



universität
wien

Emmerich Kelih

Institut für Slawistik

Lehnwörter im Slawischen: Empirische und crosslinguistische
Perspektiven
(Fallbeispiel: Slowenisch)

Workshop
Wien, 25. bis 26. September 2014

Überblick

- Allgemeines aus der der Lehnwortforschung
- Problemstellung – Fallbeispiel Slowenisch
- Einbettung: „Loanwords in the world's languages“ (Haspelmath/Tadmor 2009)
- ... viele offene Fragen u.a.
 - a.) Auswahl der Wortlisten
 - b.) Zusammensetzung der Wortarten
 - c.) Identifikation von Lehnwörtern
- Zusammenfassung & Perspektiven
- Ein fernes Ziel (1): Ein quantitatives Lehnwortprofil
- Ein fernes Ziel (2): Von der Deskription zum Vergleich (Methoden?)

Lehnwörter/Entlehnungen in den Sprachen der Welt

- Wortformen (Lexeme, Teile von Wörtern, Morpheme u.a.) werden von einer Sprache entlehnt/übernommen
- keine „Entlehnung“ im eigentlichen Sinne, sondern eine Inkorporation/Adaption
- terminologische Vielfalt, z.B. im Deutschen bzw. in der Germanistik: “äußeres” Lehnwort: direkte Entlehnungen (Fremdwort, Lehnwort), Scheinentlehnungen (lexikalisch, semantisch), “inneres” Lehnwort: Lehnbedeutung, Lehnübersetzung, Lehnübertragung etc.)
- unterschiedliche Relevanz und Wichtigkeit in unterschiedlichen linguistischen Traditionen
- Entlehnungen: generell eine der zentralen Möglichkeiten der Vergrößerung/Bereicherung eines Wortschatzes einer Sprache
- Sprache als ein offenes, dynamisches System mit einer Vielzahl von Systembedürfnissen

Fallbeispiel Slowenisch (ein kurzer Überblick):

- ungebrochenes Interesse der slowenistischen Sprachwissenschaft an LW seit den Anfängen
- Vielzahl von Studien und Aspekten: slawisch – nicht-slawisch, innerslawische Kontakt
- Entlehnungen im weitesten Sinne und deren historische Genese

Fallbeispiel: Slowenisch

Miklosich, F. (1867): *Die Fremdwörter in den slavischen Sprachen*. Wien.

Schuchardt, H. (1884): *Dem Herrn Franz von Miklosich zum 20. Nov. 1883. Slawo-deutsches und Slawo-italienisches*. Graz.

Breznik, Anton (1909): *Slovanske besede v slovenščini*. Ljubljana.

Striedter-Temps, H. (1963): *Deutsche Lehnwörter im Slovenischen*. Wiesbaden.

Jakopin, F. (1968): O deležu ruskih elementov v razvoju slovenskega knjižnega jezika, *Slavistična revija*, 16, 65-90.

Lägreid, A. (1973): *Die russischen Lehnwörter im Slovenischen*. [...] München.

Thomas, G. (1997): The Role of German Loanwords in the Slavic Languages of the Former Habsburger Empire, in: *Canadian Slavonic Papers*, 39,3, 333-359.

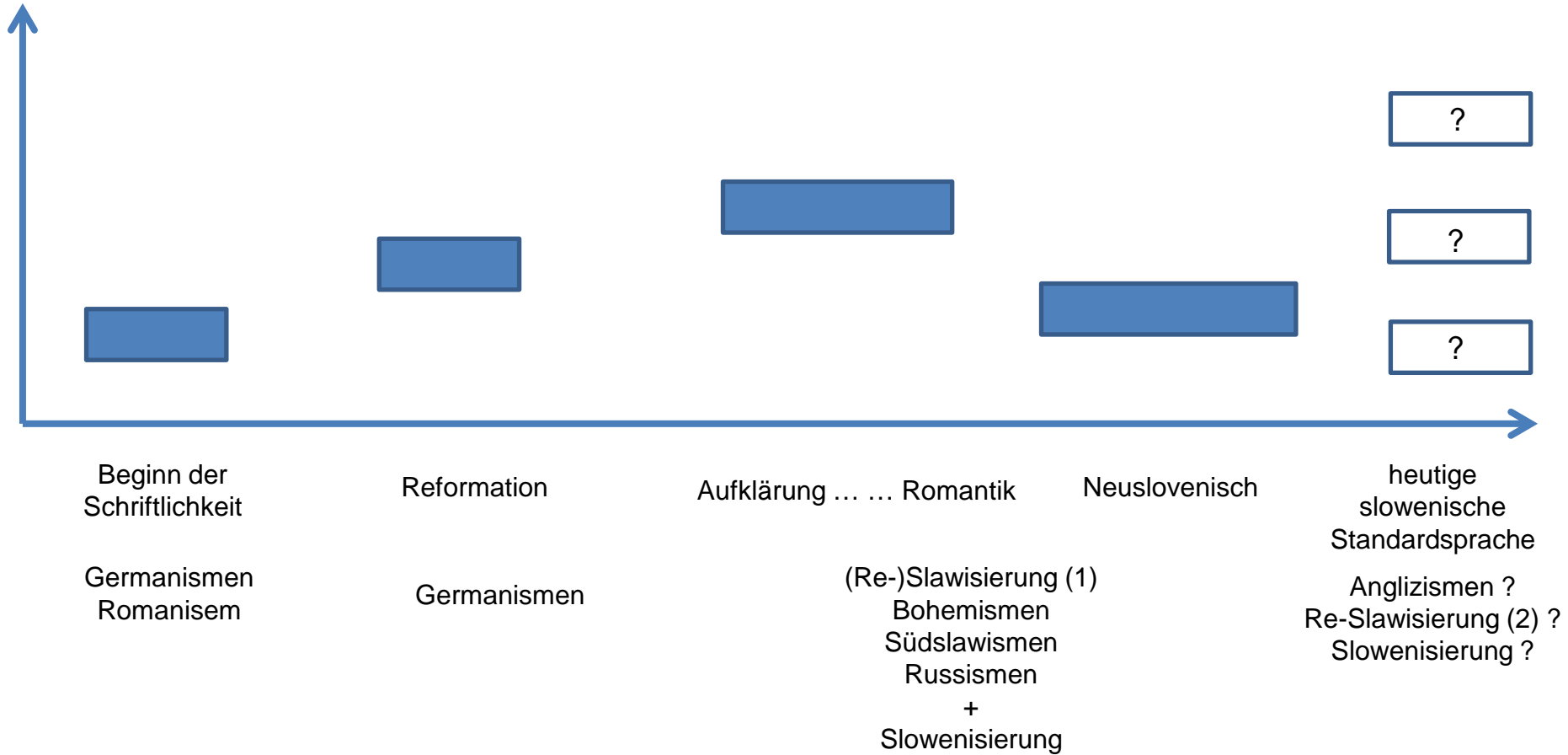
Snoj, M. (2006): O tujkah in izposojenkah v slovenskem jeziku, in: *Slavistična revija* 54, 343–349.

Jelovšek, A. (2009): Srbohrvatizmi nekoč in danes: Analiza tiska med letoma 1945 in 2005. Ljubljana.

... ..

... ..

Versuch einer „Modellierung“: Zunahme/Abnahme von Lehnwörter in der slowenischen Standardsprache





Wie kann man die Frage nach einem bestimmten Lehnwortanteil systematisch, empirisch begründet untersuchen?

reduktionistischer Ansatz:

- Aussagen über ein Sprachsystem als Ganzes nur eingeschränkt möglich
- Reduktion auf einzelne Register/Funktionalstile
- Reduktion auf einzelne Autoren
- Reduktion der Untersuchungsbasis auf Basiswortschatz/Grundwortschatz

→ besondere **Resistenz** des Basiswortschatzes gegenüber Entlehnungen! (vgl. Hock/Joseph (1996: 257), Weinreich (1964) etc.)

Inspirationen aus der allgemeinen bzw. typologischen Sprachwissenschaft

- neuer Input und Ansatz durch das Projekt von Haspelmath/Tadmor (2009)
- Basisdaten über die Frequenz von Lehnwörtern/Entlehnungen



The World Loanword Database (WOLD)

It provides vocabularies (mini-dictionaries of about 1000-2000 entries) of 41 languages from around the world, with comprehensive information about the loanword status of each word. It allows users to find loanwords, source words and donor languages in each of the 41 languages, but also makes it easy to compare loanwords across languages. Each vocabulary was contributed by an expert on the language and its history. The database can be accessed by language, by meaning or by author.

The list of 1460 meanings on which the vocabularies are based is called the Loanword Typology meaning list, and it is in turn based on the list of the Intercontinental Dictionary Series.

The World Loanword Database is the result of a collaborative project coordinated by Uri Tadmor and Martin Haspelmath between 2004 and 2008, called the Loanword Typology Project (LWT).

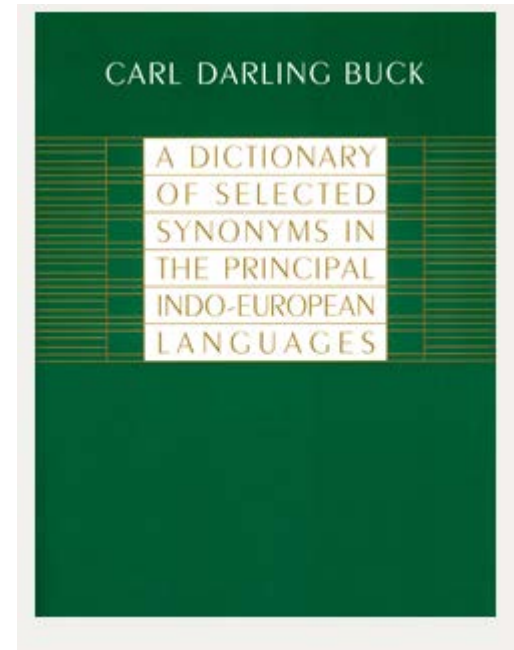
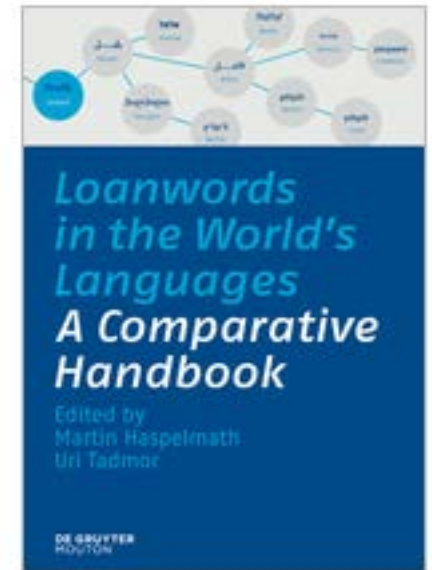
The World Loanword Database (WOLD)

Liste von ca. 1500 Bedeutungen (auf Englisch)

Quelle und ursprüngliche Liste:

Buck, Carl Darling (1949): *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago.

- spezielle Fokus: Untersuchung des Basis-Vokabulars
- Gruppierung des Materials in 22 lexikalische/semantische Gruppen
- mittlerweile Wortlisten, inkl. Metainformationen zu den einzelnen Lehnwörtern für ca. 40 Sprachen vorhanden



einige Details:

| Nr. | lexico-semantical group | Nr. of meangings |
|-----|--------------------------------|------------------|
| 1 | The physical world | 75 |
| 2 | Kinship | 82 |
| 3 | Animals | 96 |
| 4 | The body | 157 |
| 5 | Food and drink | 81 |
| 6 | Clothing and grooming | 59 |
| 7 | The house | 48 |
| 8 | Agriculture and vegetation | 74 |
| 9 | Basic actions and technology | 78 |
| 10 | Motion | 82 |
| 11 | Possession | 46 |
| 12 | Spatial Relation | 75 |
| 13 | Quantity | 38 |
| 14 | Time | 57 |
| 15 | Sense perception | 49 |
| 16 | Emotions and values | 48 |
| 17 | Cognition | 51 |
| 18 | Speech and Language | 41 |
| 19 | Social and political relations | 36 |
| 20 | Warfare and hunting | 40 |
| 21 | Law | 26 |
| 22 | Religion and belief | 26 |
| Sum | | 1460 |

Untersuchte Sprachen:

Swahili, Iraqw, Gawwada, Hausa, Kanuri, Tarifiyt Berber, Seychelles Creole, Romanian, Selice Romani, Lower Sorbian, Old High German, Dutch, English, Kildin Saami, Bezhta, Archi, Manange, Ket, Sakha, Oroquen, Japanese, Mandarin Chinese, Thai, Vietnamese, White Hmong, Ceq Wong, Indonesian, Malagasy, Takia, Hawaiian, Gurindji, Yaqui, Zinacantán Tzotzil, Q'eqchi', Otomi, Saramaccan, Imbabura Quechua, Kali'na, Hup, Wichí, Mapudungun

(!) außer Niedersorbisch keine slawischen Sprachen vertreten

EXTRA-Kategorie – eingeführt durch Haspelmath/Tadmor (2009)

- moderne Welt
- diverse Funktionswörter

Welche Formen von Meta-information?

| | Upper-Sorbian | | |
|--------------------|---------------|------------------------------|--------------------------------------|
| meaning | word form | borrowed status | source words |
| world | swět | 5. no evidence for borrowing | land 'land, country' New High German |
| land | land | 1. clearly borrowed | kraj 'country, land' Upper Sorbian |
| land | kraj | 2. probably borrowed | |
| soil | zemja | 5. no evidence for borrowing | |
| dust | proch | 5. no evidence for borrowing | |
| mud | kuř | 5. no evidence for borrowing | |
| sand | pěsk | 5. no evidence for borrowing | |
| mountain or hill | góra | 5. no evidence for borrowing | |
| cliff or precipice | skała | 5. no evidence for borrowing | |
| plain, | rownina | 5. no evidence for borrowing | |
| field | plonina | 5. no evidence for borrowing | |
| valley | doł | 5. no evidence for borrowing | |
| island | kupa | 5. no evidence for borrowing | |
| ... | ... | ... | ... |

1. absolute Frequenz von Lehnwörter innerhalb einer lexiko-semantischen
2. Gesamtanzahl von Lehnwörtern pro Sprache innerhalb des Basiswortschatzes

→ **Basis für ein quantitatives Lehnwortprofil**

Möglichkeiten der „inneren“ Differenzierung

„However, different languages display a remarkable degree of consistency which 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. (Tadmor 2009: 64)

häufig LW in:

1. religion and belief
2. clothing and grooming
3. house/living

niedrige Frequenz von LW in

1. the body
2. spatial relation
3. sense perception
4. function words with deictic function

- Probleme der Universalienforschung
- generelle Tendenzen?
- Wie passen die slawischen Sprachen in diese Tendenz?



→ Einige Anmerkungen zur Auswahl der Wortlisten !

Alternative (1): Swadesh-Liste(n)

Morris Swadesh (1909-1967) Begründer der Lexikostatistik, Glottochronologie,
Idee: Basisliste von Wortformen für Sprachvergleich, unterschiedliche Swadesh-Listen im Umlauf,
Reduktion von ursprünglich 200 auf 100 Wortformen,



<http://rosettaproject.org/>

für viele Sprachen (300) verfügbar



für slawische Sprachen de-facto vorhanden:

Carlton, Terence R. (1991): *Introduction to the phonological history of the Slavic Languages*. Columbus, Ohio: Slavica.

bzw.

Mel'nyčuk, O. S. (Hg.) (1966): *Vstup do porivnjal'no-istoričnoho vyvčennja slov'jans'kich mov*. Kyïv: Naukova Dumka.

Ausgangspunkt:

- Wortschatzliste für einen (möglichst exakten) Vergleich von Sprachen
- „Basiswortschatz“: hohes Alter, wenig Entlehnungen,

- (1) allgemein gebrauchte Adjektive,
- (2) bekannte Tiere und Vögel,
- (3) bekannte Pflanzen,
- (4) allgemein bekannte Verben,
- (5) Verwandtschaftsbezeichnungen,
- (6) Natur, Werkzeuge und häusliches Wirtschaften,
- (7) Ernährung und Nahrung
- (8) Körperteile (Mensch, Tier).

COMMON PLANTS (cont'd)

| <i>PSl</i> | <i>OCS</i> | <i>Ukr</i> | <i>R</i> | <i>Br</i> | <i>Blg</i> | <i>M</i> | <i>SC</i> |
|------------|------------|------------|----------|-----------|------------|----------|-----------|
| льпъ | — | льон | лен | лѣн | лен | лен | лѣн |
| malina | — | малина | малина | маліна | малина | малина | маліна |
| орѣхъ | — | (г)оріх | орех | арэх | орех | орев | òрах |
| овьсь | — | овес | овес | авѣс | овес | овес | òвас |
| просо | просо | просо | просо | проса | просо | просо | прòсо |
| ръшеніца | пшеница | пшениця | пшеница | пшаніца | пшеница | пшеница | пшєница |
| гъжь | — | — | рожь | — | ръж | ’рж | рѣж |
| sliva | — | слива | слива | слива | слива | слива | шльйва |
| smе/ьгk | — | смерека | смерч | смерч | смрека | смрека | смрѣка |
| trava | тpава | тpава | тpава | тpава | тpева | тpева | тpѣва |
| vьrba | вьрба | верба | верба | вярба | вьрба | врба | врба |
| želodь | желѣдь | жолудь | желудь | жолуд | желъд | — | жѣлуд |
| žito | жито | жито | — | жыта | жито | жито | жйто |

aus Carlton (1991: 340)

- keine „Repräsentativität“ aufgrund der geringen Anzahl von Testformen
- fehlende Äquivalente bzw. semantische Übereinstimmung
- geringe interne Differenzierung des Materials, d.h. geringe Aussagekraft
- keine Funktionswörter
- keine Gewichtung der Wortarten
- vermutlicher Lehnwortanteil sehr gering
- keine validen quantitativen Vergleiche möglich

Alternative (2): erweiterter Grund- und Erweiterungswortschatz

Annahmen der Sprachdidaktik :

3000 Lemmata Grundwortschatz + 2000 Lemmata Aufbauwortschatz = 5000 L.

- pragmatische Herangehensweise (viele „Vorbildprojekte“ für Einzelsprachen)
- Wortschatz gewährleistet eine funktionierende Kommunikation in unterschiedlichen Kommunikationssphären (mit variabler Klassenbildung)
- kein verbindlicher Rahmen vorhanden
- Beispiel:

Kelih/Vučajnk (ed.): *Grund- und Aufbauwortschatz Slowenisch – Deutsch*.
[in Arbeit, seit 2011]

- | | |
|---|--|
| (1) Personalien, Informationen zur Person, | (10) Ausbildung und Schule, |
| (2) Wohnen, | (11) Arbeit und Beruf, |
| (3) Umwelt, Pflanzen, Tiere | (12) Sprachen und Fremdsprachen, |
| (4) Reisen und Verkehr, | (13) Freizeit und Unterhaltung, |
| (5) Speisen und Verpflegung, | (14) Persönliche Beziehungen und Kontakte, |
| (6) Einkaufen, | (15) Politik und Gesellschaft. |
| (7) Öffentliche und private Dienstleistungen, | (16) häufigste Verben |
| (8) Körper, Gesundheit und Hygiene, | (17) Strukturwörter |
| (9) Wahrnehmung und Motorik | |

orientiert sich an Glaboniat, M.; Müller, M.; Rusch P. (2002): *Profile Deutsch*. Berlin: Langenscheidt.

Auswahl? - welche Metainformationen ? + phraseologische Komponente

Alternative (3): Lehnwörter innerhalb von Häufigkeitswörterbüchern

- Häufigkeitswörterbücher für viele Sprachen und Subregister vorhanden
- lexikalisches Minimum (hohe Grad an Textdeckung)
- hohe Abhängigkeit vom Textinput

Häufigkeitswörterbücher für Slowenisch?

- weitaus schlechtere Ausgangslage als für Russische, Tschechisch, Slowakisch usw.
- kein „repräsentatives“ Häufigkeitswörterbuch für das Slowenische vorhanden (!)
- z.T. Absichtserklärungen (T. Erjavec, P. Jakopin)
- Jakopin, P. (1995): Nekaj številčk iz Slovarja Slovenskega Knjižnega Jezika, in: *Slavistična revija* 43 (3), S. 341–375.
- Liste der 500 häufigsten Wortformen aus dem *Slovar Slovenskega Knjižnega Jezika* 1970 ff.
- Jakopin, P. (2002): *Entropija v slovenskih leposlovnih besedilih*. Ljubljana: Založba ZRC.
- Wortlisten für zwei Prosa-Korpora des Slowenischen vorhanden (Eigenarbeit), d.h. viel Basisarbeit noch notwendig
- Vergleichbarkeit von Häufigkeitswörterbüchern (methodologische Herausforderung)

Auszug aus dem kroatischen Häufigkeitswörterbuch

Moguš, Milan; Bratanić, Maja; Tadić, Marko (1999):
Hrvatski čestotni rječnik. Zagreb: Školska knjiga.

| Rang | Wortform | Rang | Wortform | Rang | Wortform |
|------|----------|------|----------|------|--------------|
| 1 | biti | 25 | ovaj | 158 | zašto |
| 2 | i | 26 | o | 159 | oblik |
| 3 | u | 27 | ali | 160 | organizacija |
| 4 | da | 28 | jedan | 161 | razvoj |
| 5 | on | 29 | moj | 162 | vrlo |
| 6 | na | 30 | ili | 163 | živjeti |
| 7 | koji | 31 | kako | 320 | različit |
| 8 | ja | 32 | imati | 321 | upravo |
| 9 | za | 33 | sav | 322 | dinar |
| 10 | sebe | 34 | mi | 323 | sistem |
| 11 | ne | 35 | tako | 324 | tišina |
| 12 | s | 36 | samo | 356 | poseban |
| 13 | a | 37 | kad | 357 | problem |
| 14 | htjeti | 38 | znati | 358 | funkcija |
| 15 | što | 39 | pa | 359 | davati |
| 16 | taj | 40 | po | 360 | mjera |
| 17 | od | 41 | još | 361 | materijal |
| 18 | to | 42 | drugi | 362 | umrijeti |
| 19 | kao | 43 | do | 363 | djelo |
| 20 | ti | 44 | onaj | 364 | bojati se |
| 21 | moći | 45 | vi | 365 | također |
| 22 | svoj | 46 | ni | 366 | trava |
| 23 | iz | 47 | velik | 367 | turistički |
| 24 | sve | 48 | više | 415 | tiho |
| 25 | ovaj | 49 | neki | 417 | politički |
| 26 | o | 50 | li | 418 | tekst |



Vorteile: ?

Nachteile?

Alternative (4): Liste von Haspelmath/Tadmor (2009)

Appendix: The Loanword Typology meaning list

| | | | |
|--------------------------|------|-------------------|-------------------|
| 1. The physical world | 1.44 | stone or rock | 2. Kinship |
| 1.1 world | 1.45 | earthquake | 2.1 person |
| 1.21 land | 1.51 | sky | 2.21 man |
| 1.212 soil | 1.52 | sun | 2.22 woman |
| 1.213 dust | 1.53 | moon | 2.23 male (1) |
| 1.214 mud | 1.54 | star | 2.24 female (1) |
| 1.215 sand | 1.55 | lightning | 2.25 boy |
| 1.22 mountain or hill | 1.56 | thunder | 2.251 young man |
| 1.222 cliff or precipice | 1.57 | bolt of lightning | 2.26 girl |
| 1.23 plain | 1.58 | storm | 2.261 young woman |
| 1.24 valley | 1.59 | rainbow | 2.27 child (1) |
| 1.25 island | 1.61 | light | 2.28 baby |
| 1.26 mainland | 1.62 | darkness | 2.31 husband |
| | 1.63 | shade or shadow | |

- adaptieren auf „slawische Verhältnisse“ ?
- Umgang mit Wortbegriff – Wortverbindungen – mehrgliedrigen Verbindungen
- „problematische“ Begriffe (*grass-skirt, men's house, garden-house, older brother, older sister, fermented drink*) ersetzen bzw. streichen?
- Anzahl an Internationalismen (*ocean, lagoon, savanna, manioc bread, toucan, radio, telephon*) reduzieren?
- die Auswahl an Funktions- bzw. Strukturwörtern reflektieren ?

II. Wortliste – Verteilung der Wortarten?

bekannte Befunde der Lehnwort-Forschung: „Hierarchie“ der Entlehnungen:

1. Substantive (innerhalb dessen Konkretes vor Abstraktem)
2. Verben (weitaus weniger häufig)
3. Adjektive/Adverben
4. Weitere Funktionswörter

→ Die Reihenfolge ist umstritten!

by category compared to other languages

- (6) Haugen (1950: 224), on Norwegian and Swedish immigrant speech in the US: nouns > verbs > adjectives > adverbs, prepositions, interjections
- (7) Muysken (1981), on Spanish in Quechua (repeated by Winford 2003: 51): nouns > adjectives > verbs > prepositions > coordinating conjunctions > quantifiers > determiners > free pronouns > clitic pronouns > subordinating conjunctions

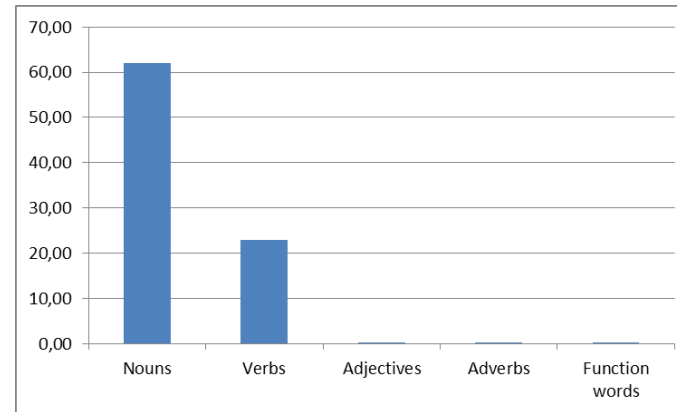
nouns, conjunctions > verbs > discourse markers > adjectives > interjections > adverbs > other particles, adpositions > numerals > pronouns > derivational affixes > inflectional affixes

aus: Matras, Yaron (2009): *Language contact*. Cambridge: Cambridge Univ. Press, S. 157

Wie sieht die quantitative Verteilung von Wortarten aus?

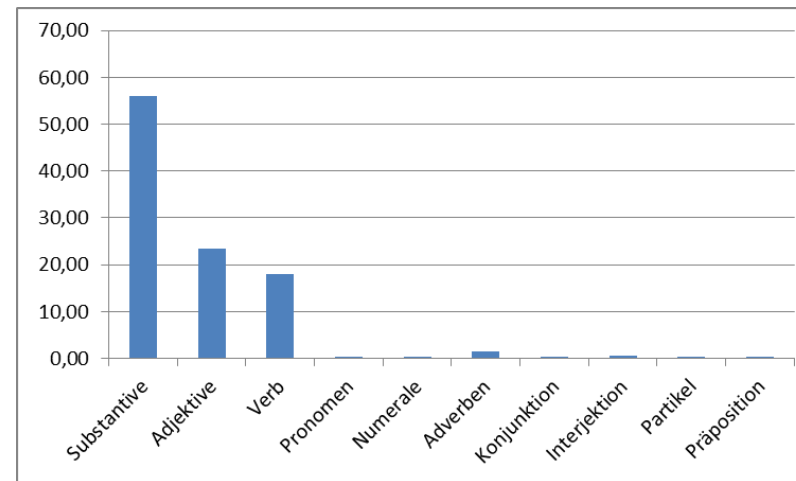
1. Liste von Haspelmath/Tadmor (2009)

| Semantic word classes | | abs. | % |
|-----------------------|----------------|------|-------|
| 1 | Nouns | 905 | 61,99 |
| 2 | Verbs | 334 | 22,88 |
| 3 | Adjectives | 120 | 8,00 |
| 4 | Adverbs | 4 | 0,01 |
| 6 | Function words | 97 | 7,00 |
| | | 1460 | 100 |



2. Slovar slovenskega knjižnega jezika 1970 ff.

| Wortart | abs. | % |
|--------------|-------|--------|
| Substantive | 51448 | 56,02 |
| Adjektive | 21516 | 23,43 |
| Verb | 16479 | 17,94 |
| Pronomen | 130 | 0,14 |
| Numerale | 122 | 0,13 |
| Adverben | 1325 | 1,44 |
| Konjunktion | 76 | 0,08 |
| Interjektion | 615 | 0,67 |
| Partikel | 9 | 0,01 |
| Präposition | 115 | 0,13 |
| Summe | 91835 | 100,00 |

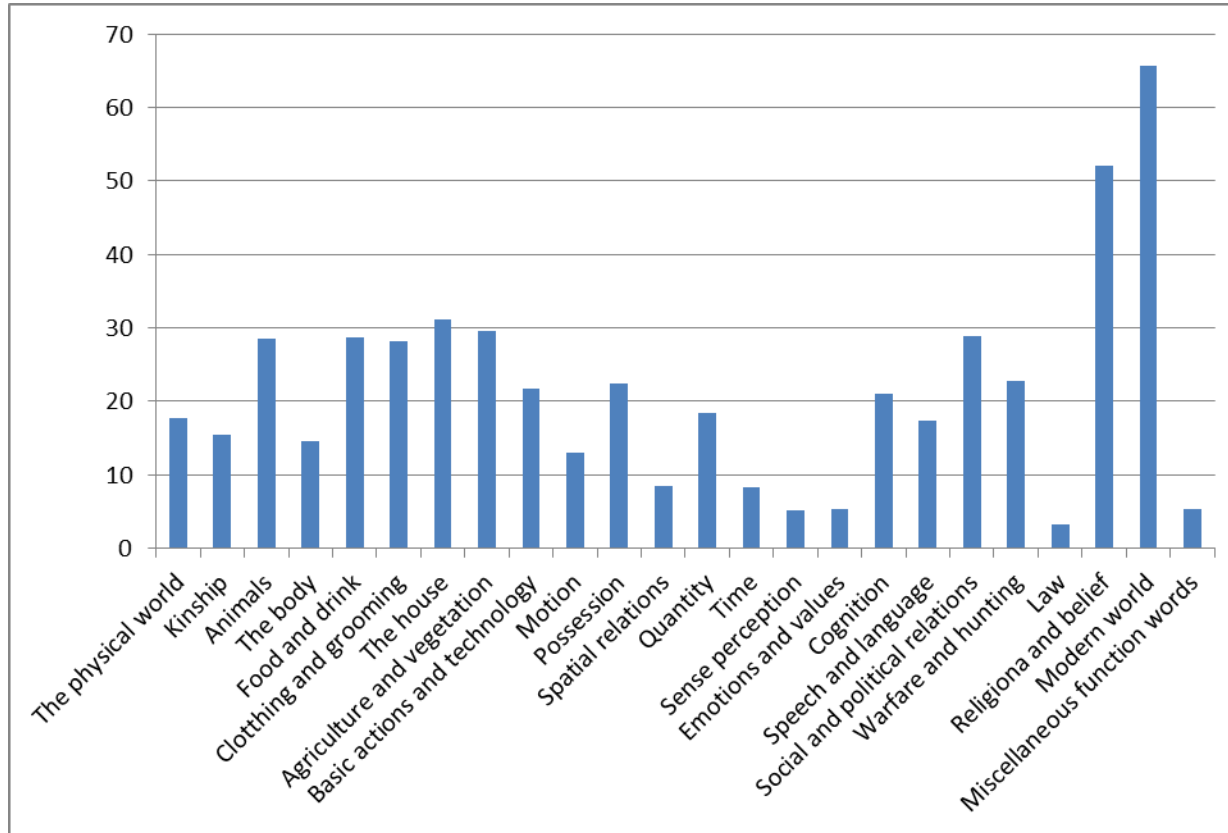


Zwischenergebnisse & Perspektiven:

- einheitliches Schema für slawische Sprachen zu entwickeln
- Adaption und Spezifizierung der Wortliste
- Anzahl von Synonymen pro Begriff (Einschränkung?)
- Auswahl der Meta-Informationen
- Unterschiedliche Wahrscheinlichkeitsgrade von “Entlehnungen” – theoretisch gerechtfertigt?
- Unterscheidung von Registern
- zeitliche Tiefe

... .. noch ein langer Weg zu einem

validen “quantitativen“ Lehnwortprofil:
z.B. Niedersorbisch (Daten basierend auf Bartels 2009):



und einem Vergleich ... mit anderen Sprachen

