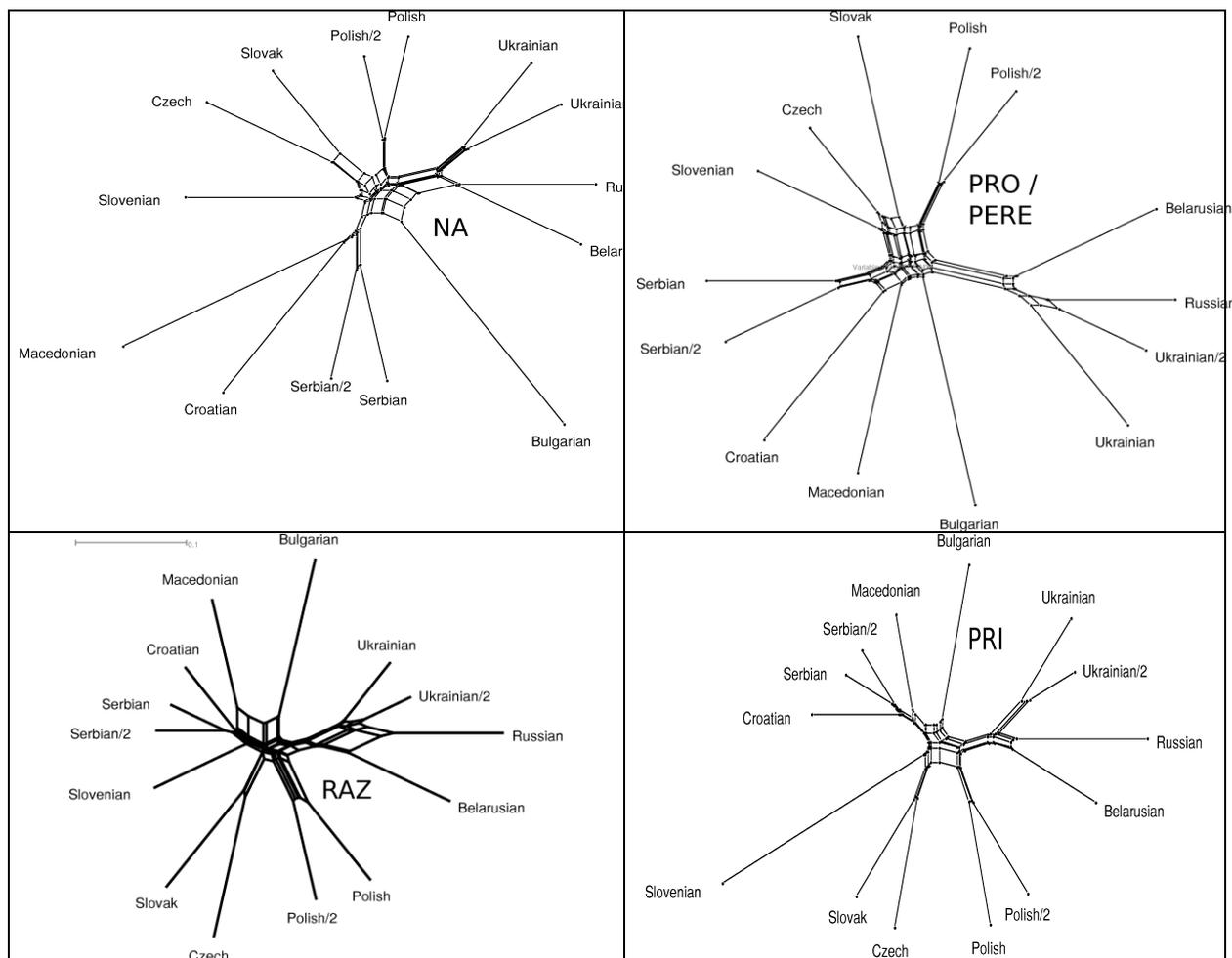


Figure 3.1–3.10: NeighborNet graphs of prefix types. Ukrainian/2, Polish/2 etc. refer to multiple translations into the same language

The graphs are mere visualizations of the distance matrices we have described above. Shortest paths in these graphs are proportional to the distances in the matrices. In addition to showing these distances, they also visualize similarity in distances: Not only do they show that, say, two translations into Polish are similar in that the distance between them is small, but that they are also similar in respect to the distances they have to other languages, say, Russian or Czech. In this way, we may find that the distance between some doculects is substantial, but that they are grouped together, because they are alike in respect to distances to other languages.

3. Data Analysis

3.1 Diagnostic Questions

For the purposes of this article, I focus on the following question: In what respect do the graphs show unexpected results in respect to genetic affiliations, and may these results be due to language contact?

I evaluate the graphs in respect to a naive hypothesis based on the received view of the genetic make-up of Slavic. This hypothesis expects East, West, and South Slavic to be coherent and distinct branches of the family, with subdivisions into Lechitic, Sorbian and Czecho-Slovak in West Slavic, and into Slovenian/BCS and Macedonian/Bulgarian in South Slavic. This is summarized in the tree in

figure 4 that represents the received genealogy as found in Sussex/Cubberley (2006). The network in figure 5, in contrast, is based on empirical data concerning the retention of common Slavic vocabulary in different Slavic languages; it is a visualization of distance data computed by Žuravlev (1994: 129–133) on the basis of the Etymological Dictionary of Slavic Languages (ÉSSJa 1974–). The two graphs show rather far-reaching, but not complete, compatibility and thus constitute a kind of baseline for the analysis below.

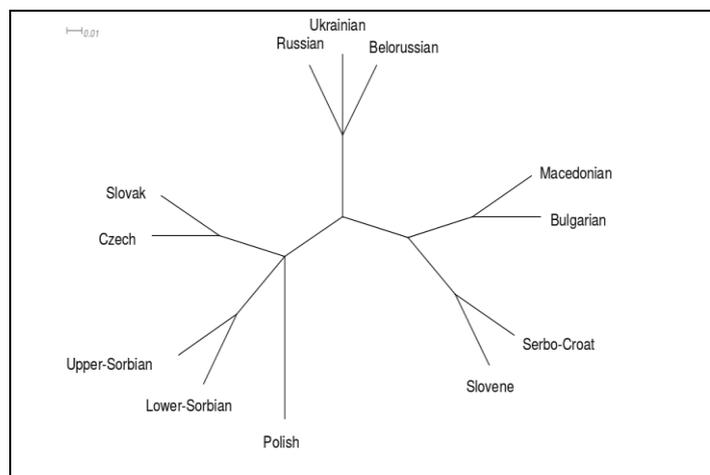


Figure 4: A tree-like representation of the Slavic genus following Cubberley/Sussex (2006)

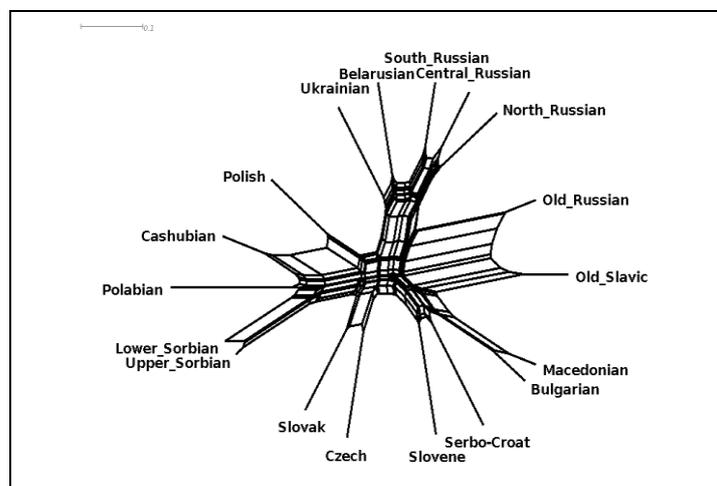


Figure 5: NeighborNet representation of the genetic make-up of Slavic based on the shared retention of common Slavic vocabulary across Slavic (visualization of data in Žuravlev 1994:129–133)

The overarching question is operationalized as a series of specific subquestions with increasing granularity. The first issue is whether the main branches of Slavic are represented in coherent groups, i.e. whether its members are more similar to each other than to members of a different branch. Then, specific issues are addressed that concern individual languages that cross the boundaries between these groups and or that may not pattern with their nearest genetic siblings according to the received interpretation.

Thus, in respect to each affix, I ask the questions:

- Is East Slavic a coherent, separate branch?
- Is there an East Slavic language which does not cluster with the other two, and what does it group with?
- Is West Slavic a coherent, separate branch?
- Does Polish pattern with Czech/Slovak or rather with East Slavic/Russian?
- Is South Slavic a coherent, separate branch?
- Does Slovene group with BCS or with Czech/Slovak?
- Does Macedonian group with BCS or with Bulgarian?
- Do Bulgarian and Macedonian form a subdivision of South Slavic, that is, an internal branch?
- Does Bulgarian group with South Slavic or with East Slavic (Russian)?

3.2 Data Evaluation

Tab. 3 presents the answers to the lead questions in table format, with answers that violate the expectations based on the received view highlighted in bold.

Tab. 3: Evaluation of neighbor net graphs based on a list of diagnostic questions. For each branch of Slavic, the lead question is whether the branch is coherent (coher.) or fragmented (fragm.) and is followed by a number of more specific questions.

| | EsI | Which language separate | WSI | pl with EsI (ru)? | SSI | sl w. bcs or w. cz/sk? | mk w. bg or w. bcs? | do bg and mk form a branch? | is bg w. SSI or w. EsI? |
|----------|--------|-------------------------|----------------------|--------------------|--------------|------------------------|---------------------|-----------------------------|-------------------------|
| UV | coher. | | fragm. | between | coher. | between | w. bg | branch | w. SSI |
| ZA | coher. | | fragm. | between | fragm. | w. cz/sk | w. bcs | no | between |
| DO | coher. | | rath. coher. (w. sl) | no | fragm. | w. cz/sk | w. bcs | no | between |
| NA | coher. | | coher. (w. sl) | no | fragm. | w. cz/sk | w. bcs | no | rath. EsI |
| OT | coher. | | fragm. | between | fragm. | between | w. bcs | no | between |
| O | coher. | | fragm. | towards EsI | fragm. | between | between | partly | rath. SSI |
| PO | coher. | | coher. (w. sl) | no | fragm. | rath. w. bcs | w. bcs | no | rath. EsI |
| PRI | coher. | | fragm. | between | fragm. | w. cz/sk | w. bcs | no | rath. SSI |
| PRO-PERE | coher. | | rath. coher. | no | rath. fragm. | w. cz/sk | between | no | rath. SSI |
| RAZ | coher. | | coher. | no | rath. fragm. | w. bcs | rath. Bcs | partly | between |
| | | | | | | | | | |
| AR | fragm. | ru (w. pl) | fragm. | w. ru | fragm. | rath. w. bcs | w. bcs | | w. EsI |
| CIJA | coher. | | fragm. | between | fragm. | w. cz/sk | w. bcs | | w. EsI |
| EC | coher. | | fragm. | between | coher. | | w. bg | branch | w. SSI |
| eK | fragm. | by | fragm. | between | coher. | | w. bg | branch | w. SSI |

| | | | | | | | | | |
|------|-------------------|---------------|-----------------------------------|---------|----------------------------|-------------------------|-----------------|--------|-----------------------|
| | | (w. pl) | | | | | | | |
| ICA | coher. (infr.) | | fragm. (infr.) | | coher. | | w. bg | branch | w. SSl |
| IK | rath. coher. | by (w. pl) | coher. (if by is re- moved) | | coher. | | w. bcs | | alone |
| KA | coher. | | fragm. | between | rath. coher. | w. bcs | w. bg | branch | w. SSl |
| NICA | fragm. | ru (w. bg) | coher. (infr.) | | rath. coher. (w. ru) | w. bcs | between | | towards ru |
| NIE | fragm. | ru (w. bg) | coher. | | fragm. (w. ru) | w. bcs | w. bcs | | w. ru |
| NIK | fragm. | ru (w. bg) | coher. | | rath. coher. (w. ru) | w. bcs | between | | w. ru |
| OST | fragm. | ru (w. bg) | rath. coher. | | rath. coher. (w. ru) | w. bcs | rath. w. bcs | | w. ru |
| STVO | fragm. | ru (w. bg) | fragm. | between | fragm. | rath. w. bcs | w. bcs | | w. ru |

I will summarize these answers and provide commentary for them in the remainder of this section, discussing each branch in separate, as far as this is practical.

The first question concerns the coherence of East Slavic. Across all the verbal prefix classes, East Slavic is a separate, well distinguished branch in the graph. With suffixes, we see a quite different picture; East Slavic is often fragmented. In some cases, this is due to systematic errors: In the case of AR, the reason for a fragmented East Slavic group is mostly the conflation of unstressed etymological /o/ and /a/ in Belarusian (e.g. *дырэктар* ‘director’) and no palatalization of final /r/ in Ukrainian (in Ukrainian, both *мова* and *зосудар* fall into this group, while Russian has *мова*, but *зосударь*). With most suffixes, however, no such problems are apparent. With eK, Polish moves very much into the East Slavic group, while with IK, there is a strong attraction between Belarusian and Polish that crosses the boundaries between East and West. The strongest effect, however, is evident with NIE, NIK, OST and STVO (and to a lesser degree with NICA), where Russian moves out of the East Slavic cluster to form a separate group with Bulgarian. These suffixes are to a certain extent characteristic of a higher register and rather prone to be influenced both by Church Slavonic in Russian and by standard Russian in Bulgarian. We thus see a much larger influence of direct or indirect contact between Russian and Bulgarian with the suffixes than with the prefixes.

As to West Slavic, we find that in general, this group is less coherent than East Slavic with the Polish and the Czech and Slovak translations being almost always rather far from each other. This fits well with the observation that the split between Lechitic and Czech/Slovak is much earlier than that between the East Slavic languages; note that Žuravlev’s data concerning the inherited lexicon as visualized in figure 5 also involve a rather large distance between Polish and Czech/Slovak. Polish is to different degrees attracted by Russian or other varieties of East Slavic, with which it shares a long

contact history and an essentially open border uninhibited by natural boundaries. This tendency of Polish to be similar to East Slavic holds across both prefixes and suffixes, and contact induced and inherited similarity seems to be difficult to distinguish in this case.

South Slavic is likewise little coherent, which is generally to be expected from the literature. However, our findings do not primarily involve a division between BCS and Slovenian on the one hand, and Macedonian and Bulgarian, on the other, although this is the split most often adduced for South Slavic, most prominently in Mareš's (1980) tetrachotomy of Slavic. Only little evidence for such a contrast is found among the verbal prefixes (perhaps most clearly with UV, but not very strong even in this case). Among the suffixes, a split where we find Macedonian and Bulgarian on the one hand, and Slovenian and BCS on the other is present only with EC, KA and ICA. With eK, we find that Macedonian and Bulgarian pattern together, but Slovenian and BCS do not.

A more significant split that does not figure as prominently in the literature concerns the position of Slovenian in respect to Czech/Slovak or BCS. Across the verbal prefixes, Slovene only forms a common group with BCS in respect to RAZ, and to some extent in respect to PO. With U/V, OT and O, Slovenian is situated between Czech/Slovak and BCS. In most cases, however, namely with ZA, DO, NA, PRI and PRO/PERE, Slovenian joins the West Slavic branch, and specifically Czech/Slovak. This grouping is posited sometimes in the literature and there are several well-known isoglosses that unite Slovene with West Slavic, e.g., the formation of the future tense with an auxiliary derived from 'be', rather than 'want', as in the rest of South Slavic (see, e.g., Bernštejn 1961: 78–79 for a general assessment of this isogloss region as one that „in respect to its numbers of features does not have a rival“ [my transl.]). However, this close proximity is still unexpected.

The verbal prefix data and the nominal suffix data differ in this respect. With the suffixes, a Czech-Slovak-Slovenian group is found clearly only for the suffix CIJA, and to some extent with AR and STVO. With EC, eK, ICA and IK, Slovenian is grouped inside of South Slavic, but separate from BCS; however, most frequently, namely with KA, NICA, NIE, NIK and OST, Slovenian clearly patterns with BCS. The situation is thus reverse of what we see with the verbal prefixes, where the tendency is clearly for Slovene to pattern with West Slavic.

Macedonian patterns with BCS for most of the verbal prefixes (ZA, DO, NA, OT, PO, PRI); the exception is with U/V, where Macedonian and Bulgarian form the expected branch inside of South Slavic. With nominal suffixes, in turn, Macedonian sometimes patterns with BG (EC, eK, ICA, KA), sometimes with BCS (AR, CIJA, NIE and to some extent OST), and sometimes falls in-between (NIK and NICA). The main factor here seems to be the attraction of Bulgarian by Russian, which does not in general extend to Macedonian: If Bulgarian moves into a group with Russian, it is separated from Macedonian, which then has a tendency to be part of the BCS-cluster.

We see such an attraction of Bulgarian and Russian only with two of the verbal prefixes, and only as a weak tendency. With nominal suffixes, however, this attraction is striking: Bulgarian and Russian form an exclusive group in a majority of cases (namely NIE, NIK, OST and STVO), and Bulgarian

shows a tendency towards Russian or East Slavic in three more cases (AR, CIJA, NICA). Only with EC, eK, ICA and KA is Bulgarian unambiguously part of the South Slavic group. In these cases, it forms an exclusive branch of South Slavic together with Macedonian.

In other words, we tend to see two patterns with the nominal suffixes concerning East South Slavic: First, a pattern where Bulgarian and Macedonian form a group inside of South Slavic, in many cases with an opposing group composed of BCS and Slovenian; second, a pattern where Macedonian is very similar to BCS and in many cases Slovenian, and Bulgarian groups with Russian.

3.3. Summary and Discussion of the Data

The above evaluation of the graphs along the lines of the lead questions allows a number of interesting observations. While a full explanation of these findings is difficult in view of the restricted data basis of a single text, some observations can be made and tentative explanations can be given already at this point.

On the most general level, it clearly emerges that, notwithstanding the diversity within each of the variables, verbal prefixes as opposed to nominal suffixes behave differently as groups – there are clearly configurations that are restricted to, or much more prominent in, one of the two groups. This is interesting because it shows that investigating change in respect to different subsystems of derivational morphology, or more generally, in respect to different parts of speech, may be important for our understanding of language change in Slavic (and beyond) in general.

The results fit well to the well-established fact that nouns are more easily borrowed than verbs (Tadmor 2009: 61–63), thus leading us to expect that at least part of the difference we see between the two derivational types is due to a different general propensity to be changed by language contact. Note that this would show that such a generalization also holds for a closely related genus such as Slavic where questions of morphological complexity, i. e. grammatical factors, provide only a weak explanation (cf. Tadmor 2009: 61 for this hypothesis).

How exactly do the two derivational domains differ, and what are the most significant patterns? Two cases seem to be particularly relevant.

First, we note the clear coherence of East Slavic across all verbal prefixes, and its fragmentation, in contrast, with a number of nominal suffixes where Russian patterns with Bulgarian. This is most clearly the case with NIE, NIK, OST and STVO (it is less clear with NICA, where the type frequency is in addition lower and the error margin thus higher). NIE, OST and STVO involve abstract vocabulary, and all suffixes are part of the lexical layer that is strongly influenced by Church Slavonic in Russian. Bulgarian in turn has more direct historical links to Old Church Slavonic, and was in addition strongly influenced by Standard Russian in the 19th century. For this reason, this similarity can be readily explained by language contact: between Church Slavonic and Russian, on the one hand, and between Russian and Bulgarian, on the other (with a more complicated, but close relationship

between Bulgarian and Church Slavonic forming the background for this). Interestingly, no such influence can be observed with the prefixes, which show much less impact of this contact history.

The case of Russian and Bulgarian forming a single group is significant, because it is clearly not due to genetic association, i. e., the border between the traditional branches of Slavic are crossed without a strong suggestion that the border itself may be open to question.

This is different in the second case where a language has changing affiliations to branches of Slavic depending on the variable: The South Slavic language Slovenian is grouped with Czech/Slovak rather than with BCS in most of the prefixes (ZA, DO, NA, PRI, PRO/PERE), and clearly grouped with BCS only in one case (RAZ); the other prefixes fall between. With suffixes, this relationship is reverse: Slovenian patterns with Czech only in one case (CIJA), while in the other cases it is more or less clearly associated with BCS or at least falls into the South Slavic branch.

This case is different from the one above, because it is well known that there are deep early links between Slovenian and Czech-Slovak stemming before the end of the first millenium when the land link between the two areas where interrupted. In addition, Czech, as the main Slavic language of the Austrian empire had a strong influence on Slovene in the 19th century when modern standard Slovene developed.

If we subscribe to the assumption that verbal prefixes are more conservative, the contrast can be readily explained: The prefix system or, by extension, the verb system has preserved much of the similarity that Slovene had with West Slavic to begin with. The association of Slovene with BCS, or South Slavic in general, in respect to the suffixes would then be an outcome of subsequent language contact, along the same lines that language contact is the driving factor for the similarity of Bulgarian and Russian. This fits well to the observation that CIJA is the only suffix where Slovene patterns with Czech/Slovak – this could be explained, by the same token, with the hypothesis that many modern terms ending in CIJA may have entered the language during the 19th century when Czech was an influential contact language for the written medium.

Note that while such a view of Slovenian as consisting of an older layer with West Slavic affiliations, and a newer layer that is due to strong contact with South Slavic dialects fits well to other observations (say, the reflex of the Jers or the future tense formation), it is not in line with Žuravlev's data that is based on shared retentions of Common Slavic vocabulary. While it could be argued that contact influence of Štokavian dialects may have had an impact on these retention rates, more evidence for West Slavic retention patterns would be expected. However, this issue requires a more detailed research.

A final point to be discussed is the tendency of Macedonian to be closely grouped with BCS with verbal prefixes, and much less so for nominal suffixes. If we take the above assumption for granted that verbal prefixes represent an earlier layer, it follows that Macedonian is genetically closer linked to BCS than normally assumed, but has been under contact influence of Bulgarian. This is a problematic view, however, and other factors, such as the tendency to coin norms that differed from both BCS and

Bulgarian during the standardization of Macedonian would clearly need to be taken into account for a complete picture.

4. Summary and conclusions

The present paper has employed a corpus driven method for the comparison of the use of derivational morphology across many Slavic translations of a single text. Based on the evaluations of NeighborNet graphs for 10 verbal prefixes and 12 nominal suffixes, it was shown that the domains differ considerably in the way Slavic languages are grouped within these two domains. The relations within the domain of verbal prefixes arguably draws a more conservative picture of the genetic relations, while the nominal suffixes show more evidence for contact influence and generally seem to be more volatile. Most strikingly this is the case with Russian and Bulgarian, which clearly belong to different branches in respect to the use of verbal prefixes, but form a common branch for many nominal suffixes. Somewhat simplifying, and put drastically, the data suggest that in respect to verbal prefixes, Slovenian patterns look like the ones in a West Slavic language, while in respect to nominal suffixes, it is clearly a South Slavic language. While this is in line with earlier and well known observations about isoglosses that group Slovenian with West Slavic, and about the contact influence of Church Slavonic on Russian, the results of this study show that these conflicting affiliations affect different subsystems of derivational morphology and indeed the lexicon in quite different ways.

Acknowledgement: Funding by the Swiss National Science Foundation is gratefully acknowledged.

5. References

- BERNŠTEJN, SAMUIL B. (1961): *Očerk sravnitel'noj grammatiki slavjansikich jazykov*. Moskva: Izdatel'stvo Akademii Nauk SSSR.
- ĚSSJA (1974–): *Ětimologičeskij slovar' slavjanskich jazykov*. Moskva: Nauka.
- HASPELMATH, MARTIN; SIMS, ANDREA (2008): *Understanding Morphology*. Second edition. London: Arnold.
- HUSON, DANIEL. H.; BRYANT, DAVID (2006): Application of phylogenetic networks in evolutionary studies. In: *Molecular Biology and Evolution*, 23, 254–267.
- KOPEČNÝ, FRANTIŠEK (1973): *Etymologický slovník slovanských jazyku. Slova gramatická a zájmena. Svazek 1: Předložky, koncové partikule*. Praha: Academia.
- MAREŠ, FRANTIŠEK (1980): Die Tetrachotomie und doppelte Dichotomie der slavischen Sprachen. In: *Wiener Slavistisches Jahrbuch* 26, 33–45.
- NICHOLS, JOHANNA; WARNOW, TANDY (2008): Tutorial on Computational Linguistic Phylogeny. In: *Language and Linguistics Compass* 2, 760–820.
- SAHLGREN, MAGNUS (2008): *The Distributional Hypothesis*. In: *Italian Journal of Linguistics/Rivista di Linguistica*, 20, 33–53.
- SUSSEX, ROLAND; PAUL CUBBERLEY (2006): *The Slavic languages*. Cambridge: Cambridge University Press.
- TADMOR, URI (2009): Loanwords in the world's languages: Findings and results. In: Haspelmath, Martin; Tadmor, Uri (eds.): *Loanwords in the world's languages. A comparative handbook*. Berlin: de Gruyter, 55–75.

- TIEDEMANN, JÖRG (2003): Combining Clues for Word Alignment. In: *Proceedings of the 10th Conference of the European Chapter of the ACL (EACL03)* Budapest, Hungary, April 12–17, 339–346.
- VAILLANT, ANDRÉ (1974): *Grammaire comparée des langues slaves. Tome IV, La Formation des noms*. Paris: Klincksieck.
- VON WALDENFELS, RUPRECHT (2011): Recent Developments in ParaSol: Breadth for Depth and XSLT based web concordancing with CWB. In: Majchráková, Daniela; Garabík, Radovan (eds.): *Natural Language Processing, Multilinguality. Proceedings of Slovko 2011*, Bratislava: Tribun, 156–162.
- VON WALDENFELS, RUPRECHT (2015): The ParaViz Tool: Exploring Cross-linguistic Differences in Functional Domains Based on a Parallel corpus. In: Grigonyté, Gintaré; Clematide, Simon; Utká, Andrius; Volk, Martin (eds.): *Proceedings of the Workshop on Innovative Corpus Query and Visualization Tools at NODALIDA 2015, May 11–13, 2015, Vilnius, Lithuania*. Linköping University Electronic Press, Linköpings universitet, 32–36.
- VAN ZUNDERT, JORIS ET AL. (2012): Cultures of Formalization – Towards an encounter between humanities and computing. In: Berry, David M. (ed.): *Understanding Digital Humanities*. London: Palgrave Macmillan, 279–294.
- ŽURAVLEV, ANATOLIJ F. (1994): *Leksikostatističeskoe modelirovanie sistemy slavjanskogo jazykovogo rodstva*. Moskva: Indrik.