

GiF:online

Giessener Fremdsprachendidaktik: online 11

Franz-Joseph Meißner

Die Vermessung des Kern- wortschatzes der romani- schen Mehrsprachigkeit

**Eine didaktische Analyse zur
interlingualen Transparenz- und
Frequenzforschung**

GIESSENER FREMDSPRACHENDIDAKTIK: ONLINE 11

Herausgegeben von Eva Burwitz-Melzer, H el ene Martinez, Franz-Joseph Meißner

Franz-Joseph Meißner

**Die Vermessung des
Kernwortschatzes der
romanischen Mehrsprachigkeit**

Eine didaktische Analyse zur
interlingualen Transparenz- und
Frequenzforschung

Bibliografische Information der Deutschen Nationalbibliothek

Die deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie. Detaillierte bibliografische Daten sind im Internet unter <http://dnb.de> abrufbar.



Diese Veröffentlichung ist im Internet unter folgender Creative-Commons-Lizenz publiziert: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

ISBN: 978-3-944682-31-0

URN: <http://nbn-resolving.de/urn:nbn:de:hebis:26-opus-135898>

Inhaltsverzeichnis

Abbildungen	3
1. Vorbemerkung und Verortung	6
2. Die pädagogische Bedeutung von Formkongruenz und Formdivergenz und die didaktische Mehrsprachenlexikographie	12
3. Zählungen aus „primären“ und „sekundären Zielsprachen“ nach Formkongruenzen und Formdivergenzen	16
3.1 Zu dem, was, wie, wann gezählt wird.	16
3.2 Ziele der Zählung	18
3.3 Zählungen <i>ohne</i> Differenzierung nach sekundären (oder nachgelernten) Zielsprachen.....	20
3.4 Zählungen <i>mit</i> Differenzierung nach sekundären Zielsprachen.....	20
3.5 Legende und Methode der Zählungen	21
3.6 Komputationen mit Access und Excel	28
3.7 Triangulation der Zählungen	31
4. Vorstellung der Selektionen und ihre Quantifizierung im Überblick vorab	32
5. Bemerkungen zu einzelnen Sprachen aus Sicht der Transferforschung	36
5.1 Grundwortschatzliche Transferanteile nach Sprachen	37
5.2 Der englische Wortschatz als Brücke zu den romanischen Sprachen	39
5.3 Latein als gemeinsame Mutter der Romania aus Sicht des interlingualen Transfers.....	44
5.4 Die Mutter- oder Zweitsprache Deutsch als Grundlagensprache für den romanischen Transfer.....	49
6. Frequenzintervalle aus erweiterten Grundmengen für primäre (romanische) Zielsprachen.....	49
7. Frequenzintervalle in Kombination mit Transparenzstufen in primären Zielsprachen.....	55
8. Interlinguale Transparenzintervalle: Methodische Vorüberlegungen und Daten.....	57

8.1 Grundmengen (aus sprachlichen Variablen)	57
8.2 Zählungen aus gemischt-sprachlichen Kombinationen (MIT-Selektionen).....	60
8.3 Komputationen zur Transparenzstufe 3: zum Transparenzanteil aus Opaken-Zählungen	63
Syntaktisches zur Ermittlung der T3-Kohorte.....	65
8.4 \$-Markierungen	72
8.5 FaX-Marker zur Kennzeichnung minimaler morphologischer Distanz	76
9. Methodisches zu sprachlich-kombinierten Komputationen	77
9.1 Mögliche Sprachkombinationen für den zwischensprachlichen Transfer	77
9.2 Syntax der Selektoren.....	78
10. Komputationen von Transparenzprofilen in der Kombination von zwei und mehreren Sprachen	83
10.1 Grundmengen aus intralingual Formdivergenten und Reduplikationen	83
10.2 Dreisprachige Grundmengen aus gemischt-mehrsprachig einfacher und gemischt-mehrsprachig erweiterter Komposition im Vergleich	87
10.3 Transparenzprofile von intralingualen Formdivergenzen in monolingualen und gemischt-bi- oder trilingualen Grundmengen.....	90
11. Darstellung und Interpretation der gesamten Daten aus bilingualen MIT-Ziehungen	93
11.1 Vermessung morphologischer Kongruenz- und Divergenz-Relationen in gemischt-sprachlichen Serien.....	99
11.2 Über die Romania hinaus zum europäischen Wortschatz	102
12. Ergebnisse und Diskussion	105
12.1 Was zu erbringen war.....	105
12.2 Didaktische Rahmung	107
12.3 Was konkret wozu zu vermessen war	109
12.4 Lexikodidaktische und -statistische Gesamtcharakteristik.....	111
12.5 Welche methodischen Konsequenzen sind aus der quantitativen Analyse des KRM zu ziehen?	115
13. Anwendungen (Beispiele).....	116
Literatur	130
Index	137

Abbildungen

Abbildung 1: GWS und KWS im Diaregister (vgl. Meißner 2016: 28)	8
Abbildung 2: Beispiel für eine Serie (ohne Etyma)	9
Abbildung 3: Syntax	25
Abbildung 4: Alle Markierungen im Überblick aus dem Gesamt des KRM-Inventars abzüglich der "Spezifizierung"	33
Abbildung 5: die wichtigsten Selektionen (gepr. am 24.04.2018)	34
Abbildung 6: Transparenz-Kategorien in % (gerundet)	35
Abbildung 7: Transparenzselektionen in % ohne sprachliche Spezifizierung	36
Abbildung 8: Grundwortschatzliche Transferpotentiale von Englisch und Französisch im Vergleich zueinander und zu romanischen Sprachen	37
Abbildung 9: Grundwortschatzlicher Ausschnitt aus „R“ in Rfr, Rpg und Rsp bzw. „M“ in Rit	38
Abbildung 10: Englische Transferbasen in Französisch, Italienisch, Portugiesisch, Spanisch (aus X)	41
Abbildung 11: Fälle und Prozentwerte englischer Transferbasen aus (X)	42
Abbildung 12: Transferbasen mit Englisch und Deutsch in Fällen und % (Zählung 01.12.2017)	43
Abbildung 13: zufällig gewählter Ausschnitt aus dem KRM mit den Rubriken ROS- und ETYMA	47
Abbildung 14: Frequenzintervalle Kernwortschatz Spanisch	50
Abbildung 15: Frequenzintervalle Französisch	51
Abbildung 16: Frequenzintervalle Italienisch	51
Abbildung 17: Frequenzintervalle Portugiesisch	52
Abbildung 18: Frequenzintervalle zu den GrdM nach intralingual formdivergenter Zählung	53
Abbildung 19: Frequenzintervalle des Italienischen aus I und I OR Iw	54
Abbildung 20: X-, Xw-, X OR Xw-Anteile nach Sprachen und Fällen	58
Abbildung 21: Interlinguale Transparenzanteile in Grundmengen ohne sekundäre Zielsprachen	59
Abbildung 22: Zählung T-Kohorten aus (X)	60
Abbildung 23: Selektionen aus unterschiedlichen Grundmengen zu Französisch	63
Abbildung 24: Anteil von T3 Transparenzfälle in den Symbolfarben für die Zielsprachen	64
Abbildung 25: T3 mit RoSED-X	65
Abbildung 26: Maximale Opaken-Anteile aus monolingualer X-Zählung	67
Abbildung 27: Ausschnitt aus der Selektion RoS-* bzw. *-*	69
Abbildung 28: Opazität und Transparenz in T3-Selektionen aus X MIT Y	69

Abbildung 29: Transparenzraten von T3 für unterschiedliche *-X aus gemischt-sprachlicher Ziehung	71
Abbildung 30: Prozentanteile von \$-Markierungen	74
Abbildung 31: \$-Markierungen in T2 und T1.....	75
Abbildung 32: Mögliche Sprachen-Paare und Sprachkombinationen	77
Abbildung 33: Komponentenanteile der erweiterten Grundmengen in %.....	83
Abbildung 34: Erweiterte Grundmenge mit einer sekundären Zielsprache, Selektoren aus der RoS-Reihe mit T3- Differenzierung für IT und PG.....	85
Abbildung 35: Erweiterte Grundmenge mit einer sekundären Zielsprache, Selektoren aus der RoS-Reihe mit T3-Differenzierung für IT und PG.....	86
Abbildung 36: Selektoren-Kohorten aus 3sprachiger Grundmenge (P OR Pw) MIT (I) MIT F)	87
Abbildung 37: Vergleich aus dreisprachigen Grundmengen, gemischt-mehrsprachig, einfach und erweitert.....	89
Abbildung 38: A: Abhängigkeit von Variablen.....	90
Abbildung 39: B: Variablen und Größenverhältnisse (gepr. 01.05.18).....	91
Abbildung 40: Gezielte Selektion aus mono-, gemischt-bi- und -trilingualen Grundmengen.....	92
Abbildung 41: RoS-Selektion aus viersprachigen Grundmengen	93
Abbildung 42: MIT-Zählungen (12.12.2017).....	94
Abbildung 43: Transparenzstufen nach Fällen zu Abb. 39	96
Abbildung 44: die intralingual-divergenten Lemmata in sprachlich-gemischter Kombination in %	98
Abbildung 45: Transparenzklassen in gemischt-sprachlichen Grundmengen aus X zu Y	100
Abbildung 46: Transparenzraten Klassen aus trilingual gemischt-sprachlicher Zählung ohne T3	101
Abbildung 47: Synopse der Transferpotenziale	114
Abbildung 48: Ausschnitt aus der Opaken-Liste für Italienisch.....	117
Abbildung 49: Ausschnitt aus der Selektion: englische Transferbasen für Italienisch.....	118
Abbildung 50: portugiesische Profilwörter im Kontrast mit englischen Transferbasen.....	119
Abbildung 51: Ausschnitt englischer Transferbasen für romanische Sprachen.....	120
Abbildung 52: Phonologische Korrespondenz FR = Konsonant+L ~ IT = Konsonant+I	121
Abbildung 53: Alle FR = Konsonant+L ~ IT = Konsonant+I-Korrespondenzen aus KRM	122
Abbildung 54: Ausschnitt aus der Selektion FL="*pl*" MIT IT="*pi*"	122

Abbildung 55: Phonologische Korrespondenzen für FR = *ch* und IT `“c“	124
Abbildung 56: Minimale Formabweichungen von Portugiesisch	126
Abbildung 57: até logo/hasta luego	126
Abbildung 58: Ausschnitt aus Formkonvergenzen Portugiesisch/Spanisch	127
Abbildung 59: Sprachlich-paarige Selektion	127

Vorbemerkung für eilige Leser

Lexikostatistische Ausführungen, die auf eine Datenbank und deren Syntax rekurren und immer wieder Zählungen aus unterschiedlichen Grundmengen präsentieren, sind keine leicht eingängliche, sondern eine anspruchsvolle Lektüre.

Wer mehr an der Relevanz der Ergebnisse für die Praxis des Fremdsprachenunterrichts oder für das eigene Lernen interessiert ist, sei direkt auf die Kapitel 1, 2 und 11 verwiesen.

1. Verortung

Der *Kernwortschatz der romanischen Mehrsprachigkeit* (KRM) wurde in der gleichnamigen elektronischen Mehrsprachen-Datenbank zur Förderung der Lesekompetenz in vier romanischen Sprachen – Französisch, Italienisch, Portugiesisch und Spanisch (sowie in Englisch und Deutsch¹) – zusammengestellt. Hinzu kommen die Etyma. Der KRM ist sowohl ein Instrument der romanistischen Interkomprehensionsdidaktik als auch der lexikodidaktischen Fremdsprachenforschung und der Praxis des Fremdsprachenlehrens und -lernens. Da die Selektion des KRM-Vokabulars, die Struktur sowie die Frequenzierung und Behandlung der Lemmata² andernorts detailliert erörtert wurden, erübrigen sich an dieser Stelle weitere Angaben (Meißner 2016; 2018).

In Übereinstimmung mit zahlreichen frequenzlexikographischen Glossaren ist unter Kernwortschatz das Wortschatzsegment unterhalb eines Frequenzrangs von 5000 zu verstehen (u.a. Davies 2006; generell: Kühn 1990; Zöfgen 1994; 2006; Tschirner 2005; Siepmann 2006; Perkuhn et al. 2012). Die folgende Abbildung zeigt den Grund- (GWS) und Kernwortschatz (KWS) im Ensemble des Diaregisters:

¹ Vgl. Tschirner (2005: 134): „Entgegen den Annahmen der Lesedidaktik der 70er und 80er Jahre des letzten Jahrhunderts haben empirische Untersuchungen ergeben, dass nicht nur 60-70 % der Wortformen eines Textes bekannt sein müssen, um ihn zu verstehen, sondern 95 % und mehr (Carver 1994; Laufer 1997; Hu/Nation 2000; Qian 2002). Um Wörter aus dem Kontext erraten zu können und um unbekanntes Vokabular auf implizite Art und Weise zu lernen – wichtige Ziele der fremdsprachlichen Lesedidaktik – müssen sogar mindestens 97 % der laufenden Wortformen eines Textes verstanden werden (Swanborn & de Glopper 1999). Nach Laufer (1997) benötigt man für das Englische einen Wortschatz, der die 5000 häufigsten Lexeme umfasst, um 95 % der Wortformen eines durchschnittlichen Zeitungstextes oder Fachaufsatzes zu verstehen und um ihn damit mit Verständnis und ohne größeren Zeitverlust lesen zu können.“ – Aus diesen Bemerkungen lässt sich unschwer ableiten, mit welchem großem Nutzen sich ein mehrsprachlicher Kernwortschatz verbindet.

² Unter ‚Lemma‘ wird ein Eintrag in das Sprachinventar des KRM verstanden.

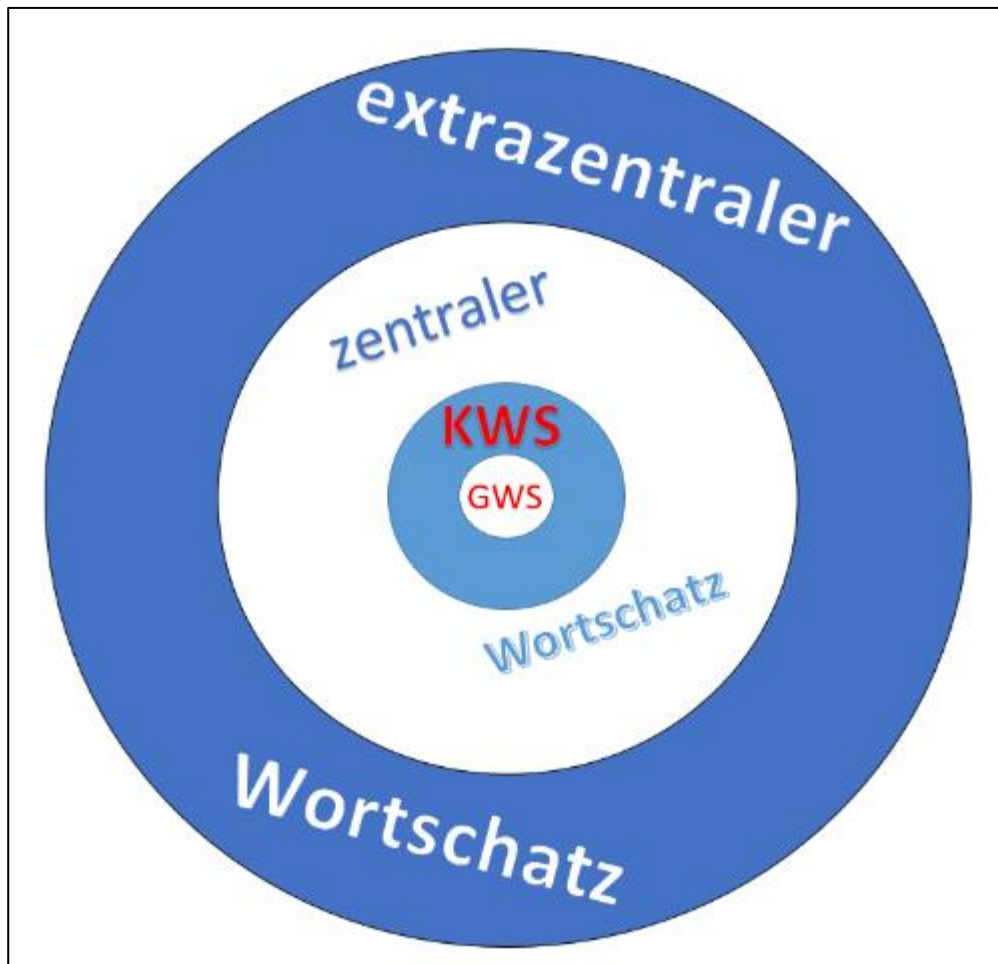


Abbildung 1: GWS und KWS im Diaregister (vgl. Meißner 2016: 28)

Als Instrument der Interkomprehensionsdidaktik sind der KRM und die aus seinem Bestand generierten [Lernapps](#) vor allem ein Mittel zur Förderung der Lesekompetenz in mehreren romanischen Sprachen³ und eines reflexiven Zugriffs auf Sprachen und Sprachenlernen. Dem entsprechen die Struktur und die notwendigen lerner-adressierten Informationen des KRM. Dass indes die Lesekompetenz auch Ressourcen liefert, deren Verfügbarkeit den Ausbau produktiver Fertigkeiten befördert, ist trivial und wurde detailliert andernorts erläutert.

³ Der KRM-Bestand liefert die Grundlage für sog. Lernapps (vgl. www.eurocomdidact.eu), die unter den o.g. Adressen beschrieben wurden. Ihre Nutzung ist gegen eine geringe Gebühr für alle Interessenten offen.

Die im Weiteren ausgewerteten Quantifizierungen beleuchten den romanischen Kernwortschatz unter dem Gesichtspunkt des interlingualen Identifikationstransfers: Es werden Parameter dazu entwickelt und geprüft, inwieweit die Kenntnis von Lexemen des Kernwortschatzes der Sprache X helfen kann, inhaltliche und formähnliche Entsprechungen in der Sprache Y und/oder Z zu identifizieren, d.h., inwieweit Bestände der romanischen Sprache X lexikalische Transferbasen für die Schwestersprachen stellen können.

Der KRM stellt Bedeutungsadäquanzen zu den sechs erwähnten Sprachen und den jeweiligen Etyma in Serien zusammen.

451 CHOIX	SCelta	ESCOLHA	elección	choice	(Aus)wahl
-----------	--------	---------	----------	--------	-----------

Abbildung 2: Beispiel für eine Serie (ohne Etyma)

Maßgeblich für die Auswahl der Lexeme, ihrer Zuordnung bzw. Serialisierung qua Übersetzung waren erstrangig die Aussagen der Frequenzforschung zu den Lemmata jeder einzelnen Zielsprache und zweitrangig deren interlinguale Transparenz bzw. Formkongruenz. Eine eigene Rubrik zeigt zudem interseriell semantische Bezüge (Synonyme), z.B. zwischen fr. *hésiter* und *tergiverser* und deren Serien. Hierneben weist der KRM die Rubriken ETYMA und INTERLIGALEXE auf. Die Etyma bzw. die Hinweise in einer weiteren Rubrik ETYDIST⁴ erlauben neben den Angaben zur Art der Serialität (in der Spalte ROS-) und zu kontrastsprachlichen Transparenzkombinationen (in PROFIL) Hinweise auf den Grad der Trennschärfe zwischen den interlingual transparenten oder opaken Formen und den Klassen der Serie. Benutzt werden die in der interkomprehensionsdidaktischen Literatur üblichen Begriffe *Formkongruenz* (Typ: fr. *ange*, pg. *anjo*, sp. *angel*, it. *angelo*, dt. *Engel*, en. *angel*) und *Formdivergenz* (fr. *balai*, pg. *vassoura*, sp. *escoba*, it. *ramazza/scopa*, dt. *Besen*, en. *broom*) (Meißner 1993). Schon hier sei unterstrichen: Die Zahl der Etyma, die sich für eine Serie nennen lassen,

⁴ Die Rubrik ETYDIST verzeichnet die Zahl der Etyma pro Serie oder der für linguistische Laien augenscheinlich opaken Formen innerhalb derselben (Transparenzstufe 0 mit 4 stark divergierenden opaken Formen). Dies macht folgende Bemerkung erforderlich: Der Begriff *Etymon* fasst in diesem Zusammenhang nicht ganz eindeutig. Denn da sich der KRM letztlich an Lerner romanischer Sprachen wendet, ist die augenscheinliche Formdivergenz für die Nennung maßgebend, weniger die belastbare wortgeschichtliche Aszendenz. Hierzu ein Beispiel: Die Nachfahren der Wurzel von vlat. TURBULUS aus TURBULENTUS und TURBIDUS bzw. TURBULENTIA begegnen in en./fr. *trouble* und *turbulence*, dt. *Turbulenz* und als Fremdwort dem aus dem Englischen entlehnten dt. *Trouble* (Sorgen) sowie in *turbid (water)*, (*de l'eau trouble*, sp. *turbio*, *turbido*, it. *torbido*; dt. *trüb* ist mit TURBIDUS auf indoeuropäischer Stufe verwandt (Grimm & Grimm 1952: XXII, 1164; auch: Meyer-Lübke 1911: 8994 und weitere). Die modernen Vertreter sind für Lerner nicht ohne Weiteres als etymologische Verwandte erkennbar. Die Entscheidung entspricht der didaktischen Orientierung des KRM.

deutet zudem auf die morphologische Distanz zwischen den Lemmata hin: Ein Mehr an Etyma in einer Serie entspricht einem Mehr an Formdivergenz (bzw. einem Weniger an Formkonvergenz) der Deszendenten. Die ‚Interligalexe‘ isolieren jene morpho-semanticen Komponenten von Wörtern, die den zwischensprachlichen Wiedererkennungs- oder Identifikationstransfer ermöglichen. Das Ligalex fr. *tromp-* ist Kernelement seiner Erweiterungen *trompeur*, *-euse*, *tromperie*, *trompe l'oeil*. Da die zuletzt genannte Bildung in den anderen Sprachen keine formkonvergente Entsprechung hat oder Derivationen bildet, ist es kein Interligalex (wohl aber ein Lehnwort). Beispiele für interligalexikalische Bildungen und deren plurilinguale assoziativen Felder: fr. *trompé* / it. *sbagliato* / pg. *errado* / sp. *equivocado* > *tromp-*, *sbagl-* *err-*, *equivoc* > *tromper*, *trompeur*, *tromperie...* / *sbaglio*, *sbagliare* / *errar*, *error*, *erro*, *erroneamente*, *erróneo* / *equivocación*, *equivocar*, *equivocadamente*, *equivoco*. Die spanische Wurzel *equivoc-* begegnet interlingual in pg. *equivocidade* (Mehrdeutigkeit), *equivocação*, *equivoco* (zweifelhaft, suspekt), fr. *équivoque* (mehrdeutig), en. *equivocation*, *equivocal* und *equivocality*, *to equivocate* usw. Ein weiteres interlinguales Assoziationsschema ergibt sich aus *voc-* / *voz-* / *voic-* > *voice* / *voix* / *voce* / *voz* / *voix* > en. *voicing*, *voiceless*, *voicelessness* / fr. *vocable*, *vocale*, *vocatif*, *vocation*, *vociférer*, *vocodeur* / it. *vocabolario*, *vocale*, *vocalismo*, *vocalizzare* usw. Die Verbindungen zum deutschen Bildungswortschatz sind vielfach, weitere Sprachvergleiche, die auch die Bedeutungsadäquanz einbeziehen (Typ: *voiceless* > *sourd*, *sordo* usw. oder > *sin voz* oder *afonía*, *aphonie...*), sind uferlos. Insgesamt zeigen Interligalexe jene Bausteine, die dem romanischen Kernwortschatz (der weitgehend mit dem europäischen zusammenfällt) gemeinsam sind⁵. Sie sind die materiellen Kerne für jeden interlingualen Identifikationstransfer. Allerdings weisen diese Elemente in der Schreibung verschiedener Sprachen oft Abweichungen voneinander auf, die interlingual selbst wiederum eine immer wiederkehrende Systematik zeigen: für den Laut /f/ en. *elephant*, *telephone*, *orthography* / fr. *éléphant*, *téléphone*, *orthographe* / it. *elefante*, *telefono*, *ortografia* / sp. *elefante*, *telefono*, *ortografía*. Die orthographische Korrespondenzregel lautet hier: en. -ph- ~ fr. -ph- ~ it./pg./sp. -f-; bzw., wie schon erwähnt, *-cracy*, *-cratie*, *-crazia*, pg./sp. *-cracia* (oder *-autocracy*, *autocratie*, *autocrazia*, pg./sp. *autocracia*). Zugleich weisen die Ligalexe über den Kernwortschatz hinaus. Pädagogisch gewendet heißt dies: Die interlinguale Lernökonomie (die ja den Mehrsprachenerwerb so erleichtert) findet ihre materielle Basis letztlich in den Interligalexen. Die Identifikation dieser Elemente – Kernbausteine des romanischen Wortschatzes bzw. des gemeinsamen Wortschatzes einer Sprachenfamilie – stellt eine der Innovationen der Interkomprehensionsdidaktik und des KRM dar.

⁵ Die Rubrik INTERLIGALEXE ist zum Zeitpunkt der Erstellung dieses Aufsatzes noch nicht vollständig bearbeitet, so dass Ausführungen hierzu noch nicht erfolgen.

Indem sich die Analysen allein auf die Darstellung der Faktoren ‚Sprachen‘ bzw. zwischensprachliche Ähnlichkeiten (und Unähnlichkeiten) – seien sie nun semantisch und/oder morphologisch – beschränken und die Lemmata um mögliche Kontexte reduzieren, klammern sie weitere Determinanten aus, die für die Realisierung eines interlingualen Identifikationstransfers notwendig sind. Dies betrifft vor allem lernerseitige bzw. personenbezogene Faktoren (Motivation, Resilienz, Vorwissen) sowie sprechsituative und kontextliche Variablen. Allerdings sind diese Faktoren individuell, oft einzigartig und hochgradig variierend. Gleiches lässt sich von dem Faktor Sprache und Sprachen nicht sagen. Wörter und sprachliche Strukturen sind – psycholinguistisch gesprochen – kognitive Schemata, deren überindividuelle Merkmale (Bedeutungen, Funktionen und Formen) Sprache als soziales Institut erst ermöglichen.

Die Kompilation und Quantifizierung des KRM-Inventars nach den Kriterien Bedeutungsadäquanz, Frequenzrang, interlinguale Transparenz und Opazität ist nicht ohne Belang für die Praxis des mehrsprachlichen Wortschatzerwerbs: Denn die Komputationen erlauben die Eingrenzung und pädagogische Einschätzung neuen lexikalischen Inputs auf der Grundlage des den Lernern schon bekannten Wortwissens sowie seine Einstufung im Hinblick auf die lernrelevanten Kriterien ‚Transparenz‘ (=Lernerleichterung aufgrund einer interlingual greifenden Transferbasis) und ‚Opazität‘ (=tendenzielle Lernerschwerung aufgrund einer fehlenden Transferbasis). Dass eine solche Konturierung in vielfacher Weise für die Konstruktion didaktischer Materialien von [Belang](#) ist, liegt auf der Hand.

Es erscheint angeraten, auch jene Fragen anzusprechen, die zu Irritationen führen könnten. Dies betrifft zunächst die Beziehung zwischen Polymorphie und Polysemie bzw. Unklarheiten bei der morphosemantischen Zuordnung der Adäquanzen. Ausgehend von einem in der Frequenzlexikographie zur Sprache X mit dem Rang <5000 eingetragenen Lemma werden im KRM die Entsprechungen über Übersetzungen und deren Frequenzrang in den romanischen Nachbarsprachen ermittelt. Ist schon die Zuordnung zwischen einem Bezeichnenden (Signifikant) und einem Bezeichneten (Signifikat) innerhalb einer einzigen Sprache „arbiträr“ im Sinne Saussures, so gilt dies erst recht für deren Übertragungen in vier bzw. sechs Sprachen. Ob man z.B. fr. *piquer* mit dt. *klauen*, *stibitzen*, *wegschnappen*, *rapschen*, *klemmen*, mit sp. *despabilar*, *birlar*, *gatear*, *afanar* oder sonstwie übersetzt, ist u.a. eine Frage der subjektiven Einschätzung des Übersetzers und des Kontextes von *piquer*⁶. Hier gibt es keine in allen möglichen Fällen zwingende Zuordnung. Der KRM versucht derlei Ambiguitäten aufzulösen, indem er ‚disambiguierend‘ Synonyme und Intersynonyme in unterschiedlichen Serien bietet und damit zugleich für einzelne Lemmata Wiederholungen (Redundanzen) in Kauf

⁶ Hierneben besteht natürlich die Reihe zu fr. *piquer* (stechen), pg. *picar*, sp. *picar*, an der das Italienische nicht formkongruent teilnimmt (*pungere*, *beccare*).

nimmt. Diese betreffen natürlich nie die Serien, sondern immer nur einzelne Lemmata in Kombination mit anderen. Mögen Redundanzen generell stilistischen Kriterien abträglich und in der Wissenschaft verpönt sein, so sind Wiederholungen jedoch pädagogisch bzw. mnemotechnisch von Vorteil. Dies zeigt nicht zuletzt der Bau der Sprachen selbst, deren zentrale Elemente allesamt im alltäglichen Gebrauch hohe Wiederholungsraten zeigen. Der menschliche Spracherwerb greift auf diese in sehr starker Weise zu.

Adressaten dieses Beitrages sind in erster Linie Lexikographen und Linguisten sowie Fachdidaktiker der romanischen Sprachen, des Englischen und des Deutschen sowie Lehrerinnen und Lehrer der betroffenen Sprachen. Natürlich schließt diese Aufzählung auch Studentinnen und Studenten sowie Interessenten überhaupt ein.

Das Inventar des KRM ist mit einem Copyright belegt, so dass eine komplette Veröffentlichung der Liste absehbar nicht erfolgt. Es ist Zweck dieses und anderer einschlägiger Beiträge, alle das Korpus betreffenden lexikographisch und didaktisch relevanten Fragen und Daten der Fachöffentlichkeit vorzulegen.

2. Die pädagogische Bedeutung von Formkongruenz und Formdivergenz und die didaktische Mehrsprachenlexikographie

Das Thema war vielfach Gegenstand fremdsprachendidaktischer Publikationen und wird daher an dieser Stelle nur soweit behandelt, wie es hilft, die weiteren Ausführungen nachzuvollziehen.

Den Kerngedanken der Interkomprehensionsmethode hat der amerikanische Lernpsychologe David Ausubel (1968) avant la lettre formuliert: "If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor in influencing learning is what the learner already knows." (vi) Gilt dies generell für alle Wissensdomänen – von der Mathematik bis zur Musik –, so erst recht für den natürlichen (Mehr)spracherwerb und das gesteuerte Sprachenlernen. Gilt es für den Lerngegenstand Sprachen im Allgemeinen, so ist seine Wirkung im Falle interkomprehensibler Sprachen besonders augenfällig. Dies ist z.B. zwischen Deutsch und Niederländisch der Fall⁷:

⁷ Zu Recht empfiehlt Wenzel (2014: 267-281) in ihrer *Fachdidaktik Niederländisch* den interkomprehensiven Ansatz. Zur Interkomprehension in der Zielsprache Schwedisch auch: Meißner (2015).

Het dagblad **de Volkskrant** is een van oorsprong rooms-katholiek dagblad. De krant richt zich sinds de jaren zestig op de hoger opgeleide lezers en plaatste zich tot aan de jaren negentig links van het politieke midden. Later schoof de krant meer op naar rechts.

Das Tageblatt **Die Volkszeitung** ist ein vom Ursprung her (ursprünglich) römisch-katholisches Tageblatt. Die Zeitung richtet sich seit den Jahren 60 (60er Jahren) an (höher opgeleide?) (höher opgeleide? aufleiten?) höher gebildete Leser und platziert sich bis in die neunziger (negentig: PRÜFEN) Jahre (hinein) links von der politischen Mitte. Später schob sich die Zeitung mehr und mehr nach rechts. (Warum „op“?)

Deutschsprachige verstehen den niederländischen Text, weil sie hinreichend über Transferbasen verfügen.

Ähnliches lässt sich für die romanischen Sprachen (und für andere Sprachfamilien) sagen. Die Feststellung ist trivial: Ein Mehr an (erkennbaren) interlingualen Form- und Funktionskongruenzen (Transferbasen) erleichtert sowohl das ad-hoc-Verstehen in interkultureller Kommunikation z.B. zwischen Romanophonen, die sich einer ‚fremden‘ romanischen Sprache bedienen, als auch das Erlernen einer weiteren Sprache einer Familie, von der ein Individuum bereits eine beherrscht. Es ist auch nicht falsch festzustellen: Die Kenntnis von mehr Sprachen befördert die Einsicht in sprachliche Zusammenhänge und erleichtert den Erwerb von Sprachen.

Erkannte Transferbasen verkürzen den Weg des Inputs zum Intake. In einer ‚neuen‘ Sprache identifizierte Schemata beschleunigen das Verstehen, vergrößern den verstandenen Input (*comprehensible input* im Sinne Krashens, z.B. 1995), vertiefen und erweitern die mentale Verarbeitung durch Vergleichen sprachlicher Systeme, erleichtern wiederum das Leseverstehen, erhöhen damit den potentiell neuen Input, ermöglichen und vergrößern den Kontakt mit der Zielsprache... und der Kreislauf bewegt sich weiter. Hier liegt, sehr verkürzt gesagt, der Grund für die hohe Lernwirksamkeit interkomprehensionsdidaktischer Ansätze (wie in zahlreichen deutschen und internationalen empirischen Studien dokumentiert).

Analysen zum Verlauf von Prozessen des interlingualen Identifikationstransfers belegen immer wieder, dass die mentale Verarbeitung der neuen Zielsprache beim Lesen (oder Hören) beginnt. Sie verläuft (wie nicht anders zu erwarten) von der Oberfläche der sprachlichen Form hin zu den hinter oder unterhalb dieser liegenden funktionalen und semantischen Zusammenhänge. Diese ‚bloßzulegen‘ und sie in die Systematik einer Hypothesengrammatik zu bringen ist das Geschäft des interkomprehensiven Ansatzes. Anders gesagt: Interkomprehension beginnt in den allermeisten Fällen bei dem

Verstehen (oder Nichtverstehen) von Wortschatz. Ohne die Stufe der Wortschatzverarbeitung ist verbale Kommunikation nicht möglich (Lutjeharms 2002). Soweit, so gut: Hieraus folgt, dass jede prädagogische Analyse auch die Dimension des didaktisch relevanten Wortschatzes berücksichtigen muss. Dies betrifft Fragen von Frequenz, Dispersion und Disponibilität des Vokabulars sowie die Mehrsprachigkeit der Lerner. Die folgenden quantitativen Analysen des KRM-Inventars setzen genau hier an. Ein Weiteres ist zu beachten: Weil Wortschatz in der Architektur einer Sprache sowohl den umfangreichsten als auch den dynamischsten Teil darstellt, muss sein Erwerb systematisch betrieben werden. Kasuistisches Vorgehen alleine, nur durch die Auswahl von Lehrtexten bedingt, greift zu kurz. Hier eine Übersicht und Hilfen bereitzustellen, ist eine zentrale Aufgabe des KRM. Die folgenden quantifizierenden Beschreibungen helfen hierbei.

Wer in didaktischer Perspektive über die Konstruktion von Mehrsprachencurricula nachdenkt, weil die so wichtige Rolle des lernrelevanten Vorwissens für die Lernwirksamkeit bekannt ist, kommt nicht umhin, sich zu fragen, wie Lernern schon zu einem möglichst frühen Zeitpunkt ihres Erwerbsweges von Mehrsprachigkeit jenes relevante Wissen vermittelt werden kann⁸. Genau hier haben die sog. didaktischen Wortschatze ihren Platz. Aber wie sind solche Wortschatze – hier der KRM – beschaffen? Wie einsetzbar? Was braucht es, um den relevanten Input zu identifizieren?

Die didaktische Mehrsprachenlexikographie – soll sie denn mehr als eine von einer Sprache in mehrere Sprachen übersetzte Wortliste sein – hat nur wenige Vorbilder und Beispiele. Zu den Ausnahmen zählt Helen S. Eatons in den 1940er Jahren erstellter *English-French-German-Spanish Word Frequency Dictionary* mit dem bezeichneten Untertitel *A correlation of the first six thousand words in four single-language frequency lists*. Eaton ging es nicht um die Nutzung von interlingualer Transparenz, und die heutigen Möglichkeiten eines Rückgriffs auf die Ergebnisse der Korpuslinguistik waren nicht gegeben.

EATON	<i>feeling (sensitivity)</i>	<i>sensibilité</i>	<i>Gefühl</i>	<i>sensibilidad</i>	-	-
KRM	<i>sensibility</i>	<i>sensibilité</i>	<i>Sensibilität Empfindsamkeit</i>	<i>sensibilidad</i>	<i>sensibilità</i>	<i>sensibilidade</i>

1: Beispiel aus Eaton

Weder kannten die 1940er Jahre die heutigen Ergebnisse der Leseforschung noch der Wörterbuchgebrauchsforschung mit der Unterscheidung zwischen Rezeptions- und Produktionswörterbüchern. In diesem Sinne ist die Eaton'schen Liste aus jetziger Sicht

⁸ Dass dies Fragen des Fremdsprachenangebots, der Sprachenfolge und der Sprachcurricula betrifft, wurde an anderer Stelle dargelegt und bleibt hier im Weiteren ausgeklammert.

in verschiedenen Bereichen defizient. Weitere einschlägige Mehrsprachen-Minimalwortschätze bleiben hier unerwähnt, weil sie ihre Selektionsverfahren nicht offenlegen oder aber lediglich eine Wortliste für eine bestimmte Zielsprache in mehrere Sprachen übersetzen.

Der gemeinsame Kern dieser Fragen zielt auf die Auswahl des lexikalischen Inputs im Hinblick auf das lernerseitig vorhandene relevante Wissen und auf das für eine rasche Lektüre in der (neuen) Zielsprache notwendige Vokabular. Wer z.B. die opaken Wörter eines Kernwortschatzes Lernern frühzeitig verfügbar macht, erleichtert ihnen das rasche Lesen – wie im obigen holländischen Text die interlingual sperrigen Wörter *krant, hoger, opleiden* signalisieren: Kennt man sie (die wenig transparenten und nicht ganz einfach erschließbaren Formen), entlastet dies den Arbeitsspeicher, der sich anderen Aufgaben, darunter der Aufbau der Hypothesengrammatik, zuwenden kann.

oFk	harceler	seccare	molestar	incomodar	bother	belästigen, anpöbeln
oFk	enlèvement	sottrazione	rapto	sustracción	abduction	Entführung

2: Serie ohne Formkongruenz

RoSED	ANGOISSE	angoscia	angústia	ANGUSTIA	anguish	Angst
RoSED	ANIMAL	ANIMALE	ANIMAL	ANIMAL	animal	Tier (animalisch)
RoSED	ANNONCER	annunziare	ANUNCIAR	ANUNCIAR	announce	ansagen, annoncieren
RoSED	anticiper	anticipare	antecipar	anticipar	anticipate	antizipieren
RoSED	antique	ANTICO	ANTIGO	ANTIGUO	antique, ancient	antik, althergebracht
RoSED	anxieux	ansioso	ANSIOSO	ansioso	anxious	ängstlich, erpicht (englisch)
RoSED	APPAREIL	APPARECCHIO	APARELHO	APARATO	apparat, device	Apparat
RoSED	APPARTEMENT	APPARTAMENTO	APARTAMENTO	apartamento	apartment	Apartment

3: formkongruente Serie mit Deutsch und Englisch

FS\$IP	SCÈNE	palco	PALCO	escenario	stage	Bühne
FS\$IP	entretien	manutenzione	manutenção	entretenimiento	maintenance, upkeep	Pflege, Wartung
FS\$IP	exploiter	sfruttare	APROVEITAR	explotar	exploit	ausbeuten, ausnutzen

4: Spanisch-französische Formkongruenzen aus gemischten Serien

Vor diesem Hintergrund genauere Einsichten in Umfang und Art interlingual-romanischer Transparenz und Opazität, zu Frequenzrängen und dem mehrsprachigen romanischen Kernwortschatz zu gewinnen, ist Ziel dieses Beitrages.

3. Zählungen aus „primären“ und „sekundären Zielsprachen“ nach Formkongruenzen und Formdivergenzen

Vorab: Für eine primäre Transfersprache gilt die Definition: Eine primäre Transfersprache ist eine Sprache, die für eine weitere Sprache als starke Transfersprache benutzt wird. Dies kann im Prinzip eine Erst-, Zweit- oder Fremdsprache sein. Eine Zielsprache zweiter Ordnung oder sekundäre Transfersprache ist die nachgelernte Sprache, die unter starkem Rückgriff auf andere, einem Individuum ‚vorbekanntes‘ Sprachen erlernt wurde/wird. Natürlich können Transferprozesse mehrsprachlich und multipel greifen, so dass mehrere Sprachen als primäre oder sekundäre Transfersprachen fungieren. Zudem können Transferergebnisse bidirektional, pro- und retroaktiv wirken. Deshalb kann ein Transfereffekt durchaus eine Überformung eines großen Teils des gesamten mentalen Lexikons nach sich ziehen und dieses verändern.

Wir benutzen also den Begriff der *Zielsprache* sowohl als Synonym zu Transfersprache (oder Brückensprache) als auch zu ‚Sprache, die aktuell im Fokus des Lernens steht‘, weil das umfassende Lehr- und Lernziel *Mehrsprachigkeit* heißt. Unter dieser Vorgabe ist jede Fremdsprache (in der Fremdsprachler) ihre Kompetenz weiterbauen, eine *Zielsprache*. Denn das eigentliche übergeordnete Ziel lautet: Mehrsprachigkeit auf möglichst hohem Niveau halten.

3.1 Zu dem, was, wie, wann gezählt wird.

Grundsätzlich ist zum Einen zwischen intralingual nicht-formdifferenzierenden Komputationen (erfassbar über den Selektor X OR Xw im Feld Lx; X steht für Sprache und w für Wiederholung derselben Wortform), und intralingual [lemmatisch-formdivergenten](#) Zählungen (über X im Feld Lx) zu unterscheiden, zum Anderen zwischen intra- und interlingualen Zählungen beider Selektionsweisen. Das Repertoire des Italienischen etwa zählt intralingual 8064 formdivergente Lemmata versus 1486 Wiederholungen oder Reduplikationen ein und derselben Form. Diese treten z.B. in Fällen auf, wo eine intraserielle Polysemie (in einer der Zielsprachen) die Anlage einer neuen Serie mit einer neuen lemmatischen Verteilung notwendig macht. Ein weiteres Beispiel: Der romanisch-formkongruente Anteil (ohne englische und deutsche Entsprechungen) im KRM zählt 1226 Serien, der aus der Kohorte der intralingual-formdivergenten italienischen Lemmata jedoch nur 1051. Der Kohorte liegen eben unterschiedliche Grundmengen zugrunde. Bezieht man in den Vergleich weitere Merkmale ein, so verändert sich die Differenz mit jedem weiteren Merkmal. Halten wir fest: Wir haben es mit unterschiedlichen [Grundmengen](#) zu tun. Grundsätzlich ist zwischen erweiterten Grundmengen (aus X OR Xw) und einfachen Grundmengen (aus bloßem X) zu unterscheiden.

Wie wird gezählt? Prinzipiell bestehen Serien aus Bedeutungsadäquanzen. Neben dieser apriorischen Festlegung werden die Lemmata, wie gesagt, nach den morphologischen Kriterien von Formkonvergenz und Formdivergenz unterschieden (vgl. auch Müller-Lancé 2003: 155ff, Meißner 2016 zu morphologischen Wortähnlichkeiten). Zur Konvention der Komputation: In intralingualer Zählung liegt eine Formdivergenz schon dann vor, wenn sich die Buchstabenfolge eines Wortes verändert. Ein Beispiel: *Demokrat*, *demokratisch*, *Demokratie* werden intralingual als drei verschiedene Lemmata gezählt. Die offensichtlichen pädagogischen Nachteile einer solchen Zählung werden durch die Interligalexe kompensiert, die die Reihe *Demokrat*, *demokratisch*, *Demokratie* beispielsweise in *demo-*, *-krat* und *-kratie/démo-*, *-crate*, *-cratie/-crata*, *-crazia...* auflösen. Diese Formen verbinden die Reihe z.B. mit *Demoskopie*, *Demographie*, *plutokratisch*, *Autokrat*, *Bürokrat* u.v.a.m. sowie deren romanische Entsprechungen. Wie man sieht, besteht zwischen den Schreibungen der Ligalexe der verschiedenen romanischen Sprachen in formkongruenten Serien eine Systematik, die sich nach Typen auszählen und pädagogisch nutzen lässt.

Ein Beispiel: Die spanisch-portugiesisch-französischen Kombinationen aus *ambito-âmbito-cadre* und *ambito-âmbito-domaine* sind im spanisch-portugiesischen Teil formkonvergent (oder homonymisch), im französischen formdivergent. Interlingual gilt: Zwischensprachliche Formkongruenz toleriert morphologische Abweichungen zwischen den Lemmata verschiedener Sprachen, sofern die Abweichungen den phonologischen Regularitäten des entsprechenden sprachlichen Systems folgen. Ein weiteres Beispiel: Die Vertreter der Serie dt. *Enthusiasmus*, en. *enthusiasm*, fr. *enthousiasme*, it./pg./sp. *entusiasmo*, ru. энтузи́зм, poln. *entuzjasm* gelten als interlingual formkongruent, weil ihre unterschiedlichen Graphien im Wesentlichen den grapho-phonemischen Eigenheiten der jeweiligen Sprache folgen und sie eine gemeinsame europäische Wortgeschichte sowie ein gemeinsames Etymon (ἐνθουσιασμός) teilen (Meißner 1979).

Formdivergenz impliziert ihrerseits zwei Differenzierungen: Entweder meint sie, wie gesagt, eine Veränderung der Buchstabenfolge zwischen zwei Formen mit identischem Wortstamm (Typ: it. *repubblica*, *repubblicano*) einerseits oder Veränderung der Buchstabenfolge zwischen zwei Formen mit unterschiedlichem lexematischen Teil (*repubblica*, *monarchia*) andererseits. Die Trennung beider Fälle wird wiederum in den Ligalexen sichtbar: *Repubblica* und *repubblicana* sind im Stamm identisch, *repubblica* und *monarchia* nicht. Zwei komponierte Lexeme mit der Endung *-aggio* z.B. (*vantaggio*, *salvataggio*, *paesaggio*, *passaggio* usw.) sind 'ligasuffigal' identisch. Interligalexe bezeichnen den morphologischen Kern, der Formkongruenzen unterschiedlicher Sprachen gemeinsam ist.

Die Unterscheidung zwischen den Zählweisen ist nicht nur lexikologisch relevant, sondern auch didaktisch: Formidentische oder formkongruente Wortstämme bewirken

in der Vielzahl der Serien lernpsychologisch eine Wiederholung ein und desselben Signals (mit einer besseren Wirkung auf die Memorisierung), divergente hingegen nicht. Dies bedeutet: Ist Lernen im Fokus (z.B. bei der Herstellung von Übungen), so ist in den meisten Fällen das serielle, Reduplikationen implizierende Kriterium anzuwenden; steht hingegen die Identifikation von Formen (Lexeme, Prä- und Suffixe) im Fokus, so ist zumeist eine Zählung nach formdivergenten Lemmata innerhalb der einzelnen Sprachen angezeigt.

Die Liste der Interligalexe versammelt alle Wortbildungselemente des KRM-Inventars. In der Kohorte der Formkongruenzen belegt die Liste Art und Frequenz der interlingual-orthographischen Abweichungen. Führt man diese auf einen Archetyp zurück, so entpuppt sich ihre Zahl als deutlich geringer als die der Lemmata. Betroffen sind Lexeme in 8700 Serien bzw. ca. 38280 Lexeme! Lerner, die mit diesen Elementen vertraut sind oder sie identifizieren können, verfügen über ein hohes Maß an für das Leseverstehen in mehreren romanischen Sprachen notwendigen Ressourcen – durchaus übrigens auch für das Hörverstehen (Baqué & Estrada 1999; Meißner & Burk 2001; Jamet 2007; Meißner 2010; Murillo Puyal 2011).

Im Einzelnen verlangt die Klassifikation von Wortähnlichkeiten eine Vielzahl von Entscheidungen, deren Grundlagen hier nicht explizit gemacht werden können (vgl. [Methode](#)).

Eine letzte Bemerkung: Die Datenbank des KRM zählt ca. 9555 Serien, 38220 romanische Lemmata und 57330 Lemmata insgesamt. Ein solcher Umfang von Lemmata bedarf selbstverständlich ständiger Pflege. Die hiermit erfolgenden Verbesserungen, manchmal Ergänzungen, erklären neben anderen Gründen (vgl. [Abb. 4](#)), weshalb identische Selektionen in sehr geringem Umfang (<1%) voneinander abweichende Ergebnisse erbringen können.

3.2 Ziele der Zählung

Die folgenden Tabellen greifen auf pädagogisch relevante Merkmale der KRM-Datenbank zu und weisen entsprechende Merkmalsanteile nach Sprachen differenziert aus.

Sie behandeln die folgenden lexikodidaktischen Aspekte zum Kernwortschatz:

1. Welches und wie umfangreich ist der Anteil des interlingual-transparenten, im Sinne des Mehrsprachenwachstums nutzbaren sowie des interlingual opaken Wortschatzes in den vier zielsprachlichen Inventaren des KRM?

2. Welches und wie umfangreich sind die Anteile der Formkongruenzen bzw. der Formdivergenzen zwischen *bestimmten* romanischen Sprachen, z.B. zwischen Französisch und Italienisch oder Portugiesisch und Spanisch usw. (zu den KRM-basierten Apps: Meißner 2018)?
3. Erfassung der vom englischen Wortschatz gelieferten Formkonvergenzen bzw. Transferbasen innerhalb des romanischen Kernwortschatzes.

Im Unterschied zu den vier romanischen Zielsprachen sind die ‚Korrespondenzwortschätze‘ des Deutschen und Englischen nicht erstrangig nach dem Kriterium des Frequenzrangs ausgewählt⁹. Stattdessen spielt neben der Bedeutungsadäquanz innerhalb der Serie die Formkongruenz der englischen Entsprechung eine Rolle, da sie den Identifikationstransfer und das Behalten der Wörter erleichtert. Zudem indizieren die Selektoren RoSE* und RoSED*, dass die so gezeichneten Lemmata auch in weiteren europäischen Sprachen, etwa den germanischen oder slawischen, formkonvergent auftreten.

4. Messung von Frequenzrängen in Kombination mit intra- und interlingualer Opazität und Transparenz.

Fragen zur Methodik der Komputation von lexikalischen Transparenzgraden und von Frequenzrang-Intervallen

5. Entwicklung eines statistischen Verfahrens zur Messung des Mehrsprachenwortschatzes im Sinne der Punkte 1 bis 5.
6. Entwicklung und statistische Überprüfung eines Verfahrens zur Vermessung interlingualer morphologischer bzw. morphosemantischer Transparenz und Opazität nach fünf Transparenzstufen.
7. Entwicklung eines Verfahrens zur Identifikation von Transferbasen und Transferrichtungen in gemischten Serien.

Nach Ausweis der interlingualen Transferforschung ist es nicht unerheblich, in welchem Kontext ein Transferimpuls erfolgt. Neben situativen Faktoren spielt hierbei die augenfällige Nähe einer gerade mental verarbeiteten Form zu einer interlingual zu identifizierenden Form eine Rolle. Bestimmte Markierungen

⁹ Die Zusammenstellung der etymologischen Korrespondenzen, insbesondere des Lateinischen, weicht hiervon ab, da es ja darauf ankommt, die Entwicklung der morphosemantischen Transferbasen nachzuzeichnen. Frequenzränge spielen nur soweit eine Rolle, als dies das zu vermutenden Schülern und Nicht-Fachstudierenden bekannte Inventar betrifft.

innerhalb der Datenbank können solche Relationen von seriell partiell formkongruenten oder -divergierenden Lemmata erfassen.

Der Anwendungsbezug der Selektionen besteht in der pädagogischen Verwaltung des Wortschatzes, d.i. die Selektion, Präsentation und Bereitstellung der jeweiligen Lemmata-Auswahl für die Konstruktion von Übungen und Aufgaben des nach Sprachen, Transparenz und Opazität kombinierbaren Kernwortschatzes.

3.3 Zählungen *ohne* Differenzierung nach sekundären (oder nachgelernten) Zielsprachen

Alle Komputationen beziehen sich *ausschließlich* auf die Kernwortschätze der romanischen Sprachen. Diese fungieren entweder als Transfer- oder Brückensprachen (primäre Zielsprache) oder als Zielsprachen für einen interlingualen Transfer. Eine solche Transfer- oder ‚Zielsprache erster Ordnung‘ kann z.B. „Französisch als Fremdsprache in dem einem Lerner bekannten Umfang“ sein: Deutschsprachige Lerner des Portugiesischen, die über Französischkenntnisse der GeR-Niveaustufe B1-(plus) oder höher verfügen, werden den im KRM verzeichneten französischen Wortschatz oft in Kombination mit dem Englischen und dem deutschen Bildungswortschatz als Transferbasen für die neue Zielsprache Portugiesisch (sekundäre Zielsprache) nutzen können¹⁰. Daneben verfügen erwachsene Lerner selbstverständlich über sog. Interlexeme oder Kognaten aus ihrer Muttersprache und zumeist auch über die des englischen Kernwortschatzes¹¹; sehr oft auch weiterer Sprachen.

Zusammenfassend heißt dies: Die Werte der Tabellen „ohne Differenzierung nach sekundären Zielsprachen“ erlauben eine lexikostatistische Aussage zur Reichweite und zum überhaupt möglichen Nutzen eines einzelsprachlich-romanischen Kernwortschatzes als Transferbasis für weitere romanische oder andere indoeuropäische Sprachen bzw. der zuvor- oder nachgelernten Sprachen Englisch und Deutsch¹².

3.4 Zählungen *mit* Differenzierung nach sekundären Zielsprachen

Diese Komputationen erfassen die Transferbasen, welche die Kombinationen von Zielsprachen erster Ordnung für bestimmte, später erlernte romanische Zielsprachen

¹⁰ Natürlich sind Transferbasen nicht nur im Bereich des Wortschatzes, sondern in der gesamten Architektur einer Brücken- oder Zielsprache anzutreffen.

¹¹ Empirische Erfahrungen mit Interkomprehensionsunterricht an deutsche Schulklassen verschiedener Altersstufen belegen, dass sich der sog. Bildungswortschatz nicht zuletzt mit dem Unterricht der verschiedenen Sachfächer stark herausbildet, und zwar in der Sekundarstufe I. Jüngere Lerner verfügen über diesen nur in Ansätzen.

¹² Vgl. hierzu die Literatur zu EuroComGerm, DaFnE (Hufeisen 2006).

(zweiter Ordnung) bieten; z.B. Italienisch nach Französisch. Eine nach Sprachen getrennte Zählung ist notwendig, da etwa das Transferpotenzial von Spanisch für das nachgelernte Portugiesisch ungleich größer ist als das jeder anderen, im KRM inventarisierten primären Transfersprache. Da in einfachen Zählungen die Transferbasen von jeweils zwei Sprachen in beide Richtungen transparent oder opak sind, sind die ausgezählten Mengen gleich groß. Es ergeben sich folgende mögliche Zähl- bzw. Sprachkombinationen ([Kombinierbarkeit:Sprachen](#)):

Französisch mit Italienisch
Französisch mit Portugiesisch
Französisch mit Spanisch
Italienisch mit Portugiesisch
Italienisch mit Spanisch
Portugiesisch mit Spanisch

Der pädagogische Nutzen dieser Zählungen ist augenfällig: Die Komputationen geben zusammen mit den Selektoren Hinweise darauf, in welchem Umfang und wie sehr bestimmte Sprachkombinationen aus romanischen Zielsprachen erster und zweiter Ordnung den Zugang zu den entsprechenden nachgelernten Sprachen erleichtern. Sie helfen Konstrukteuren von Lehrmaterialien, Lehrerinnen und Lehrern sowie selbstständig Lernenden bei der Auswahl eines lernökonomisch optimalen Inputs zur Beschleunigung des Spracherwerbs.

3.5 Legende und Methode der Zählungen

Sprachkronyme: DE (Deutsch), EN (Englisch) FR (Französisch), IT (Italienisch), PG (Portugiesisch) und SP (Spanisch). In der Syntax der verwandten Software treten diese Abkürzungen nochmals verkürzt auf: DE, EN, FR, IT, PG, SP.

Formkongruent sind Vertreter einer Serie mit interlingualer Ähnlichkeit der Signifikanten. *Formkongruenz* und *Formtransparenz* werden synonymisch benutzt. (Bsp.: *Polizei, police, polizia, policia...*)

Formdivergent sind **interlingual** Vertreter mit gänzlich fehlender oder sehr distanter, für den linguistischen Laien kaum erkennbarer Ähnlichkeit (Bsp.: it. *dare* und dt. *geben*).

Als **intralingual** formdivergent gelten in der Zählung alle Formen, die sich nach einem Merkmal (oder mehr) unterscheiden: fr. *dû* und *dûment* werden als zwei verschiedene Formen gezählt.

Bedeutungsadäquat sind die interlingualen Bedeutungsäquivalente einer Serie. Man kann auch von Intersynonymen sprechen.

Semantische Distanz, abgekürzt „SD“: Bezeichnet Bedeutungsunterschiede unterhalb der Polysemiegrenze **innerhalb der Serie** werden in *gravierend/nicht-gravierend* unterschieden (1 oder 2).

Serie bezeichnet eine Anreihung bedeutungsadäquater Wörter verschiedener Sprachen. Das KRM-Repertoire besteht aus Serien.

Reihe fasst eine Kategorie von Selektoren, z.B. zur Ermittlung von Transparenz oder Opazität zusammen: RoS-, RoSE-* -Reihe (s.u.).

Profilwort: ein formdivergentes Lemma, das innerhalb einer Serie in nur einer einzigen Sprache auftritt. Ein Profilwort ist seriell opak: *beaucoup – molto – muito – mucho – much*.

Transferbasis: Wissensressource, die die Realisierung eines Transfers erlaubt. Die T. für einen interlingualen Identifikationstransfer stellt ein zumindest in zwei Sprachen präsent (und einem Individuum mental zugängliches) Element dar (s. auch Transferpotenzial).

Serialisierung: Die Anordnung des Materials geschieht in Serien bedeutungsadäquater Lexeme. Polysemien innerhalb der einzelnen Zielsprachen als auch zwischen denselben führen zu lemmatischen Redundanzen. Um diese kenntlich, komputierbar und vermeidbar zu machen, werden pro Zielsprache eigene Rubriken (LF, LI, LP, LS) eingeführt. („L“ steht für *lingua*).

Einzelzielsprachlich betrifft eine einzelne Zielsprache.

Zwischensprachlich oder **interlingual** meint ‚auf mehr als eine Sprache bezogen‘.

Formkongruenz kann z.B. einzelzielsprachlich auftreten – nicht nur in Fällen von Synonymie (dt. *Streichholz* und *Zündholz*; fr. *la mousse* [1. Moos, 2. Schaum, 3. Schaumkrone], *le mousse* [Schiffsjunge]), sondern auch interlingual (fr. *mousse/pg. moço* [junger Mann]). Die Bedeutungen ‚Schiffsjunge‘ und ‚junger Mann‘ sind einander semantisch ähnlich (da Schiffsjungen in der Regel junge Männer sind). Sie sind eingeschränkt bedeutungsadäquat. Das Portugiesische hält für *Schiffsjunge* wie das Spanische *grumete* parat, das Italienische hingegen die Formkongruenz *mozzo*.

Kennungen oder Selektoren der Transferstufen in der Rubrik ROS¹³:

RoS serielle romanische Transparenz, z.B.

abattre	abbattere	abater	abatir
---------	-----------	--------	--------

5: Beispiel für serielle Transparenz

TRANSPARENZSTUFE 4 (Formkongruenz in 4 Lemmata)

¹³ Die in der Legende aufgeführten Informationen erlauben den Nachvollzug der folgenden Ausführungen. Eine weiterführende Darstellung findet sich in Abschnitt [„Methodisches“](#).

- RoSD** serielle romanische Transparenz und Deutsch
RosE serielle romanische Transparenz und Englisch
RoSED serielle romanische Transparenz von Deutsch und Englisch

ADOPTER	ADOTTARE	ADOPTAR	ADOPTAR	adopt	Annehmen adoptieren
---------	----------	---------	---------	-------	------------------------

6: Beispiel für serielle Transparenz mit Deutsch und Englisch

COMMANDER	comandare	COMANDAR	-	command	kommandieren
COMMANDER	comandare	ENCOMENDAR	encomendar	order, trust (command)	bestellen, anvertrauen, (kommandieren)

7: Beispiel für serielle Transparenz mit minimal morphologischer Abweichung

Reihe aus „RoS OR RoSD OR RoSE OR RoSED“

TRANSPARENZSTUFE 4 (Formkongruenz in 4 Lemmata)

RoS*-X serielle Transparenz mit Ausnahme der bezeichneten Sprache (RoS*-I, RoSE*-F, RoSED*-F z.B.)

TRANSPARENZSTUFE 3 (Formkongruenz in 3 Lemmata)

(dementsprechend RoSE-*, RoSED-* [RoSD-* ohne statistische Signifikanz] Da der Selektor Opazität identifiziert, errechnet sich die Transparenz aus *-* minus *-X. Da das Zeichen „-“ ausschließlich in den RoS-X-Formen verwandt wird, ist das Kürzel *-* möglich.

RoS2 zweistufige, paarige Formdivergenz

TRANSPARENZSTUFE 2 (Formkongruenz in 2x2 Lemmata)

PORTER	PORTARE	LEVAR	LLEVAR
--------	---------	-------	--------

8: Beispiel für ein T2-Serie

RosE2, RoSED2 romanische Serie mit englischer bzw. mit englischer und deutscher Transferbase bei 2 intraserialen Formkongruenzen

RoS3, RoSE3, RoSED3 dreisprachliche Formdivergenz

TRANSPARENZSTUFE 1 (Formkongruenz in 1 Lemma)

983	FENÊTRE	FINESTRA	JANELA	VENTANA
-----	---------	----------	--------	---------

9: Beispiel für eine T1-Serie

oFk ohne Formkongruenz bzw. Opazitätsstufe 4

TRANSPARENZSTUFE 0

fillette	bimba	menina	chiquilla
----------	-------	--------	-----------

10: Beispiel für eine T0-Serie

FELD: Dem KRM liegt eine relationale Datenbank zugrunde. Eine solche Datenbank hat die Struktur einer Tabelle, die aus Feldern (oder Rubriken) besteht. Wichtig ist, dass in der Rubrik ROS- nur eine einzige Kennung verwendet werden kann.

§-Markierung signalisiert sowohl eine Transparenzgrenze innerhalb einer Serie als auch eine interlinguale Formdivergenz.

FI\$PS	RoS [~]	2	AGNEAU	agnello	cordeiro	cordero
--------	------------------	---	--------	---------	----------	---------

11: Beispiel für eine intraserielle Transparenzgrenze

FI\$PS Lies: FR und IT sind zueinander formkonvergent, aber formdivergent zu PG und SP, die gegenseitig formkonvergent sind. Die Serie hat 2 Etyma.

FI\$P\$S Lies: FR und IT zueinander formkonvergent, nicht zu PG und nicht zu SP, PG ist nicht zu SP formkonvergent (vgl. die Serie zu fr. *fenêtre*)

Sortierung Die Sortierung innerhalb der §-Markierung erfolgt streng alphabetisch: FI\$PS; z.B. FS\$IP oder FI\$P\$S oder F\$IS\$P

Transparenzstufen, Selektoren und Sprachkombinationen		Zahl	
T4: RoS OR RoSD OR RosE OR RoSED	FIPS	1	
T3: *-*-RoS-F	IPS	2	
T3: *-*-RoS-I	FPS		
T3: *-*-RoS-P	FIS		
T3: *-*-RoS-S	FIP		
T3: *-*-RoSE-F	IPS		
T3: *-*-RoSE-I	FPS		
T3: *-*-RoSE-P	FIS		
T3: *-*-RoSE-S	FIP		
T3: *-*-RoSED-F	IPS		
T3: *-*-RoSED-I	FPS		
T3: *-*-RoSED-P	FIS		
T3: *-*-RoSED-S	FIP		
Kurzform: *-X			
T2: Wie "RoS2" OR Wie "RosD2" OR Wie "RosE2" OR Wie	F	I	2X2
	F	P	
	F	S	
	I	P	

"RosED2" oder in der Kurzform *2	I		S		
	P		S		
T1: Wie "RoS3" OR Wie "RosD3" OR Wie "RosE3" OR Wie "RosED3" oder *3	F	I	P		3
	F	P		S	
T0: oFk	I	S	F	P	4

Abbildung 3: Selektoren zu den T-Klassen

Innerhalb horizontaler Felder zeigen gleiche Farben Transparenzkombinationen an, der Übergang einer Farbe zur anderen eine Transfersperre. Die Ziehung des Profils erfolgte aus der Mutterdatei, gefiltert in ROS- nach den oben genannten Kennungen.

Lies: „Transparenz für x Sprachen“. Z.B.: T4 meint Transparenz für 4 Sprachen pro Serie, T2 für 2mal 2 Sprachen, T1 für 1 Sprache X mit einer weiteren Zielsprache bei gleichzeitiger 3facher Formdivergenz, ohne Formkongruenz heißt (natürlich) „ohne Formkongruenz innerhalb der Serie.“

**Hier nochmals einige Beispiele für Transparenzstufen:
(2008)T4:**

CITÉ	CITTÀ	CITADE	CIUDAD
CLAIR	CHIARO	CLARO	CLARO
CLIENT	CLIENTE	CLIENTE	cliente
collection	collezione	COLECCÃO	colección
COMBATTRE	COMBATTERE	COMBATER	COMBATIR

12: mehrserielles Beispiel für eine T4-Transparenz

5330	RoSED-F	confrère	COLEGA	colega	COLLEGA	colleague	Kollege
6896	RoSED-F	retombée	efeito	efecto	effetto	consequences, effects	Auswirkungen, Effekt

13: Beispiel für Profilwörter in Serien

T2: Zweistufige Transparenz wird durch *2 ermittelt (2x2 Formkongruenzen innerhalb der Serie).

164	FI\$PS	RoSE2	ARRIVÉE	CHEGADA	LLEGADA	ARRIVO	arrival
-----	--------	-------	---------	---------	---------	--------	---------

14: Beispiel für eine T2-Transparenz

T1: Einstufige Transparenz wird durch den Selektor *3 (3 verschiedene Lexeme innerhalb der Serie) identifiziert. Beispiel zu T1: pg. *ordenhar* und sp. *ordeñar*.

8317 F\$!\$PS	RoS3	traire	ordenhar	ordeñar	mungere	milk
---------------	------	--------	----------	---------	---------	------

15: RoS3-Serie mit \$-Spezifizierung

T0 zeigt Serien ohne interlexikalische Transferbasis.

oFk	COIFFEUR	CABELEIRO	PELUQUERO	BARBIERE	hairdresser, barber	Frisör
-----	----------	-----------	-----------	----------	---------------------	--------

16: Weiteres T0-Beispiel

* ist ein Platzhalter, der verschiedene Werte annehmen kann, auch für die Zielsprachen oder die disambiguierenden Sprachen.

X ist in Termini für Sprachen ein Platzhalter für konkrete romanische Sprachen.

Transparenzstufenschema:

T4	G	=	G	=	G	=	G
T3	U	≠	V	=	V	=	V
T2	L	=	L	≠	M	=	M
T1	O	=	O	≠	T	≠	P
T0	A	≠	G	≠	H	≠	N

Tabelle 17: Transparenzstufen nach serieller Transparenz

Die Tabelle zeigt die Relationen innerhalb der Serien nach Formkongruenz und Divergenz. Gleiche Buchstaben signalisieren interlingual transparente Formen. Gezählt wird nach Wortstämmen. T4 und T0 greifen eindeutig, weil T4 nur formgleiche Vertreter, T0 nur formverschiedene betrifft. T3 hat innerhalb der überwiegend formkongruenten Serie eine Abweichung und diese stellt keine Transferbasis für einen anderen Vertreter der Serie. In T2 sind immer zwei von vier Lexemen formkongruent. T1 ist demgegenüber mit nur einem einzigen Nachbarn formgleich, während zwischen diesen beiden Formkongruenten und den anderen Vertretern der Serie jeweils Formdivergenz besteht.

Anzahl der Etyma pro Serie unterstützt die Aussage zur Transparenzstufung: Ein Mehr an Etyma innerhalb einer Serie deutet auf starke Formdivergenz hin.

FaX (X für F, I, P oder S): ist eine Markierung, die eine leichte Formabweichung unterhalb der R-*-Schwelle kennzeichnet. Sie kann durchaus in Kombination mit dieser gesetzt werden. Beispiel für eine formkongruente Serie mit geringer Formabweichung des portugiesischen Vertreters:

PROFIL	RoS-	FR	IT	PG	SP
FaP	RoS	commander	commandare	encomendar	comandar

18: Minimale morphologische Differenzierung

Ausschluss von Doppelzählungen der Signifikanten pro Zielsprache durch Nutzung eines Zeichens „w“ in Verbindung mit F, I, P oder S zur Kennzeichnung der Wiederholung einer schon eingetragenen Form.

Anmerkung zur Zählung: Infolge der ständigen Kontrolle und Pflege des umfangreichen Datenbestandes von ca. 40.000 romanischen Lemmata und ihren deutschen und englischen Adäquanzen sind (sehr) geringfügige Verschiebungen in den Zählungen möglich. Diese liegen deutlich unter 1 Prozent und sind damit statistisch – pädagogisch ohnehin – irrelevant.

Etyma: Da die Erhebung der Etyma – zumeist lateinischen Ursprungs – und der Interligalexe noch nicht abgeschlossen ist, bleiben Vertreter beider Kategorien in den aktuellen Zählungen noch unberücksichtigt. Eine Lieferung der Daten zu einem späteren Zeitpunkt ist geplant.

Fk: Augenscheinliche Formdivergenz. Die Rubrik Fk qualifiziert die Signifikanten nach augenscheinlicher Ähnlichkeit, wie sie dem linguistischen Laien erscheint. Gezählt wird nach dem Kriterium ‚ohne Kenntnis der historischen Lautlehre als nicht-transparent erscheinende Form‘. Fk ist eine Rubrik, die vor allem Kontrollzählungen dient.

Markierung von Frequenzrängen und -klassen: Lemmata, die in den Klett’schen Grundwortschatzlisten verzeichnet sind, stehen in der Datenbank in Majuskel. Als ihre Frequenzmarkengrenze gilt <2000. Dieser Wert trennt gemeinhin einen Grundwortschatz von einem Kernwortschatz. Die Verzeichnung eines Lemmas in den Routledge Frequenzwörterbüchern geschieht dort unter Angabe des Frequenzranges, z.B.:

“**boucle** *nf* loop, buckle, curle 3822” (Lonsdale & Le Bras 2009 : 209)

Im KRM wird die Aufnahme in einem Routledge-Wörterbuch mit der Angabe <2000 durch „R“ markiert. Ein Wort in Majuskeln, das zudem die Markierung „R“ ausweist, ist somit sowohl in den Grundwortschatzlisten (<2000) als auch den Routledge Wörterbüchern (<2000) verzeichnet. Eine weitere Markierung ist „R§“ für das Routledge Frequenzintervall 2000-5000. Sie signalisiert: Eintragung in den Routledge-Glossaren (>2000), nicht aber in den Klett’schen Listen. Selbstverständlich umfasst der Kernwortschatz sowohl die Frequenzintervalle <2000 als auch <5000 (für das Italienische <7000). Liegt der Frequenzrang oberhalb dieser Marke und ist das Wort daher bei Routledge nicht inventarisiert, so wird dies durch „R-“ signalisiert. Dem entsprechend gestaltet sich die Markierung der italienischen Lemmata auf der

Grundlage von De Mauro et al. (1992). Dass die De Mauro-Liste das kernwortschatzliche Register bis zum Frequenzrang von <7000 ausweitet, macht ein weiteres Markierungsintervall erforderlich. Für das italienische Inventar bestehen mithin die Klassen: Majuskel, M (Grundwortschatz), M§ (2000-5000) und M§§ für das Frequenzintervall 5000-7000, M- >7000 (nicht in der De Mauro-Liste inventarisiert).

Transferpotenzial und Identifikationstransfer: Für die Entwicklung der Lesekompetenz in mehreren Sprachen ist das Inbezugsetzen eines ‚neuen‘ sprachlichen Schemas (z.B. eines Wortes) zu einem schon bekannten und verfügbaren Schema notwendig (passim auch Quian 2002; Lutjeharms 2002; 2007). Der zugrundeliegende mentale Prozess generiert einen sog. interlingualen „Identifikationstransfer“. Ein solcher Transfer wirkt sowohl ‚proaktiv‘ im Hinblick auf die Konstruktion der ‚neuen‘ Sprache als auch retroaktiv. In diesem Fall führt die Erweiterung des gespeicherten archetypischen Schemas zu einer Überformung des bisherigen betroffenen Sprachenwissens bzw. des mentalen Lexikons.

Primäre und sekundäre Zielsprachen: Eine primäre Zielsprache ist (hier) eine ‚fremde‘ Sprache, die für das Erlernen einer weiteren Fremdsprache als Brücke dient. Infolge der pro- und/oder retroaktiven Wirkung des Identifikationstransfers befinden sich beide Sprachen im Fokus der auf mehrere Sprachen bezogenen Erwerbsabsicht. Die gewählte Nomenklatur unterstreicht diese für das Wachstum von Mehrsprachigkeit so wichtige Lernmotivation.

Selektion und Komputationen werden erstrangig nach Bedeutungsadäquanz Frequenzrängen und zweitrangig nach Transparenzstufen zu Formkongruenzen und Formdivergenzen durchgeführt.

Verben und Normalisierung: Ein wesentlicher Begriff ist ‚Frequenzrang‘ (hierzu Davies 2006: 1-10). Es handelt sich um einen Wert, dem die Normalisierung der Tokens (z.B. *aller* aus *allons, vas; nombreux* aus *nombreux, nombreuse, nombreuses* usw.) vorausgeht. Gerade im Bereich der Verben benötigen Lerner aber nicht nur die Kenntnis der Infinitive, sondern auch der unregelmäßigen und oft opaken konjugierten Formen. Deshalb umfasst das KRM-Repertoire separat eine Liste der Verbformen.

3.6 Komputationen mit Access und Excel

Die zur Verwaltung des KRM-Inventars eingesetzte Software – Microsoft Access und Excel – erlaubt Komputationen in Verbindung mit der merkmalsorientierten Präsentation der Lemmata bzw. der Rubriken:

	Sprache	Feld PROFIL	Feld ROS-	RX§	T-Klasse und Selektor
<ul style="list-style-type: none"> Zählung der transparenten romanischen Serien und ggf. Extraktion der einzelsprachlichen Lemmata nach Transparenzstufe 4 und weiterer didaktisch relevanter Indikatoren 	Ohne Angabe				T4: RoS OR RoSD OR RoSE OR RoSED
<ul style="list-style-type: none"> Identifikation und Zählung des exklusiv romanischen Anteils 	Ohne - Angabe		X		RoS
<ul style="list-style-type: none"> Zählung der romanischen Serien mit einer englischen oder deutschen Transferbasis 	Ohne Angabe		X		RoSE OR ROSED OR ROSE2 OR ROSED2 OR ROSE3 OR ROSED3
<ul style="list-style-type: none"> Identifikation und Zählung dreistufiger Opazität 	F, I, P, oder S		X		T3: RoS*-X ¹⁴
<ul style="list-style-type: none"> Identifikation und Zählung dreistufiger Transparenz 			X		*-* - *-X
<ul style="list-style-type: none"> Identifikation und Zählung zweistufiger sprachenpaariger Transparenz 			X		*2
<ul style="list-style-type: none"> Identifikation und Zählung einstufiger Transparenz 			X		*3
<ul style="list-style-type: none"> Identifikation und Zählung von seriellen Opazitäten oder 0-Transparenzfällen 			X		oFk
<ul style="list-style-type: none"> Identifikation und Zählung der romanischen Serien und Lemmata nach drei, für das Italienische vier Frequenzintervallen, u.a. zur Trennung zwischen dem Grund- und dem Kernwortschatz (<2000 				X	R, R§, R- oder M, M§, M§§, M- MAJUSKEL

¹⁴ X ist ein Platzhalter für Sprache, z.B. für Spanisch.

Vermessung des romanischen Kernwortschatzes

[GWS]; <5000, <7000 nur für Italienisch, >5000 bzw. <7000)					
<ul style="list-style-type: none"> Einschätzung und Kategorisierung der semantischen Distanzen der Lemmata innerhalb ein und derselben Transparenzstufe 		SD; FaX, \$-Markierungen			
<ul style="list-style-type: none"> einzelzielsprachliche und sprachvergleichende Zählungen nach romanischen Zielsprachen, deren Frequenzränge, Transparenzstufen und in unterschiedlichen Kombinationen 					T3: z.B. I- OR Iw- Sortierungen in Verbindung mit F- OR Fw; oder einfach F mit P
<ul style="list-style-type: none"> komplette intraserielle lemmatische Zählung. Sie kann einzelzielsprachlich und sprachvergleichend eingesetzt werden. 	F, I, P oder S				„I OR Iw“
<ul style="list-style-type: none"> Identifikation und Zählung formdivergenter Lemmata innerhalb der Inventare der einzelnen Sprachen (zum Ausschluss von wiederholt auftretenden Formen in kombinierten Zählungen) 	F, I, P oder S				z.B. „S“ für SP
<ul style="list-style-type: none"> Identifikation der Interligalexe und Erstellung einer Liste 	Eigene Rubrik				Hier nicht aufgeführt
<ul style="list-style-type: none"> Identifikation der Etyma aller zielsprachlichen Lemmata (nicht zuletzt zur Solidifizierung der morphologischen und semantischen Trennschärfen innerhalb 	Eigene Rubrik				Hier nicht aufgeführt

von Serien und zu ihrer Kategorisierung.					
--	--	--	--	--	--

19: Funktionen der Datenbank mit Syntax der Selektoren

Pädagogische Ziele und praktische Anwendungen (s. Ziele des KRM und der Zählung)

- Optimierung der Inputselektion im Unterricht romanischer Sprachen und der folgenden didaktischen Steuerung, insbesondere im Bereich der Konstruktion von Sprachcurricula.
- Erleichterung der ‚Zuwegung‘ zum Mehrsprachenerwerb (romanische Sprachen mit Englisch und/oder Deutsch).
- Erleichterung des Lernens und Lehrens von Sprachlernkompetenz durch Bewusstmachung interlingualer morphosemantischer Konvergenzen und Divergenzen bzw. morphologischer Regularitäten (vgl. Klein & Stegmann 1999)
- Erstellung einer Liste von Interligalexen aus Beständen des KRM (in Arbeit).
- Bereitstellung der Grundlagen für die Zusammenstellung von Lern-Apps.

3.7 Triangulation der Zählungen

Selbstverständlich liegt in der Zuordnung der Lemmata in Frequenzstufen und Bedeutungsadäquanzen in den Schwestersprachen, dem Englischen und dem Deutschen immer auch ein subjektives Moment. Um dieses klein zu halten, geschah neben der erwähnten Festlegung von Kriterien- und [Rangfolgen für Übersetzungsentscheidungen](#) eine Überprüfung (erneute [Zuordnung](#)) durch zwei weitere Personen. Dabei hatten diese die Lemmata der Serie bestimmten Kriterien zuzuordnen; darunter auch die Zahl der Etyma, die einer romanischen Serie zugrunde lagen. Hier galt, wie gesagt: Je mehr Etyma desto größer die Formdivergenz. Die Zahlenangaben sollten natürlich, wenn möglich, mit den Zuweisungen in Transparenzklassen korrelieren, d.h. die Lemmata pro Serie sollten möglichst derselben Frequenzstufe entsprechen.

Der Wert „4 (Etyma)“ legt eine volle serielle Intransparenz nahe, die regulär in dem Kürzel T0 aus „oFk“ in ROS- erfasst wurde. Die Angabe „3“ im Feld ETY entsprach der Zuordnung „RoS*3“ (T1); „2“ der Transferstufe T2, die entweder aus „RoS*-*“ (Formdivergenz 1:3) oder „RoS*2“ (paarige Formkonvergenz) komponiert wurde. Die folgende Tabelle zeigt Übereinstimmungen und Abweichungen für die einfach komponierten Stufen T4, T1 und T0:

Transparenzkl. und Selektor	Fälle	Zahl d. Etyma	Fälle aus ETY	Abweichung
T4 (RoS, RoSD, RoSE, RoSED)	5606	1	5812	3,5%

Vermessung des romanischen Kernwortschatzes

T1 (RoS*3)	1049	3	1114	5,9%
T0 (oFk)	854	4	812	5%

Triangulation zwischen Transparenzklassen und Variablen aus ETY

Wie erklären sich die Abweichungen? Kurz und knapp damit, dass das Kriterium für die Zuordnung die „augenscheinliche Formähnlichkeit“ war, wie sie bei Schülern oder Lernern vermutet wurde. So wurden Serien wie *poche-borsa-bolso-bolsillo* 3 Etyma zugeordnet, obwohl laut Dauzat et al. (1971) nur vlat. BŮRSA und fränz. *POKKA möglich sind.

Natürlich führen die Abweichungen in begründeten Fällen zu einer entsprechenden Revision des Korpus.

4. Vorstellung der Selektionen und ihre Quantifizierung im Überblick vorab

Die folgenden Tabellen und Abbildungen erlauben die Bildung eines ersten Überblicks über die Klassifikatoren und Selektoren des KRM-Inventars zur Erfassung von interlingualer Transparenz und Frequenzrängen sowie den entsprechenden Wortschatzanteilen.

Komputationen nach Selektoren ohne einzelzielsprachliche Spezifizierung

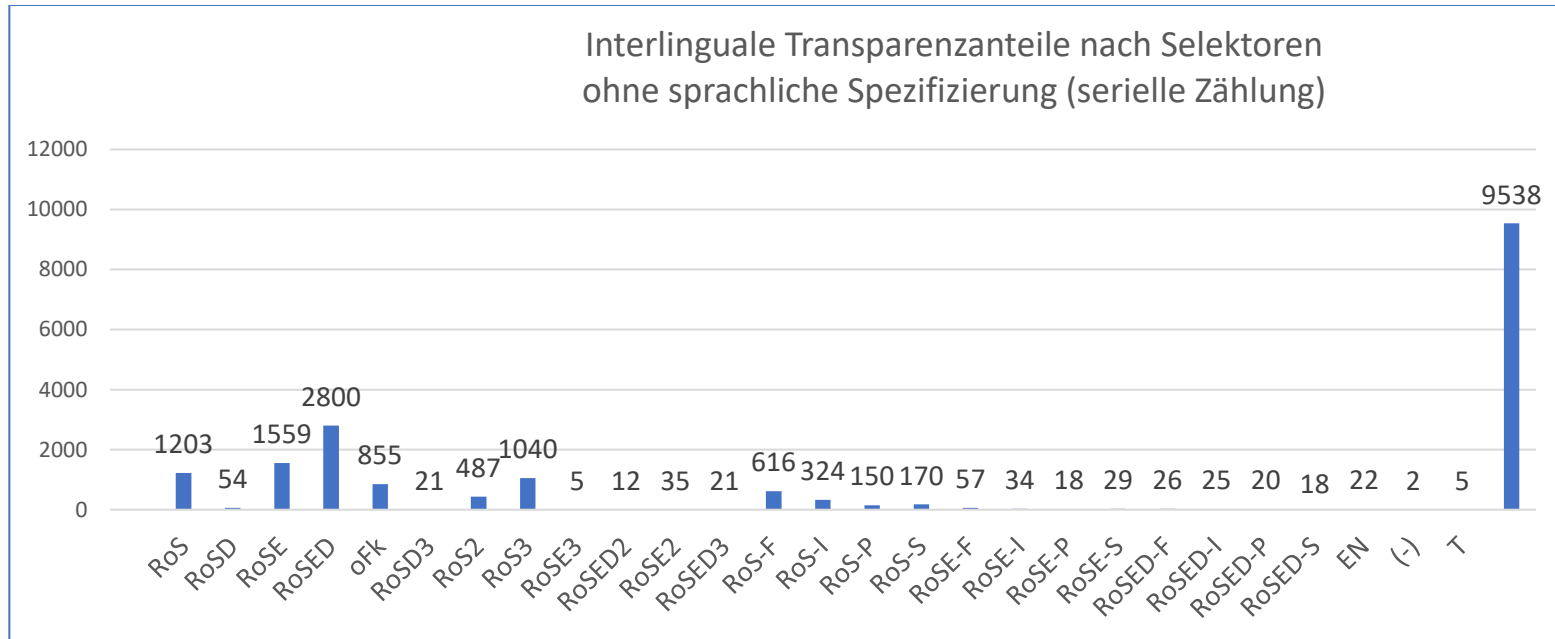


Abbildung 4: Alle Markierungen im Überblick aus dem Gesamt des KRM-Inventars abzüglich der "Spezifizierung

Die Werte für \$-Markierungen (s.u.) sind nicht in die folgenden Zählungen zu nehmen, da es keine \$-Markierung ohne Ko-Markierung in der Rubrik ROS- gibt. Geringfügige Unterschiede in der Anzahl der Serien (9538/9557) erklären sich mit der Aufnahme von englischen Wörtern in den benutzten Wortlisten. Des Weiteren ist die benutzte Opposition von „T“ (Toponym) und „-“ (fehlend) im Sprachenvergleich nicht trennscharf. Gleiches gilt für Eigennamen und deren Derivate (Typ: *gaullisme*, *peronismo* usw.). Ähnliches für regional-spezifische Wörter (sp. *Porteños* ‚Einwohner von Buenos Aires‘, *Gaditos* ‚von Cadiz‘ im Spanischen; it. *napolitano* ist viel frequenter als fr. *napolitain* usw. (Datenüberprüfung am 03.01.2019)

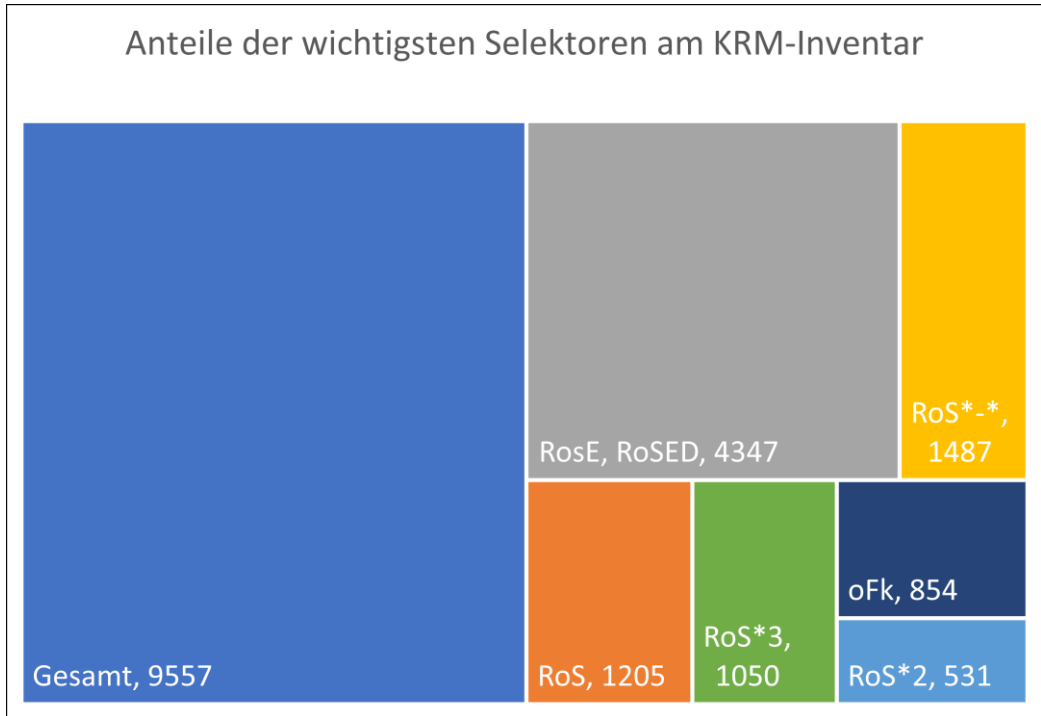


Abbildung 5: die wichtigsten Selektionen (gepr. am 24.04.2018)

Die RoSE*-Fälle schließen die RoSE2 und RoSE3-Okkurrenzen ein. Dies ergibt Doppelzählungen mit den RoS*2 und RoS*3. Da die Zahl der RoSE mit RoSED-Fälle 4348 beträgt, lautet der Bias 279 (=3% der Gesamtzahl aller Serien).

Wie die rechte Hälfte der Abb. zeigt, stellt die RoSE*-Gruppe (aus RoSE und RoSED) die größte Kohorte dar, gefolgt von der RoS-Gruppe. Dies belegt: Der Anteil der mit dem Englischen geteilten Formkongruenzen von 4347 ist mehr als dreimal so stark wie der der Panromanismen ohne Englisch. Verblüffend gering ist hingegen die Anzahl der komplett opaken Lemmata (die keine morphologische Basis für einen interlingualen Identifikationstransfer bieten).

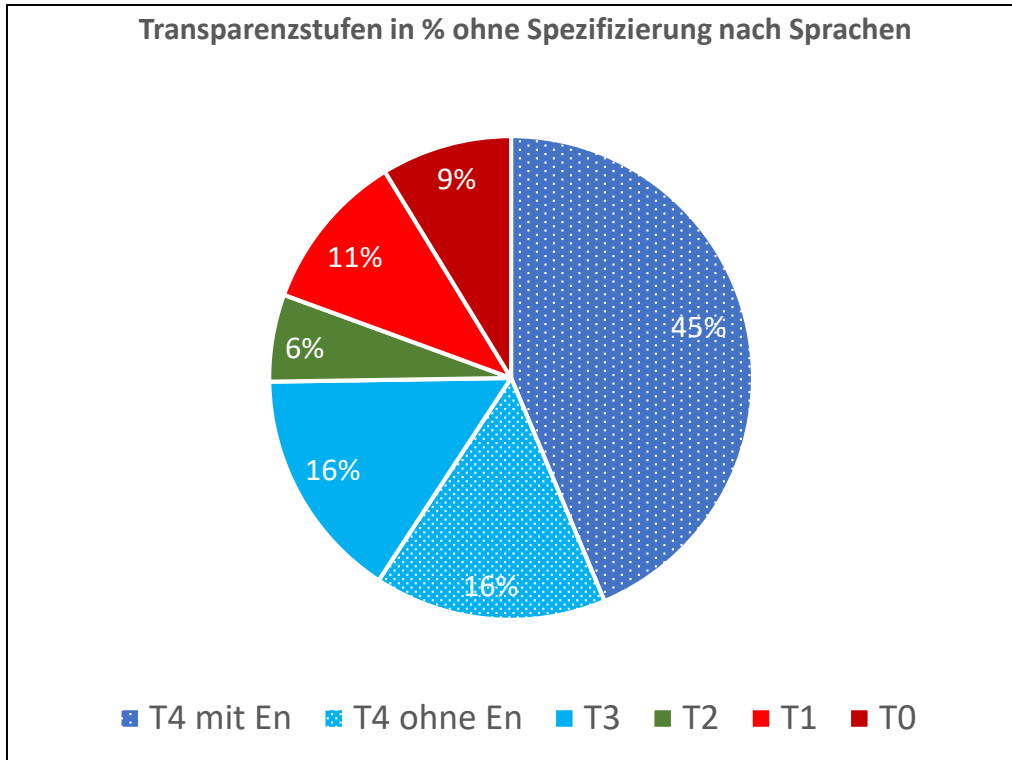


Abbildung 6: Transparenz-Kategorien in % (gerundet)

Die folgende Tabelle zeigt die Daten im Zusammenhang mit den Selektoren¹⁵.

SERIEN (9588)	Fälle	Prozent	Selektor
Transparenzstufe 4	5525	58%	RoS, RoSD, RoSE, RosED
Transparenzstufe 3	1478	15%	RoS-*
Transparenzstufe 2	534	6%	RoS*2
Transparenzstufe 1	1048	11%	RoS*3
Transparenzstufe 0	855	9%	oFk

20: (Beispiel für Selektoren mit Zählung) aus der GrdM 9557 (gerundet)

Abb. 4 hat gezeigt, dass nicht alle Selektoren zu statistisch-relevanten Kategorien führen. Deshalb stellen Abb. 5 und Abb. 6 die quantitativ relevanten Klassen nach Fällen und in Prozent zusammen.

Was sagt Abb. 5 didaktisch aus? Wenn man von einer überwiegenden interlingualen Transparenz des romanischen Kernwortschatzes spricht, meint dies grob einen Transparenzanteil von 81 Prozent (Summe aus Lemmata der Transparenzstufen 4 bis

¹⁵ Voneinander gering abweichende Werte sind rundungsbedingt.

inkl. 2). Das von romanischer Zielsprache zu romanischer Zielsprache in jedem Fall wirklich neu zu erlernende, morphologisch unbekannte Vokabular beträgt minimal 20 Prozent (Transparenzstufen 1 und 0). 6 Prozent betreffen Transparenz zwischen zwei Zielsprachen. In diesem Zusammenhang lohnt es sich, an den Grenzwert für die Interkomprehensibilität zwischen Sprachen zu erinnern: Liegt der Anteil opaker Elemente über 30 Prozent am Gesamt eines Textes, so ist die Interkomprehensionsmethode nur eingeschränkt einsetzbar.

Die Rolle der Selektoren für die Bildung von Transparenzklassen drückt sich in ihrem Anteil an der Gesamtverteilung aus, wie Abb. 7 veranschaulicht:

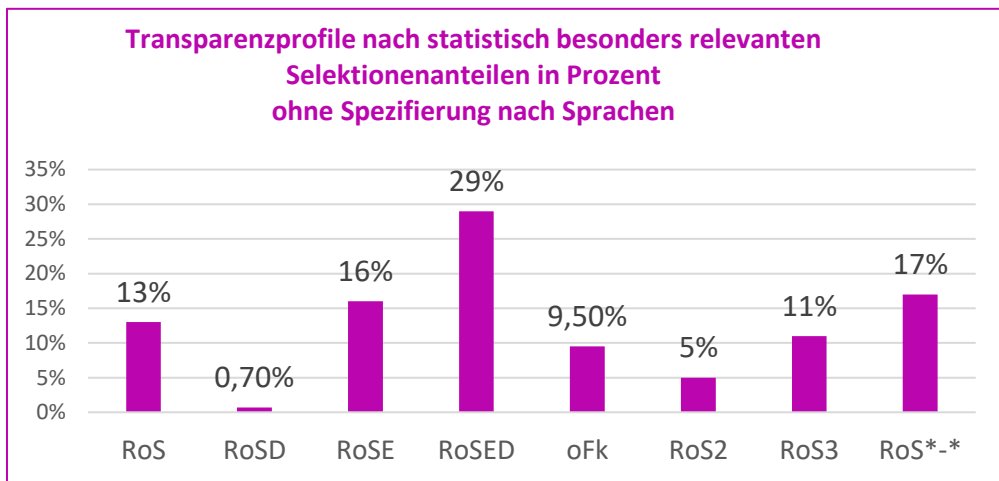


Abbildung 7: Transparenzselektionen in % ohne sprachliche Spezifizierung

Nimmt man zur Festlegung der statistischen Relevanz den kritischen Wert 100 (=1% der Gesamtzahl der Serien) an, so sind nur die Selektoren RoS, RoSE, RoSED; RoS*-* (RoS*-F, RoS*-I, RoS*-P, RoS*-S); RoS*2, RoS*3 und oFk relevant. Der vernachlässigte Wortschatzanteil beträgt 319 Fälle. d.i. 3% vom Gesamt. Dieses Bias ist geringer als das aus den Rundungen insgesamt.

5. Bemerkungen zu einzelnen Sprachen aus Sicht der Transferforschung

Es war bereits von Grund- und Kernwortschatzen die Rede. – Der Grundwortschatz, bestehend aus dem Frequenzrang <2000 in den Routledge-*frequency dictionaries* (Davies 2006; Davies & Raposo Preto-Bay 2008; Lonsdale & Le Bras 2009) bzw. den Einträgen der De Mauro-Liste von <2000 sowie der Verzeichnung in den Klett'schen Grundwortschatzlisten, umfasst folgende Anzahl an Serien pro Einzelzielsprache sowie an formkongruenten T4-Fällen sowie opaken T0-Fällen:

ZIELSPRACHE	ANZAHL	FORMKONVERGENZEN	T0-OPAKE
Französisch	3867 (40%) ¹⁶	davon T4-formkonverg. 1971 (21%)	T0-opak 389 (4%)
Italienisch	2486 (26%)	davon T4-formkonverg. 1358 (14%)	T0-opak 258 (3%)
Portugiesisch	3065 (32%)	davon T4-formkonverg. 1701 (18%)	T0-opak 309 (3%)
Spanisch	3091 (32%)	davon T4-formkonverg. 1664 (17%)	T0-opak 308 (3%)

Die folgenden Ausführungen antworten auf Fragen, die im fremdsprachendidaktischen Diskurs oft gestellt wurden. Dies betrifft sowohl die Minimalwortschätze und deren interlinguale Transferpotentiale als auch die Rolle von Englisch und Latein im Zusammenhang mit sog. Sprachenfolgen. Dass diese Fragen auf der Folie von Deutsch als Mutter- oder Zweitsprache zu betrachten ist, versteht sich von selbst.

5.1 Grundwortschatzliche Transferanteile nach Sprachen

Auszählungen auf der Grundlage der Inventare der Klett'schen Grundwortschätze ergaben als potentielle Transferpotenziale (Meißner 1989):

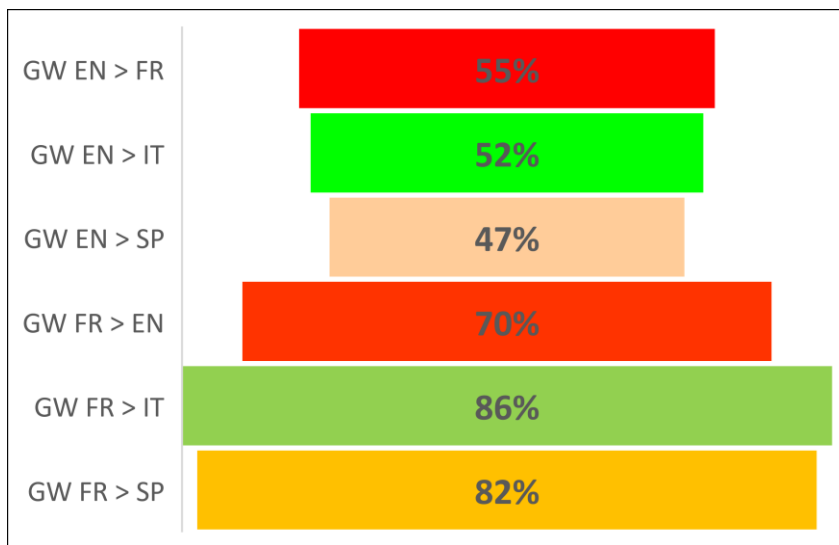


Abbildung 8: Grundwortschatzliche Transferpotentiale von Englisch und Französisch im Vergleich zueinander und zu romanischen Sprachen

Offensichtlich sind die Transparenzen aus Richtung Französisch nach Englisch deutlich größer als die von Englisch nach Französisch. Gleiches gilt erst recht von Französisch nach Italienisch und Spanisch. Dank des KRM-Inventars ist es nun möglich, die unterschiedlichen Mengen genauer zu quantifizieren.

¹⁶ Bezogen auf das Gesamt des Kernwortschatzes.

Vermessung des romanischen Kernwortschatzes

RoS-	Fr	Rf	Lfr	It	Ri	Lit	Pg	Rp	Lpt	Sp	Lsp	En	De
RoS-F	APRÈS (D ^{im})	R	Fw	SECONDO	M	Iw	SEGUNDO	R	Pw	SEGÚN	Sw	according to	gemäß, nach, laut
RoS-F	SELON	R	Fw	SECONDO	M	Iw	SEGUNDO	R	Pw	SEGÚN	S	according to	gemäß, laut
RoSE	SEMAINE	R	F	SETTIMANA	M	I	SEMANA	R	P	SEMANA	S	week	Woche
RoSE	SIGNE	R	Fw	SEGNO	M	Iw	SINAL	R	Pw	SEÑA	S	sign	Zeichen, Wink
RoSED	signal	R	F	SEGNALE	M	I	SINAL	R	Pw	SEÑAL	Sw	signal	Signal, Zeichen
RoSE	SIGNE	R	Fw	SEGNO	M	Iw	SINAL	R	Pw	SEÑAL	Sw	sign	Anzeichen, Signal
RoSE	simplement	R	Fw	semplicemente	M	I	simplemente	R	P	sencillamente	S	simply	einfach
RoS	SEIGNEUR	R	F	SIGNORE	M	Iw	SENHOR	R	P	SEÑOR	S	Lord	Herr (fr Gott)
RoS-F	MONSIEUR	R	F	SIGNORE	M	Iw	SENHOR	R	Pw	SEÑOR	Sw	mister/sir	Herr
RoS-F	MADAME	R	F	SIGNORA	M	I	SENHORA	R	P	SEÑORA	S	madam, lady	Frau, Madame (iron)
RoSE	SENS	R	F	SENSO	M	Iw	SENTIDO	R	P	SENTIDO	Sw	sense, feeling	Sinn, Gefühl, Bedeutung
RoSE	SENTIMENT	R	F	SENTIMENTO	M	I	SENTIMENTO	R	P	SENTIMIENTO	S	sentiment	Gefühl
RoS	SENTIR	R	Fw	SENTIRE	M	Iw	SENTIR	R	Pw	SENTIR	S	feel, regret	fühlen, spüre, bedauern
RoS-F	ÉPROUVER	R	Fw	SENTIRE	M	Iw	SENTIR	R	P	SENTIR	Sw	feel, experience	fühlen, spüren
RoSED	SÉPARER	R	F	SEPARARE	M	I	SEPARAR	R	P	SEPARAR	S	separate	trennen, separieren
RoS2	ÊTRE	R	Fw	ESSERE	M	Iw	SER	R	Pw	ser	S	being	Wesen
RoSED	SÉRIE	R	F	SERIE	M	I	SÉRIE	R	P	SERIE	Sw	series	Serie
RoSED	SÉRIEUX	R	Fw	SERIO	M	I	SÉRIO	R	P	SERIO	S	serious	seriös, ernst
RoSED	SERVICE	R	F	SERVIZIO	M	I	SERVIÇO	R	P	SERVICIO	S	service	Dienst, Service
RoSE	SERVIR	R	F	SERVIRE	M	I	SERVIR	R	P	SERVIR	S	serve	dienen
RoS	SI	R	F	SE	M	I	SE	R	P	SI	S	if	wenn (Bedingung)
RoS	OUI	R	F	SÌ	M	Iw	SIM	R	Pw	SÍ	S	ye	ja
RoS	SI	R	Fw	SÌ	M	Iw	SIM	R	Pw	SÍ	Sw	yes	ja
RoS-F	TOUJOURS	R	F	SEMPRE	M	I	SEMPRE	R	P	SIEMPRE	S	always	immer
RoS	SIÈCLE	R	F	SECOLO	M	I	SÉCULO	R	P	SIGLO	S	century	Jahrhundert
RoSE	SIGNIFIER	R	F	SIGNIFICARE	M	I	SIGNIFICAR	R	P	SIGNIFICAR	S	signify	bedeuten
RoS	SIGNE	R	F	SEGNO	M	I	SINAL	R	P	SIGNO	S	sign, mark, symbol	Zeichen, Markierung, Symbol
RoSE	SILENCE	R	F	SILENZIO	M	I	SILÊNCIO	R	P	SILENCIO	S	silence	Ruhe
RoSED	symbole	R	F	SIMBOLO	M	I	SÍMBOLO	R	P	símbolo	S	symbol	Symbol
RoSED	SIMPLE	R	F	SEMPLICE	M	I	SIMPLES	R	P	SIMPLE	S	simple	einfach, simpel
RoSE	SIMPLEMENT	R	F	SEMPLICEMENTE	M	Iw	SIMPLESMENTE	R	Pw	SIMPLEMENTE	S	simply	einfach
oFk	SANS	R	F	SENZA	M	I	SEM	R	P	SIN	S	without	ohne
RoSED	syndicat	R	Fw	SINDACATO	M	I	SINDICATO	R	P	sindicato	S	workers' union	Gewerkschaft, Syndikat

Abbildung 9: Grundwortschatzlicher Ausschnitt aus „R“ in Rfr, Rpg und Rsp bzw. „M“ in Rit

Doch zurück zum KRM. In Bezug auf den Bestand der einzelsprachlich-formdivergenten Lemmata des KRM baut das Englische in serieller Zählung (X OR Xw) in folgendem Umfang seriell umfassende Transferbasen (T4) für die einzelnen Zielsprachen auf¹⁷:

T4-Transparenzen mit Englisch

(F)	für Französisch:	4628 (48%) ¹⁸
(I)	für Italienisch:	4628 (48%)
(P)	für Portugiesisch	4626 (48%)
(S)	für Spanisch:	4626 (48%)

Anders gesagt: 1.) 48 Prozent der im KRM aufgenommenen Vokabeln verbinden sich im Hinblick auf die einzelzielsprachlich-französischen Lemmata mit einer Lernerleichterung (wenn Lerner des Englischen hinreichend mächtig sind). 2.) Die Kognaten des Englischen sind quasi im gleichen Umfang in allen romanischen Sprachen präsent. 3.) Natürlich gelten die lernökonomischen Effekte nicht nur für die Sprachenfolge ‚romanische Sprache nach Englisch‘, sondern auch für ‚Englisch nach einer romanischen Sprache‘ (Vgl. auch Marx 2010 zu „eag“ [English after German]. Dieser Vorteil fällt vor allem für romanophone Muttersprachler ins Gewicht oder für solche mit einer nicht-indoeuropäischen Muttersprache, die sich des Englischen bedienen, um eine weitere europäische Sprache zu erlernen.

5.2 Der englische Wortschatz als Brücke zu den romanischen Sprachen

Fraglos ist das Englische eine wichtige Transfersprache für die romanischen Sprachen generell. Es gewinnt diese Eigenschaft grundlegend, aber nicht allein aus der Qualität seines Wortschatzes (und seines Sprachbaus überhaupt): Es schuldet sie in noch größerem Maße der Zahl seiner nativen, zweit- und fremdsprachlichen Teilhaber: Sie beläuft sich global auf ca. 359 Mio. Muttersprachler und ca. 1,5 Milliarden sekundärer Sprecher. Die Zahl derer, die Englisch von Zeit zu Zeit als Fremdsprachler benutzen, dürfte noch höher anzusetzen sein. Damit muss Englisch als eine wirksame Transfersprache zu weiteren europäischen Sprachen gelten, nicht nur zu den romanischen (Phillipson & Skutnabb-Kangas 1999; Gnutzmann 2005). Dies gilt nicht zuletzt für native Sprecher nicht-europäischer Sprachen, die über Kenntnisse des Englischen als Fremd- oder Zweitsprache verfügen, wie etwa der DaFnE-Ansatz zeigt (Hufeisen 2006; vgl. zu eag. Marx 2010).

¹⁷ Syntax: *E*" in Kombination mit z.B. Lfr=F OR Fw). Ohne sprachliche Spezifizierung ergibt *E* 4649 Fälle.

¹⁸ Zahl der intralingual-formdivergenten Lemmata in der Rubrik FR: 7639; in IT: 8064; in PG: 7758; in SP: 7591.

Die größte Selektion bilden, wie gesagt, die RoSE- und RoSED-Eintragungen. Sie besteht aus Interlexemen¹⁹ unterschiedlicher Kategorien wie Kultismen (*Humor, extase, grammair, Medikament, география* (Geographie), *talento* ...), Szientismen und Neohellenismen und Neolatinismen oder hybriden Bildungen (*oxydation, oncologie, scalpel, chlore* < gr. *khlôros, scander, Liberalismus, émancipation, Komintern*), Wörtern des modernen Lebens (*Scotch, libéral/Liberalismus, rock'n'roll, chewing gum, chocolate, автомобиль* [Automobil], *Massenmedian, mass media, Internet* usw.), Bedeutungs-exotismen (dt. *Samovar* von ru. *самовár*, dt. *Ikebana* von jap. フラワー・アレンジメント²⁰ usw.). Romanische Serien sind demgegenüber von eingeschränkter Verbreitung (z.B. *arbre, albero, árvore, árbol*). Wie durchlässig das Kriterium ‚nur in romanischen Sprachen existent‘ ist, zeigt etwa die Serie *an/année, anno, ano, año*, denn *annual* (*annuel, annuale, pg./sp. anual* usw.) ist durchaus im Englischen geläufig und als ursprünglich lateinisches Wort auch Sprechern germanischer und auch slavischer Sprachen erschließbar (dt. *Annuität*, ru. *аннуитёт*, schwed. *annuitet* oder ru. [абсолютный](#), pl. *absolutny*, dän. *absolut*, vgl. Meißner 1993). Damit sind die sog. Bildungswortschätze der germanischen und slavischen Sprachen angesprochen, deren Repertoire ihren kompetenten Sprechern ungezählte Transferbasen zur Verfügung geben.

Ein Wort zur Zählung der englischen (und deutschen) KRM-Anteile, um Missverständnisse zu vermeiden: Während, wie erwähnt, der Bestand der romanischen Zielsprachen im KRM aus entsprechenden Minima-Listen oder Frequenzglossaren kompiliert wurde, ist dies für die deutschen und englischen Lemmata nicht der Fall. Sie erreichten das KRM-Inventar als Übersetzungen bzw. Bedeutungsadäquanzen von romanischen Lemmata. Maßgeblich für die Aufnahme der deutschen und englischen Lemmata waren erstrangig die Bedeutungsadäquanz mit der romanischen Serie, zweitrangig eine mögliche Formkongruenz und erst drittrangig die Verzeichnung in Jones & Tschirner (2006) bzw. in Davies & Gardner (2010)²¹. Anders gesagt: Die Verzeichnung der englischen oder deutschen Lemmata im KRM bildet nicht die Kernwortschätze dieser Sprachen ab, denn diese müssten selbstverständlich aus deren jeweils eigenem lexikalischen Repertoire generiert werden. Entsprechende Extra-

¹⁹ Wir ziehen diesen Ausdruck der gängigen Bezeichnung ‚Internationalismus‘, der im strengen Sinne nur sog. ‚Europäismen‘ – auch dieser Terminus ist umstritten (Muljačić 1979) – fasst, vor, weil er politisch unkorrekt und überdies ungenau ist.

²⁰ Unterschiedliche Schriftsysteme setzen der lesebasierten Interkomprehension natürlich erhebliche Hürden entgegen. Dies trifft umso mehr zu, wenn diese Systeme ideographisch (wie das japanische Kanji) funktionieren. Solche Graphien repräsentieren Bedeutungen (Signifikate). Die dazu gehörigen Wortformen (Signifikanten) müssen ‚zusätzlich‘ erlernt werden. Die japanischen Hiragana- und Katakana-Schriften übersetzen hingegen Laute in Schrift.

²¹ Selbstverständlich wurden alle französischen, italienischen, portugiesischen und spanischen Lemmata mit einem Frequenzrang <5000 bzw. <7000 (Italienisch) aufgenommen.

polationen sind deshalb nicht möglich. – Welches und wie groß sind nun die zwischen den romanischen Sprachen und dem Englischen liegenden Potenziale für einen Identifikationstransfer?

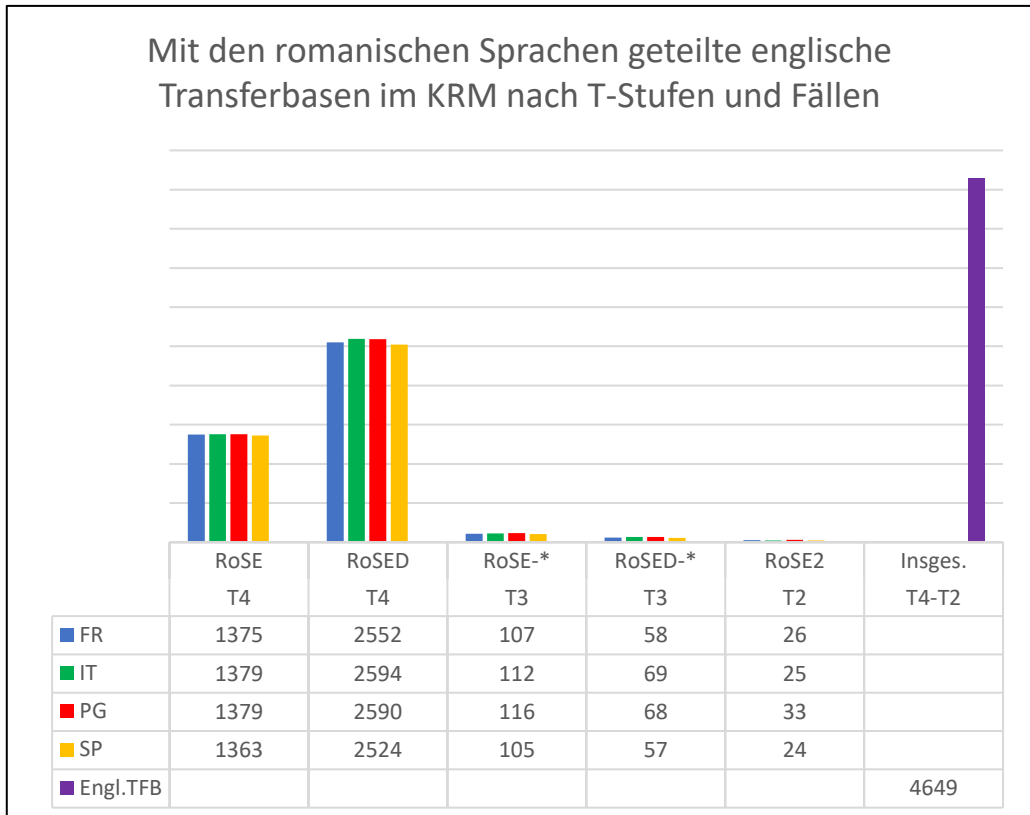


Abbildung 10: Englische Transferbasen in Französisch, Italienisch, Portugiesisch, Spanisch (aus X)²²

Der Anteil englischer Transferbasen in den romanischen Kernwortschätzen beläuft sich in den einzelnen romanischen Sprachen auf jeweils 44 bis 45 Prozent (als Divisor genommen wird die Zahl der Serien [9558]). Wie man in Abb. 11 sieht, umfassen die englischen Transferbasen quasi in gleichem Maße alle vier romanischen Sprachen.

Die folgenden Prozentwerte sind aus der einfachen Grundmenge von X (ohne Reduplikationen) generiert. Der Anteil der englischen Transferbasen im KRM erreicht in dieser Zählung folgende Prozentwerte:

²² Mit nur 5 Okkurrenzen fällt RoSE3 unter die Relevanzgrenze, was seine Nichtaufnahme in die Abb. erklärt.

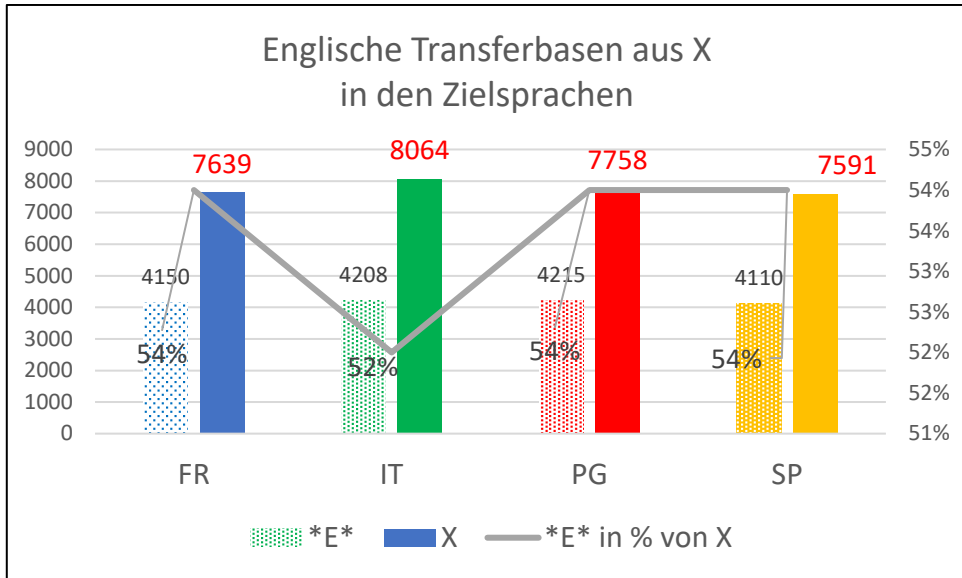


Abbildung 11: Fälle und Prozentwerte englischer Transferbasen aus (X)

Lies: Die Grundmenge zu FR aus X-Zählung umfasst 7639 Okkurrenzen; der Anteil mit englischen Formkonvergenzen 4150. – Die graue Farbe in der Legende ist ein Platzhalter für die Prozentangaben zu den verschiedenen Sprachen.

Der Anteil des englischen Transferpotentials ist keine *quantité négligeable*. Daher verwundert es nicht, wenn die Didaktik das synergetische Potenzial des Englischlernens für die Mehrsprachigkeit längst entdeckt hat (z.B. Klein & Reissner 2006; Le Besnerais 2011; Jakisch 2015). Ein weiterer Vorteil des Englischen ist zu nennen: Dies betrifft nach Ausweis der empirischen Lernerforschung das hohe Maß, in dem es dem Englischunterricht gelingt, die Motivation der Lerner auf hohem Niveau zu halten (Androulakis et al. 2007; speziell in Bezug auf Deutschland: Meißner et al. 2008) (zur Geschichte des englischen Wortschatzes: [s. unten](#))

Die folgende Abbildung zeigt detailliert die Transferbasen des Englischen in den verschiedenen romanischen Sprachen und den Anteil der einzelnen Selektoren:

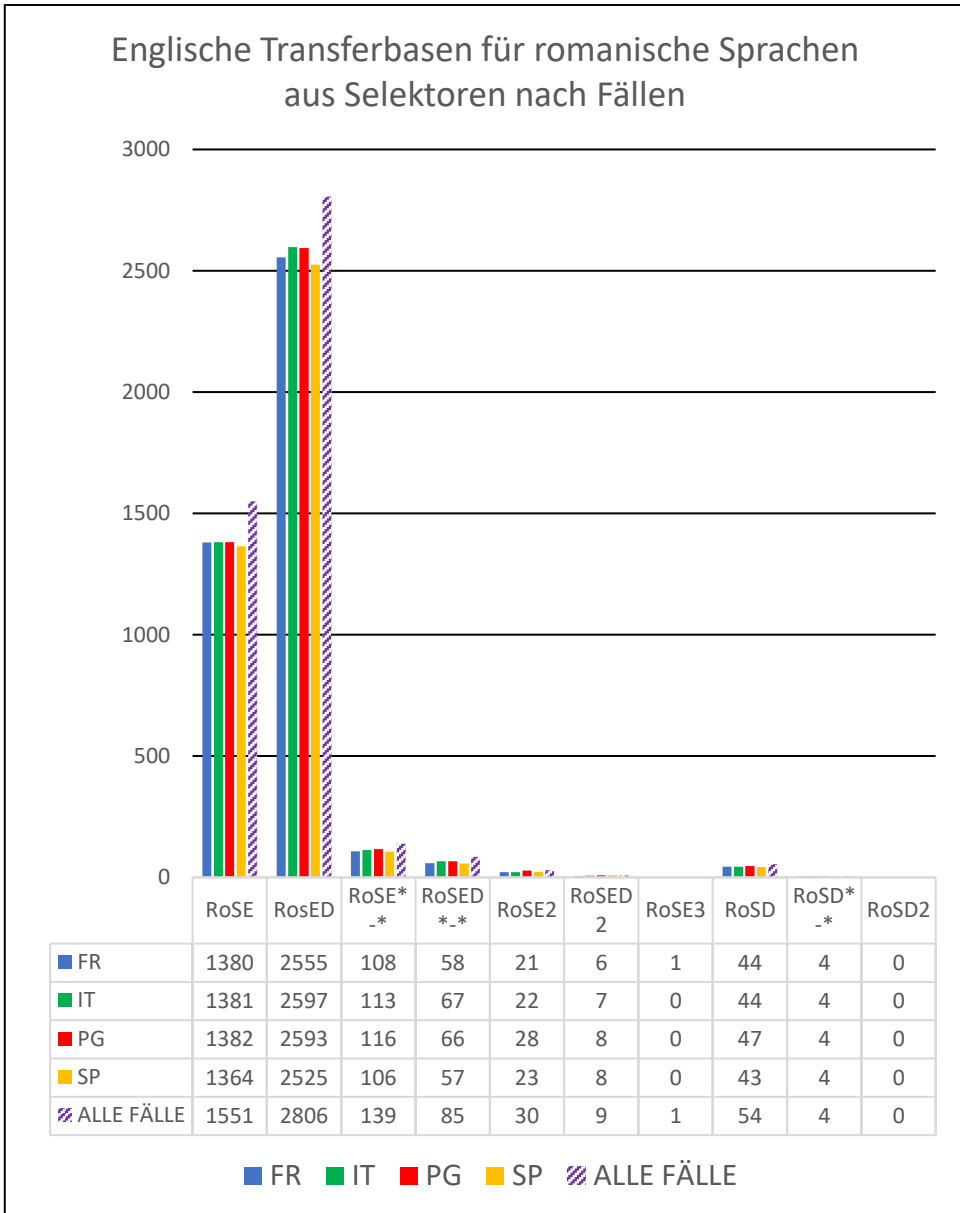


Abbildung 12: Transferbasen mit Englisch und Deutsch in Fällen und %
(Zählung 01.12.2017)

5.3 Latein als gemeinsame Mutter der Romania aus Sicht des interlingualen Transfers

„Latein ist die Mutter der romanischen Sprachen“ (und Kinder gibt es nicht ohne Mütter) – so lautet ein verbreitetes, aber im Hinblick auf die romanische Interkomprehension und das erforderliche Transferpotenzial irriges Stereotyp. Die Metapher signalisiert die Meinung, dass man Latein lernen solle/müsse, weil dann das Erlernen von Sprachen, insbesondere der romanischen, leichter falle²³. Es geht also um Transferbasen und deren Wirksamkeit für den Lernerfolg.

Über den Lernerfolg in der Domäne Fremdsprachen entscheiden freilich nicht allein linguistische Kategorisierungen von Sprachen und sprachlichen Formen (einschließlich der etymologischen Klärung von Wörtern). Vielmehr bestimmen lernerseitig Art und Umfang des lernrelevanten Vorwissens sowie die Lernmotivation weitgehend die Lernwirksamkeit. Generell gilt zu den linguistischen Ressourcen von Lernern: Nur ein solches kognitives Schema kann eine Transferbasis liefern, über das Lerner deklarativ und/oder prozedural verfügen und welches sie zielgerichtet aktivieren können. Hier spielen vor allem Analogien zu sprachlichen Formen und die an sie geknüpften kulturellen Inhalte sowie das Unterrichtserlebnis eine ganz entscheidende Rolle.

Was bedeutet dies nun im Hinblick auf das Lateinische, dessen positive Transferwirkung, obwohl oft behauptet, für das Erlernen der modernen Sprachen in der Lernforschung

²³ Auf der Homepage des *Staatsinstituts für Schulqualität und Bildungsforschung München* liest man unter der Überschrift „Transfer aus den bereits erlernten Sprachen“ (warum eigentlich nur aus diesen? Die umfangreichsten Kenntnisse haben deutschsprachige Schüler doch in Deutsch):

„Der Transfer aus den bisher gelernten Fremdsprachen und dem Deutschen spielt in den spätbeginnenden Fremdsprachen eine besondere Rolle. Für die Schüler ist die spätbeginnende Fremdsprache die 3. oder gar 4. Fremdsprache, die sie am Gymnasium erlernen. Insbesondere zwischen Latein und den romanischen Sprachen, aber auch zum Englischen und Deutschen sind vielfältige Bezüge herstellbar, die das Erlernen einer neuen Sprache erleichtern. (...) Dies gilt für das Erlernen sowohl von neuem Wortschatz (z. B. it. „miracolo“ => vgl. lat. „miraculum“; frz. „miracle“; engl. „miracle“) als auch von neuer Grammatik (z. B. die Konjugation der italienischen Verben im passato prossimo mit avere oder essere => vgl. Französisch oder Deutsch).“

(<http://www.isb-gym8-lehrplan.de/contentserv/3.1.neu/g8.de/index.php?StoryID=26940>)
(07.01.2018)

längst relativiert wurde²⁴? Die im folgenden versuchte Antwort beleuchtet allein die linguistischen Implikationen des Transfers.

Der zunächst für die Klärung dieser Frage im Kontext des Wortschatzes zu analysierende Komplex betrifft die Sprachen und ihre Register, in denen Lateinisches bzw. lateinische Wortelemente begegnen: in der (deutschen) Muttersprache als Lehn- oder Fremdwörter: *Pfirsich, Kamin, Musik, Diktat, Humor, Polizei* (aus Gr. über Lat.), *Nation, City* (aus CIVITAS über Englisch); in Zweit- und Fremdsprachen, eventuell in moderner Gestalt: *to continue, continuar, libation, liberty, liberación...*; als Elemente moderner Wortbildung (Präfixe, Infixe, Suffixe): *kontinuierlich, Kontinuität, Telephon, Telepathie, telegen*; und schließlich im Lateinunterricht (*omnium, praeter, saeculum...*). Eine von diesen Fragen zu trennende betrifft das Latein selbst: Welches Latein ist eigentlich gemeint? Das der Goldenen oder der Silbernen Latinität? Das sog. Spätlatein unter Einschluss des Vulgärlateins, das Kirchenlatein, das Mittellatein, das Humanistenlatein? Sind es gar Verballhornungen lateinischer Sprachelemente? Etwa auch das sog. Küchenlatein? Fragen über Fragen. Relevant im Sinne des Lateinunterrichts und seiner Wirkung ist vor allem das Schullatein bzw. das Latein der Schulautoren; relevant für die Lernwirksamkeit im Sinne der Sprachlernkompetenz ist hingegen letztlich das Latein der Schüler. Aber wo finden sie es? In der eigenen Muttersprache? in deren ihnen verfügbaren Bildungswortschatz? In den Fremdsprachen? Im Englischen? In einer Fachsprache? Wie weit sich der Begriff des Lateins dehnen lässt, zeigt nicht zuletzt der britische Romanist Glanville Price (1988: Einleitung): Für ihn sind die Sprachen Roms, Madrids, Paris oder Rio de Janeiros nur Varietäten ein und derselben Sprache, der lateinischen.

Relevant im Sinne der Lernwirksamkeit sind natürlich auch die modernen Sprachen. Unter diesen spielt längst die sog. Kommunikationsfähigkeit in der Zielsprache eine zentrale Rolle. Hierzu gehören als Voraussetzungen: das Hörverstehen und das Sprechen, das Leseverstehen und das Schreiben sowie die Sprachmittlung. Die hierzu notwendigen Ressourcen an Wissen und Können sind prozeduraler und deklarativer Art. Das Übersetzen aus dem Lateinischen in die Muttersprache bildet diese Ressourcen nur zu einem sehr geringen Teil aus. Was in einer fremden Sprache Kommunizieren erfordert, erfahren Schüler im Lateinunterricht nicht.

²⁴ Die erziehungswissenschaftlichen Studien zur Lernwirksamkeit von Lateinkenntnissen im Hinblick auf den Erwerb weiterer Fremdsprachen bleiben hier unerwähnt (Haag & Stern 2003; Meißner 2003; Neveling 2006; Ortner 2010; allerdings greift der von Ortner verwandte Begriff der „kognitiven Fähigkeiten“ viel zu unscharf. Er trifft die Frage des zwischensprachlichen Transfers allenfalls marginal). Der hier entgegretende forschungsmethodische Fehlgriff macht auch vor empirischen Studien nicht halt (z.B. Jessner et al. 2018). Er durchzieht weite Strecken der Diskussion um die Rolle von Latein für den Sprachtransfer.

Eine weitere wichtige Frage für die Transferabilisierbarkeit sprachlicher Schemata ist Folge des Umgangs eines Menschen mit den betreffenden sprachlichen Strukturen. In der Muttersprache – so die allgemeine Erfahrung – kommen die Wörter ‚wie von selbst‘. Dies ist Folge des tausend- und abertausendfachen Zusammenschaltens von Sprache und Welt, Wortform (Benennung) und Referenz (Inhalt); davon, dass Welt immer wieder auf Sprache gebracht werden muss und dass wir, wenn wir Hörverstehen oder lesen, den sprachlichen Code auch immer wieder auf Inhalte oder Mitteilungen zurückführen müssen. Es ist dieses ständige Wechselspiel von sprachlichen Formen und Inhalten, das die Habitualisierung von Sprache bewirkt. Besonders schnell muss dies bei dem erwähnten Hörverstehen geschehen – für das der Lateinunterricht weder Vorbilder noch Übungen bietet, an denen eine entsprechende Lernkompetenz ausgebildet werden könnte.

Wie die Fremdheitsforschung – aber auch die alltäglichen Erfahrungen – lehren, ist zwischen nahen und distanten Fremdheiten zu unterscheiden. Distanzen zeichnen sich jeweils auf den verschiedenen Achsen von Raum und Zeit ab. Das Lateinische führt im zeitlichen und zweifellos auch zivilisatorischen Sinne in eine ‚entfernte‘ Fremdheit. Die Welt des antiken Rom oder des frühen Mittelalters ist uns Heutigen viel fremder als die Kultur unserer heutigen Nachbarn. Die Autoren der *Homburger Empfehlungen* für eine sprachenteilige, d.h. mehrsprachige, Gesellschaft haben dies gesehen, als sie von den „Erschließungssprachen“ für „zeitlich oder räumlich ferne“ Kulturen sprachen:

„Im Sekundarbereich II soll eine Erschließungssprache gelehrt werden. Die Kenntnis dieser Sprache soll besonders dazu dienen, eine zeitlich oder räumlich ferne Kultur aufzuschließen. Erschließungssprachen in diesem Sinne sind z.B. das Lateinische oder Altgriechische sowie das Arabische, Chinesische oder Japanische.“ (Christ 1980: 98)

Doch was sagt der KRM zum Lateinischen als Transfersprache? Besehen wir einen beliebigen Ausschnitt, so zeigt schon dieser die Problematik der Fragestellung:

Franz-Joseph Meißner

Fr	It	Pg	Sp	RoS-	En	De	ETYMA
ABANDONNER	ABBANDONARE	ABANDONAR	ABANDONAR	RoSE	abandon, leave	aufgeben, verlassen	<germ BAND
abattre	abbattere	abater	abatir	RoS	bring down	niederschlagen, betrüben, schlachten	<BATTUERE
abîmer	guastare	ESTRAGAR	ESTROPEAR	oFk	break, damage, ruin	kaputtmachen, kaputtgehen	<ABYSSUS (jeter dans l'™); vI STRAGARE; vI STROPPIARE<DISTURPIARE
ABONDER	abbondare	abundar	ABUNDAR	RoSE	abound	im Überfluss h.	ABUNDARE
ABORD (D™)	DAPPRIMA	PRIMEIRO	PRIMERO	RoS3	first of all	zuerst	<franz BORD; PRIMUS
ABOUTIR	sboccare	desembocar	desembocar	RoS3	lead into	führen zu, enden	franz BUTTON (frapper); <BUCA
ABRI	RIPARO	abrigo	abrigo	RoS-I	shelter, cover, refuge	Schutz	ABRICARE (se m. au soleil); <REPARARE
ABSENCE	assenza	ausência	AUSENCIA	RoSED	absence	Abwesenheit, Absenz	ABSENTIA
absent	assente	ausente	ausente	RoSED	absent	abwesend	ABSENS
ABSOLU	ASSOLUTO	ABSOLUTO	ABSOLUTO	RoSED	absolute	absolut	ABSOLUTUS
ABSORBER	assorbire	absorver	absorber (consu	RoSED	absorb	absorbieren	ABSORBERE
abuser	abusare	ABUSAR	ABUSAR	RoSE	abuse	missbrauchen	AB+USARE (<USUS)
ACCENT	accento	acento	acento	RoSED	accent	Akzent	ACCENTUS
ACCEPTER	ACCETTARE	ACEITAR	ACEPTAR	RoSED	accept	akzeptieren	ACCIPERE
ACCÈS	ACCESSO	acesso	acceso	RoSE	access	Zugriff, Zugang, Erreichen	ACCESSUS
ACCIDENT	ACCIDENTE	ACIDENTE	ACCIDENTE	RoSE	accident	Unfall	ACCIDENS
ACCOMPAGNER	ACCOMPAGNARE	ACOMPANHAR	ACOMPAÑAR	RoSE	accompany	begleiten	vI COMPANIO
ACCOMPLIR	COMPIERE	CUMPRIR	CUMPLIR	RoSE	accomplish	vollenden	CUMPLERE
ACCORD	ACCORDO	ACORDO	ACORDE	RoS	according	Einverständnis, (im Akkord arbeiten)	ACCORDARE
ACORDER	accordare	ACORDAR	ACORDAR	RoS	concede, tune	zugestehen, Einverständnis geben	ACCORDARE
ACCOURIR	ACCORRERE	acorrer	acudir corriendc	RoS-S	rush up	herbeilaufen	CURRERE
accoutumer	accostumare	ACOSTUMAR	ACOSTUMBRAR	RoSE	accustom	gewöhnen	CONSUETUDO
ACCUILLIR	ACCOGLIERE	acolher	ACOGER	RoS	receive, welcome	empfangen	COLLIGERE
ACCUSER	ACCUSARE	ACUSAR	ACUSAR	RoSE	accuse	anklagen	ACCUSARE
ACHETER	COMPRARE	COMPRAR	COMPRAR	RoS-F	buy	kaufen	vI ACCAPTARE; COMPERARE/COMPARRE
ACHEVER	portare a fine	ACABAR	ACABAR	RoS-I	achieve	zu Ende bringen	vI ACCAPARE;
AIGRE	AGRO	AZEDO	AGRIO	RoSED	acid	sauer, bitter, azid	ACER/ACRIS/ACRE
acompte	ANTICIPO	antecipação	anticipo		anticipation	Vorschuss	COMPUTARE; ANTICIPARE

Abbildung 13: zufällig gewählter Ausschnitt aus dem KRM mit den Rubriken ROS- und ETYMA

Nahezu alle T4-Selektionen aus RoS, RoS2 und RoS3, RoS-*, RoSE und RoSED bzw. RoSE-* und RoSED-* belegen eigentlich die Irrelevanz der Frage für das erfolgreiche Erlernen des romanischen Vokabulars gestützt durch Latein. Denn wer dt. *Akt*, en. *act* oder it. *atto* kennt, versteht auch lat. ACTUM auf Anhieb – und umgekehrt. Es ist empirisch falsch anzunehmen, das Verständnis von en. *act*, fr. *acte*, pg. *ato* setze die Kenntnis von lat. ACTUS voraus. Viel wahrscheinlicher ist, dass Deutschsprachige auf *Akt* rekurren. Und genauso verhält es sich mit fast allen anderen Kognaten, für die das Lateinische eine Transferbasis bereitstellt bzw. die ihrerseits für das Lateinische eine solche bieten. Das gilt zuvorderst für die zuerst erlernte Sprache, die Muttersprache, und deren Potenzial für den Erwerb von Mehrsprachigkeit, welches viel zu oft versteckt bleibt (Wandruszka 1979). Ist also auch Italienisch schon Latein? Ist es dann mehr als ein *latino maccheronico*?

Apropos Lernökonomie: Die lateinische Morphologie und Syntax sind von einem Reichtum, der in den modernen (romanischen) Sprachen überwiegend keine Fortsetzung findet: Lateinlerner erwerben viele sprachliche Schemata, die im Französischen, Italienischen oder Spanischen gänzlich ohne oder nur mit sehr schwacher Entsprechung bleiben. Strukturell – morphologisch und syntaktisch – ist Latein viel weiter von den romanischen Sprachen entfernt, als es diese untereinander sind. So baut der Lateinunterricht wenig prozedural verwertbares Schemawissen auf, das sich unmittelbar auf romanische Sprachen anwenden ließe.

Wer einen Blick auf die von Freund, Schröttel & Trier (2004) zu pädagogischen Zwecken zusammengestellte Liste der 1000 häufigsten lateinischen Vokabeln des *Bamberger Query Corpus* wirft, entdeckt dort eine Menge Vokabeln, die im Romanischen keine Fortsetzung (Transferbasen) haben. Andererseits finden sich die dort genannten Lexeme nur zu einem geringen Teil im KRM-Repertoire wieder, wo sie nur einen geringen Teil der in 9556 Serien versammelten Lemmata abdecken. Zudem enthalten die KRM-Eintragungen mit der Kennung RoS, RoSE, RoSED, RoS und andere zwar eine Menge von aus dem Lateinischen abgeleiteten Formen, allerdings decken sich diese nicht dem in der Schule gelehrt Register.

Fazit: Entscheidend für die Leistung einer Wortform für den interlingualen Identifikationstransfer ist die Vertrautheit der Lerner mit ihr, und zwar in Umgang und Qualität (Müller-Lancé 2003: 401). Dies gilt für jede potentielle Brückensprache gleichermaßen. Wie jede andere indoeuropäische Sprache das Erlernen einer weiteren erleichtert, so tut dies auch Latein. Allerdings tut es dies viel weniger auf prozeduraler Ebene. Offensichtlich besteht der Grund darin, dass der Lateinunterricht keine Produktionsroutinen ausbildet – da er weder das Sprechen noch das Schreiben in der Zielsprache übt. Das Schullatein erfasst bei Weitem nicht die Breite, in der das Lateinische Etyma für die modernen romanischen Sprachen bereitstellt.

Als „Grundlage“ für das Erlernen moderner Sprachen – wie vielfach behauptet – sind die im Unterricht vermittelten Lateinkenntnisse schlicht viel zu schmal.

Als Sprache mehrerer vergangener Epochen ist Latein kulturell von unserer heutigen Welt und ihren Sprachen natürlich viel weiter entfernt, als es die heutigen europäischen Sprachen füreinander sind. Dies betrifft vor allem die Bedeutungsseite der Wörter. Umgekehrt belohnt es die Lerner, indem es deren Blick auf diese Vergangenheiten durch die Prismen der sie verbindenden Sprache und ihrer Zeiten öffnet und ihnen die Auseinandersetzung mit einer ‚alten‘ Sprache bietet.

5.4 Die Mutter- oder Zweitsprache Deutsch als Grundlagensprache für den romanischen Transfer

Als Mutter- oder Zweitsprache kann das Deutsche in seiner Wirkung auf das Potenzial des interlingualen Transfers kaum überschätzt werden. Es aktiviert im mentalen Lexikon der allermeisten Deutschsprachigen weitaus mehr Transferbasen als jede andere Sprache (sofern ihre Vertrautheit mit einer anderen Sprache das Deutsche nicht in der bezeichneten Qualität ablöst oder ihm gleichkommt²⁵). Dies gilt umso mehr, desto stärker die Sprecher mit dem Bildungswortschatz vertraut sind, den deutschsprachig aufwachsende Kinder vor allem im Verlauf der Sekundarstufe I erwerben. Er stellt zahlreiche Brücken zu den romanischen und weiteren Sprachen bereit. Natürlich verfügen die deutschsprachigen Adressaten des KRM in ihrer deutschen Erst- oder Zweitsprache über ein Kompetenzmaß, welches weit über das im KRM erfasste Segment hinausreicht. Die unter RoSD und RoSED gefassten Größen (2854 Serien) unterschreiten daher das Transferpotenzial deutscher Muttersprachler in erheblichem Ausmaß.

6. Frequenzintervalle aus erweiterten Grundmengen für primäre (romanische) Zielsprachen

Die weiteren Komputationen sind nicht nach Abfragen für einzelne ‚sekundäre‘ Zielsprachen gefiltert. Zur Erinnerung: Gezählt wird aus der jeweiligen Rubrik für die Frequenzstufe – also für Französisch in (Rfr), für Italienisch in (Rit) usw. – nach den Selektoren „R“ (<2000), „R§“ (2000-5000) oder „R-“ (<5000) (für Italienisch: „M“, „M§“ usw.) mit der jeweiligen Sprachkennung (z.B. „F“ oder „F OR Fw“). Daher

²⁵ Diese Formulierung soll keineswegs in Abrede stellen, dass hochgradig Mehrsprachige die Rolle der deutschen Muttersprache relativieren. Im Prinzip gilt, dass sich das Transferpotenzial eines Menschen aus dem Gesamt seiner sprachlichen Erfahrungen bzw. seines Sprachenwissens in Kombination mit seiner Sprachbewusstheit und Sprachlernkompetenz bildet.

beziehen sich die Daten quasi auf den gesamten Datenbestand des KRM (für Französisch etwa auf die „F OR Fw“-elizitierten Lemmata²⁶). Dies sind sowohl die formdivergenten Signifikanten als auch deren Wiederholungen in anderen Serien.

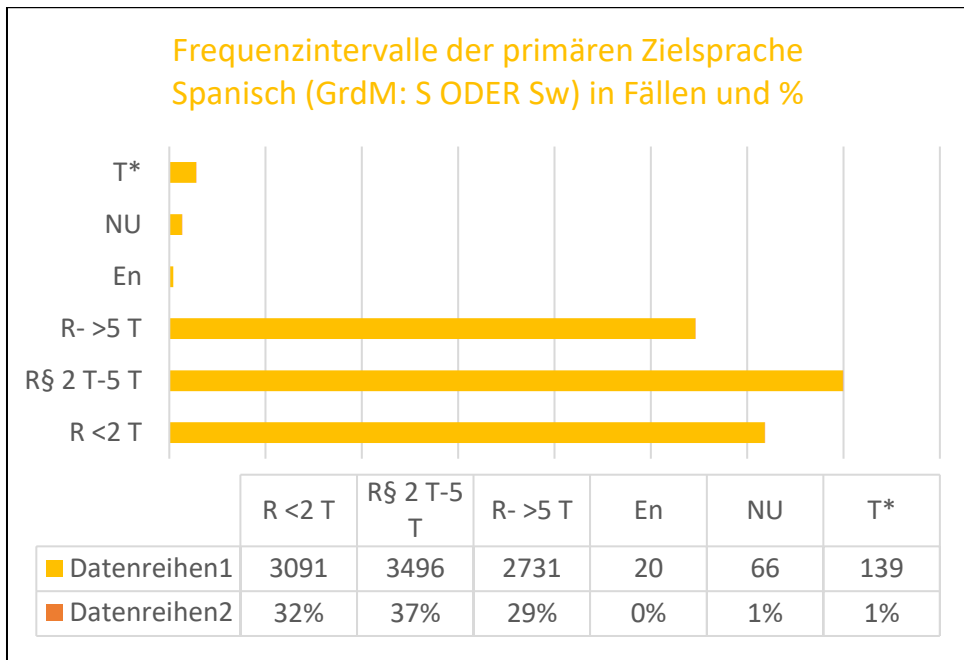


Abbildung 14: Frequenzintervalle Kernwortschatz Spanisch

Wie Vergleiche der Komputationen für die einzelnen Zielsprachen ergeben, sind die Okkurrenzen für „R“, „R§“ und „R-“ deutlich höherfrequent als z.B. die der Numeralia („Nu“), der Toponyme („T“) und der semantisch regional fixierten Wörter. Die Kennung „EN“ (Englisch) meint Verzeichnungen so markierter englischer Wörter in den benutzten Frequenzlisten der Zielsprachen. In diesen handelt es sich um Solipsismen (Typ: *indetermancy* in der De Mauro-Liste). Zudem verhalten sich die Glossare nicht einheitlich, was die Aufnahme solcher Wörter angeht (z.B. *white*, *Scotch* usw.). Didaktisch sind solche Anglizismen für Lerner, die des Englischen einigermaßen mächtig sind, natürlich transparent – sofern es sich nicht um sog. Pseudo-Anglizismen (Typ: dt. *handy*) handelt.

Die Okkurrenzen der Niedrigfrequenten betragen insgesamt 225 Fälle, d.h. 2,3 Prozent des Gesamtbestandes. Daher werden im Weiteren nur die hochfrequenten Daten zu „R“, „R§“ und „R-“ erwähnt. Sie summieren sich auf 9318 Serien.

²⁶ F steht für formdivergente französische, S für spanische Lemmata, Fw für Wiederholung eines bereits genannten französischen Lemmas.

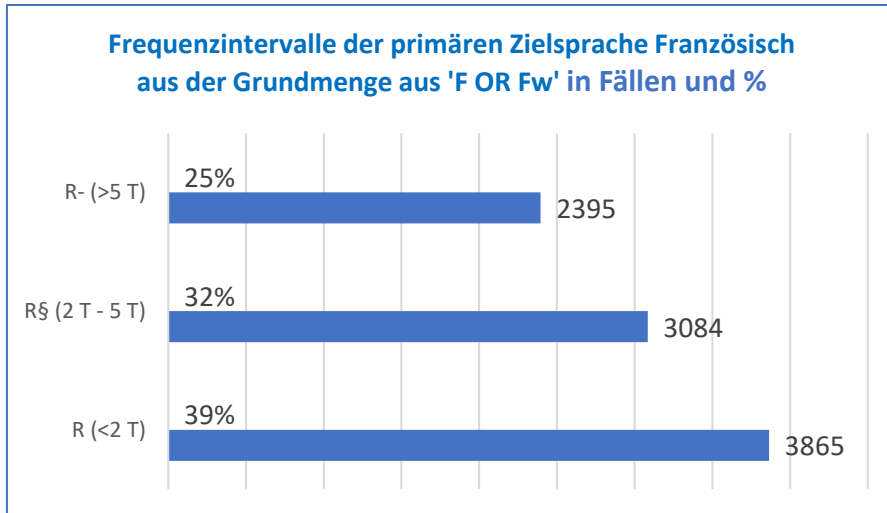


Abbildung 15: Frequenzintervalle Französisch

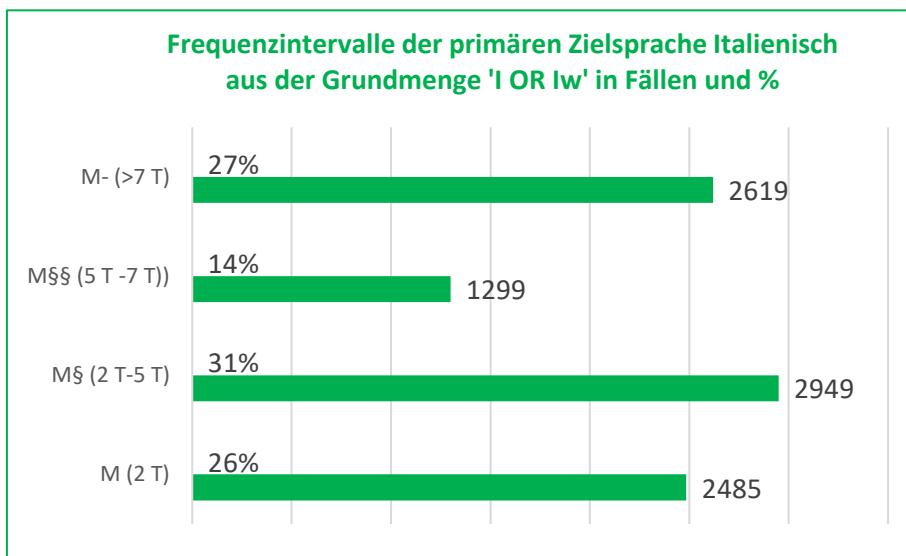


Abbildung 16: Frequenzintervalle Italienisch

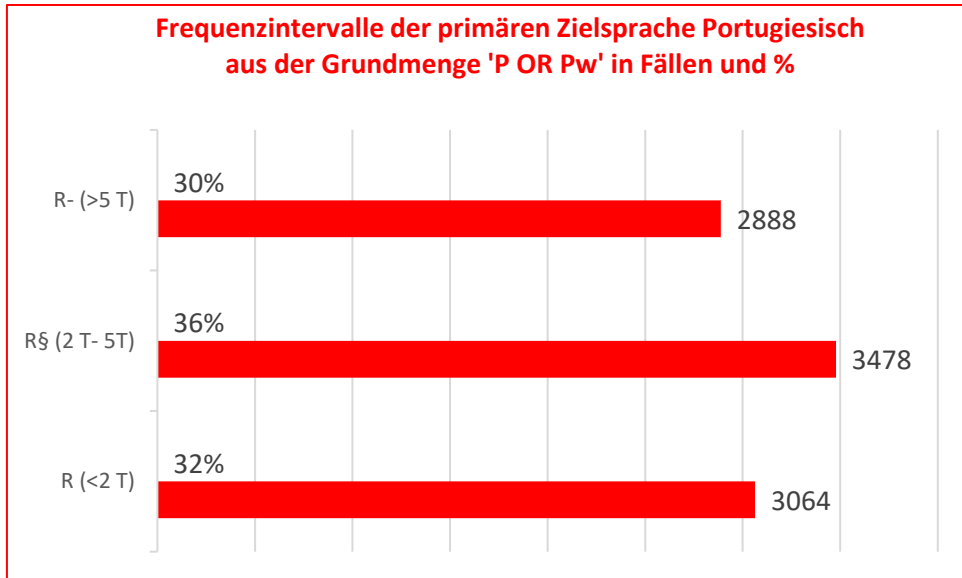


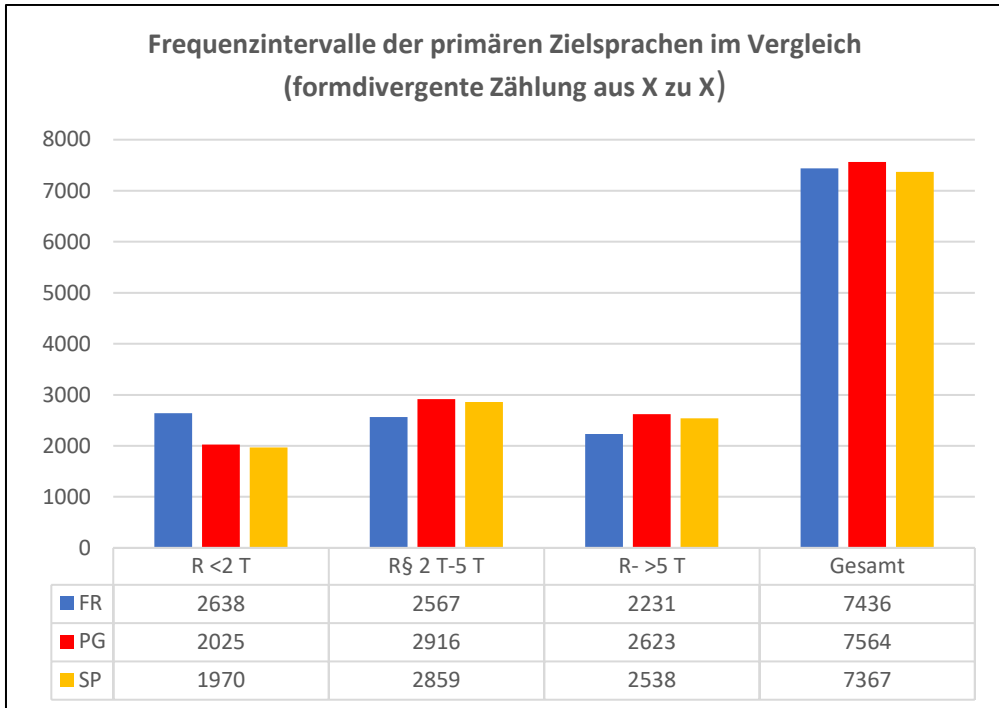
Abbildung 17: Frequenzintervalle Portugiesisch

Den größten Anteil stellt in den Kernwortschätzen von Portugiesisch und Spanisch, wie gesagt, das Intervall mit der Frequenzmarke 2000-5000, gefolgt von dem der Grundwortschätze (<2000). Die Französisch-Liste entspricht diesem Muster nicht; hier umfasst der Grundwortschatz 39 Prozent der Lemmata gegenüber 26 und 32 Prozent. Ein möglicher Grund für die Abweichung könnte – wiederum – entweder in der Machart der Klett’schen Grundwortschätze und/oder der verschiedenen Routledge-Wörterbücher liegen – die Französisch-Liste inventarisiert (nach dem Eindruck des Bearbeiters) restriktiver als die beiden die iberischen Sprachen verzeichnenden Glossare (Lonsdale & Le Bras 2009: 3).

Zur Erinnerung: Die Rubriken <5000 erfassen z.T. Übersetzungen der nachbarsprachlichen Lemmata, auch mit der Markierung >5000. Einem französischen Lemma mit dem Intervall <2000 kann also im Prinzip ein spanisches der Klasse >5000 entsprechen²⁷. Bezüglich der Zuordnung der deutschen und englischen Bedeutungsadäquanzen war das Merkmal ‚Frequenzrang‘, wie gesagt, nachgeordnet (vgl. Davies & Gardner 2010, Jones & Tschirner 2006).

Während die bisherigen Komputationen zu den Frequenzrang-Intervallen mit dem Selektor (X OR Xw) operierten, zeigt die folgende Zusammenstellung nach einzelsprachlich-formdivergent gezählte Intervalle (Selektor X):

²⁷ Man vergleiche den Fall von pg./sp. [anteceder/preceder](#).



*Abbildung 18: Frequenzintervalle zu den GrdM
nach intralingual formdivergenter Zählung*

Die Anteilswerte liegen innerhalb einer Bandbreite von 39% bis 27%. Die Differenz zwischen den romanischen Sprachen ist minimal.

	FR	PG	SP	IT
R in %	35 (25) ²⁸	27 (32)	27 (32)	(26)
R § in %	35 (32)	35 (36)	39 (37)	(31)
R- in %	30 (39)	38 (30)	34 (29)	(27+14)

Italienisch ist aus der regulären Zählung genommen, weil die De Mauro-Listen in vier Kategorien einteilt. Zum Vergleich präsentiert die folgende Abbildung die Grundmengen aus (I OR Iw) und die aus (X). Die Größe der Grundmengen ist wie folgt:

²⁸ Die in () genannten Werte zählen die Werte aus der einfachen X-Selektion.

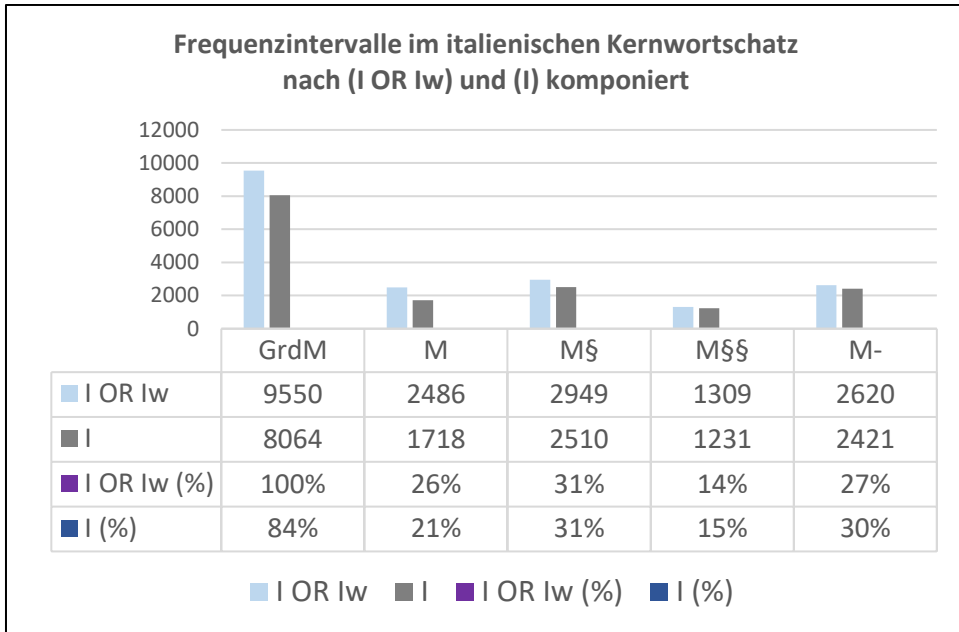


Abbildung 19: Frequenzintervalle des Italienischen aus I und I OR Iw

Infolge des erweiterten Ausschnittes des italienischen Inventars (<7000) und der Verteilung auf vier Frequenzrang-Intervalle fallen die Raten pro Rubrik geringer als in den Schwestersprachen aus. Im Unterschied zu den vorausgegangenen Tabellen zeigt Abb. 19 die Differenzen zwischen der erweiterten Grundmenge aus (I OR Iw) und der einfachen aus (I). Bei einer Differenz von 16% in den Grundmengen zeigen die Werte der Frequenzränge einen maximalen Unterschied von 5%.

In diesem Zusammenhang sind kombinierte Suchroutinen sinnvoll. Dies betrifft etwa die Kombination von ‚R‘ in Rsp mit RoSE OR RoSED in ROS- und S in Lsp. (Sprich: Zeige innerhalb des Grundwortschatzes [<2000] alle RoSE OR RoSED im Feld ROS für S [Spanisch].) Diese Selektion findet alle Fälle englisch-romanischer Formkongruenz mit dem Frequenzrang <2000 innerhalb der Sortierung von Spanisch als Zielsprache:

111	RoSE	ANCIEN	ANZIANO	ANCIÃO	ANCIANO	ancient	S	R
118	RoSED	ANGOISSE	angoscia	angústia	ANGUSTIA	anguish	S	R
119	RoSED	ANIMAL	ANIMALE	ANIMAL	ANIMAL	animal	S	R
122	RoSE	ANNÉE	ANNATA	ANO	AÑO	annual	S	R
124	RoSED	ANNONCER	annunziare	ANUNCIAR	ANUNCIAR	announce	S	R
132	RoSE	APPARAÎTRE	APPARIRE	APARECER	APARECER	appear	S	R
134	RoSE	APPARENCE	APPARENZA	APARÊNCIA	apariencia	appearance	S	R
140	RoSED	APPLIQUER	APPLICARE	APLICAR	APLICAR	apply	S	R

21: Ausschnitt aus der Selektion RoSE OR RoSED mit R in Rsp und S in Lsp

Mindestens so wichtig wie die Komputationen ist die Identifikation von Lemmata nach bestimmten Merkmalen. Auch hier können Frequenzrang-Intervalle eine wichtige Rolle spielen.

Didaktisch ist der Vergleich der Frequenzangaben relevant, weil er erstens die Befunde zu den einzelnen Zielsprachen ergänzt und weitere Interpretationen erlaubt; weil zweitens, wenn Frequenzüberlegungen bei der Selektion des didaktischen Wortschatzes eine Rolle spielen, die Kenntnis von Frequenzrang-Klassen pädagogisch sinnvolle Entscheidungen zu fällen ermöglicht. Dies sollte nicht allein aufgrund von Augenschein geschehen.

7. Frequenzintervalle in Kombination mit Transparenzstufen in primären Zielsprachen

Obwohl der Aussagewert für die Praxis des Fremdsprachenunterrichts und die Konstruktion des lexikalischen Inputs nur gering sein dürfte, sei dem Zusammenhang von Frequenzrängen und Transparenzstufen an dieser Stelle kurz nachgegangen.

Die Syntax zur Identifikation der betroffenen Strukturen ist klar, sie lautet z.B. für das Portugiesische: (P OR Pw) in Lpg MIT (R) in Rpg MIT (RoS OR RoSD OR RoSE OR RosED) in ROS-. Die Selektion erfasst die T4-stufigen Lemmata aus dem Frequenzintervall <2000. Die Syntax zeigt 1701 Okkurrenzen. Der T0-Wert aus (oFk) ist 309. Der T1-Wert aus (*3) lautet 368, der für T2: 198. Der T3-Umfang akzeptiert ein gewisses Bias, indem der Selektor (*-*) nicht nach Sprachen spezifiziert. Die Selektion umfasst 489 Serien. Bei einem Gesamtumfang von 3066 Serien für (R) (=100%) ergibt dies prozentuale Anteile von

FR		
T4=56%		
T3=16%		
T2=6%		
T1=12%		
T0=10%		
FR	IT	SP
T4=1710	1701	1701
T3=489	489	489
T2=198	198	198
T1=368	368	368
T0=309	309	309

Die im unteren Tabellenabschnitt gefassten Fallzahlen belegen interlingual die Konstanz der Suchergebnisse. Diese erklärt sich damit, dass die jeweiligen (sprach-

lichen) Grundmengen quasi das Gesamt der Serien des KRM umfassen. Die Gesamtzahl der R-Selektion erreicht, wie gesagt, 3066 Fälle.

Zu beachten ist, dass der Gesamtumfang der Serien für alle Selektionen aus (R) konstant ist. – Wie immer wurden auch hier die Werte gerundet.

Die in der Tabelle angegebenen Prozentwerte – d.h. die Transparenzwerte im Frequenzintervall <2000 – sind damit im Bereich der Grundwortschätze für die Zielsprachen dieselben.

Das Intervall 2000-5000 ist mit Ausnahme des Französischen für die Zielsprachen das jeweils größte. Die Abweichung zwischen den Sprachen Italienisch, Portugiesisch und Spanisch beträgt in diesem Segment: zwischen 31% und 36% ([Abb. 12](#) und vorausgehende).

Zur Erfassung dieser Auswahl wird der Selektor (R) gegen (R\$) ausgetauscht. Die Grundmenge aus (R\$) ist 3478.

FR		
T4=58%		
T3=15%		
T2=5%		
T1=10%		
T0=8%		
FR	IT	SP
T4=2011	2011	2011
T3=536	536	536
T2=185	185	185
T1=355	355	355
T0=287	287	287

Vergleicht man die grundwortschatzlichen (für <2000) mit den kernwortschatzlichen Beständen, so ist zunächst folgende Berechnung erforderlich: Da der Grundwortschatz eine Teilmenge des Kernwortschatzes ausmacht, liegt der zu findende Wert zwischen jenem für den Grundwortschatz und dem für das Intervall 2000-5000. Da die maximale Abweichung zwischen den beide Intervalle betreffenden Ziehungen maximal 2 Prozent beträgt und die sprachlichen Grundmengen gleich sind, kann der Wert für den Kernwortschatz ermittelt werden durch [GWS-Wert + KWS-Wert (:) 2]. Für T0 lautet demnach die Rechnung: 18%:2=9%.

8. Interlinguale Transparenzintervalle: Methodische Vorüberlegungen und Daten

Um Redundanzen zu vermeiden, verbinden die folgenden Zählungen komputativ-methodologische Überlegungen und statistisch-deskriptive Daten miteinander. Dem geht das Faktum voraus, dass die Beispiele (die ja nicht fiktiv sind, sondern der KRM-Datenbank entstammen) die Nachvollziehbarkeit der methodischen Entscheidungen erhöhen.

8.1 Grundmengen (aus sprachlichen Variablen)²⁹

Grundlegend für die Komposition der Grundmengen ist die bereits genannte Unterscheidung zwischen der intralingual-formdivergenten und der intralingual-formkongruenten Zählung. Letztere bedeutet, dass identische Formen in seriell unterschiedlichen Kombinationen mehrfach auftreten und gezählt werden müssen (Reduplikationen). Auch dies hat, was die Serienbildung angeht, natürlich viel mit Fragen um Bezeichnungen und Bedeutungen zu tun.

Hierzu vorab ein Beispiel zu it. *mantenimento*/sp. *mantenimiento* bzw. pg. *manutenção* (Wartung, Pflege, Instandhaltung; Erhaltung, Beibehalt, Lebensunterhalt) und pg. *mantimento* (Lebensmittel [sp. *viveres*], Aufrechterhaltung, Unterhalt [fr. *maintenance/maintien*, it. *mantenimento*]), Fortbestand, Beibehaltung, Aufrechterhaltung (fr. *maintien*; auch: *conservation* u.a.m.)³⁰. Zur hinreichenden Disambiguierung der portugiesischen Lemmata sind mindestens drei Serien erforderlich:

4364	entretien	manu- tenzione	manu- tenção	man- tenimiento	main- tenance, upkeep	Aufrecht- erhaltung, Unterhalt
4420	maintien	man- tenimento	manu- tenção	manteni- miento	main- taining, holding	Pflege, Versorgung, Aufrechterhaltung
6023	mainte- nance	man- tenimento	manti- mento	menteni- miento	main- tenance, servicing	Aufrechterhaltung

22: Beispiel für Mehrfachzählung infolge von Polysemie

²⁹ Natürlich können Grundmengen auch aus nicht-sprachlichen Faktoren gebildet werden. Es macht Sinn, alle Relationen für die Grundmenge des Grundwortschatzes ($R < 2000$) zu bestimmen.

³⁰ Weitere semantische Hinweise hält der KRM in der Rubrik NOTIZ bereit.

Offensichtlich konkretisieren die Sprachen jeweils eine eigene Beziehung zwischen Bezeichnung und Konzept – wie in den Bemerkungen zu Polysemie, Polynymie und Homonymie bereits angesprochen –, jeweils nach ihren eigenen lexikosemantischen Feldern, deren Merkmalen und Strukturierung. Dies bleibt nicht ohne Auswirkungen auf die Konturierung des KRM-Inventars und die Anteile seiner Komponenten, was sich bei der Komposition der Serien durch Reduplikationen (X OR Xw statt bloßem X) zeigt.

Der Anteil solcher durch Reduplikationen bedingter Mehrfachzählungen liegt bei 20% für das Französische, für das Italienische bei 16%, das Portugiesische bei 19% und Spanisch bei 21% (also im Schnitt bei 19%). Aufgrund der Markierung der ‚semantischen Distanz‘ und der distinguierenden Selektoren „X“ und „Xw“ in den Feldern Sprachen (z.B. Lpg für Portugiesisch) ist es leicht möglich, solche Reduplikationen ggf. ‚mit starker Bedeutungs differenzierung‘ zu identifizieren.

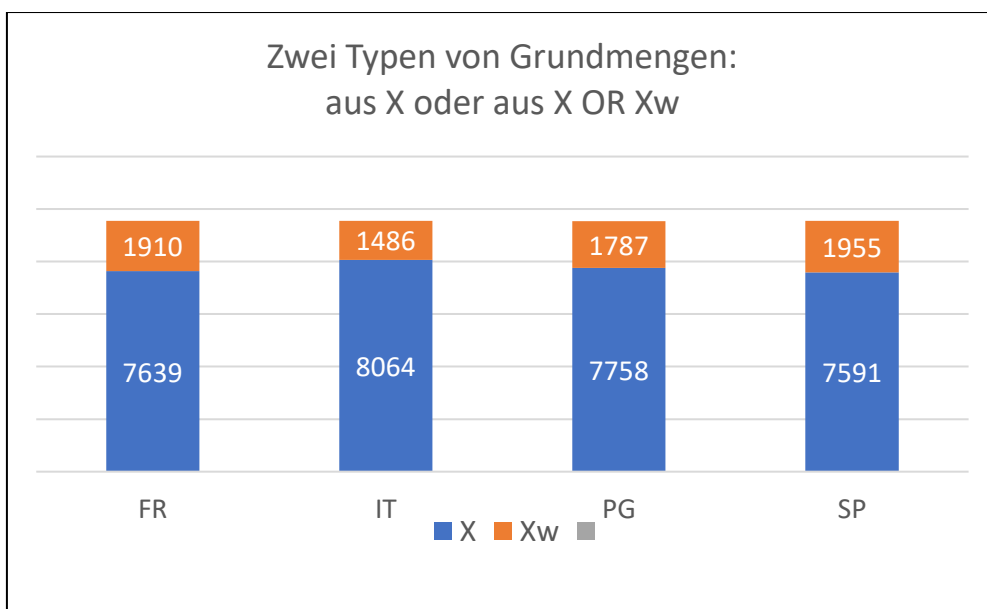


Abbildung 20: X-, Xw-, X OR Xw-Anteile nach Sprachen und Fällen

Die Gesamtzahlen aus beiden Grundmengen: FR=9549, IT=9550, PG=9545 und SP=9546. Die Menge aus X OR Xw – für FR also 9549, für IT 9550 usw. – erreicht quasi die Gesamtzahl der Serien. Prozentual belaufen sich die Anteile der Reduplikationen auf gerundet 20% (FR), 16% (IT), 19% (PG) und 20% (SP).

Die nächste Graphik zählt Transparenzfälle nach Stufen und Sprachen aus der erweiterten Grundmenge (X OR Xw)³¹.

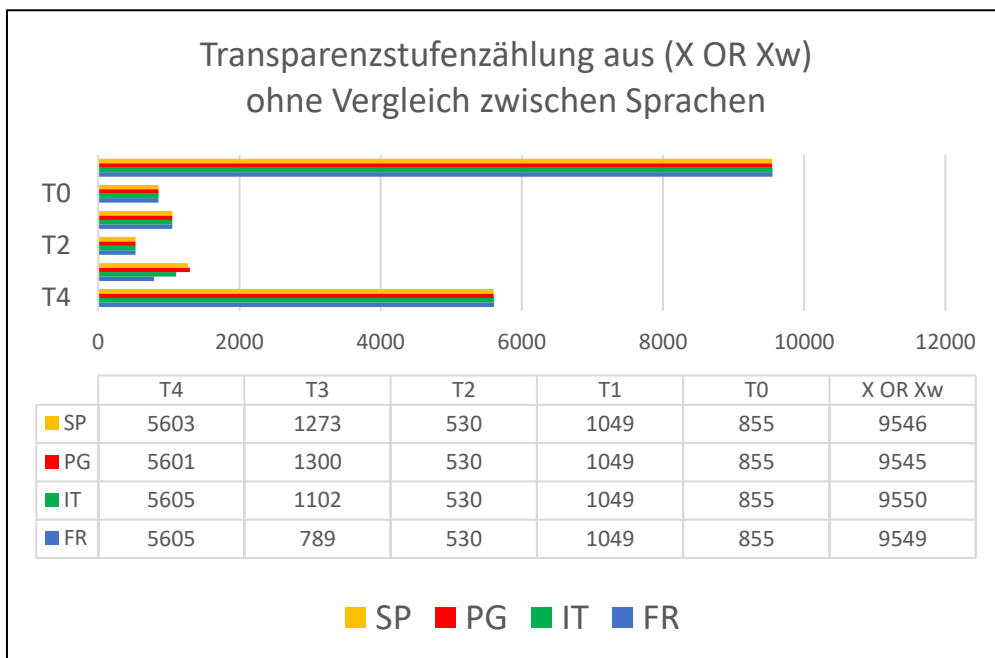


Abbildung 21: Interlinguale Transparenzanteile in Grundmengen ohne sekundäre Zielsprachen

In der folgenden Abbildung wird die Grundmenge zu (X) verändert:

³¹ Der besseren Verständlichkeit halber seien nochmals die Zählweisen verdeutlicht:

T4	Zeige aus der Grundmenge (X OR Xw) alle Markierungen von {RoS, RoSD, RoSE, RoSED}
T3 opak	Zeige aus der Grundmenge (X OR Xw) alle Markierungen von {RoS*-X}
T3 transp.	Zeige aus der Grundmenge (X OR Xw) alle Markierungen für {RoS*-* (-) RoS-X nach Sprachen}
T2	Zeige aus der Grundmenge (X OR Xw) alle Markierungen von {RoS*2 bzw. *2}
T1	Zeige aus der Grundmenge (X OR Xw) alle Markierungen von {RoS*3 bzw. *3}
T0	Zeige aus der Grundmenge (X OR Xw) alle Markierungen von {oFk}

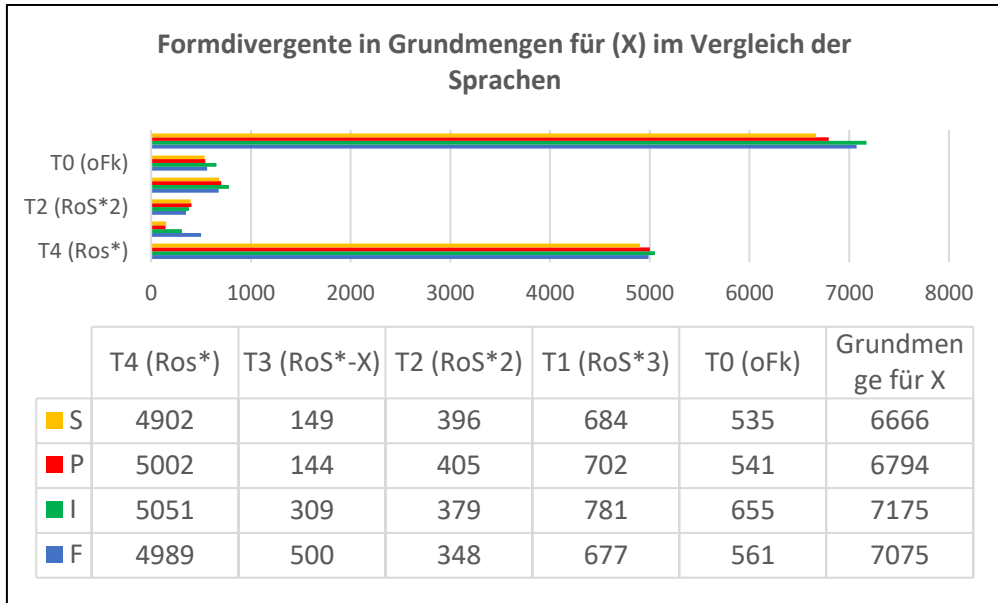


Abbildung 22: Zählung T-Kohorten aus (X)

Beide Zählweisen (vgl. Abb. 20, 21, 22) geben eine differenzierte Auskunft über das Transferpotential der Sprache X im KRM, und zwar sowohl zwischen den intralingual als auch den interlingual formdivergenten Anteilen.

8.2 Zählungen aus gemischt-sprachlichen Kombinationen (MIT-Selektionen)

kombinierten Grundmengen oder sog. MIT-Gruppen. Gemischt-sprachliche Mengen sind aus mindestens zwei Sprachen (oder mehr) komponiert. Eine solche Selektion würde etwa der Kombination von Sprache X (F) mit Sprache Y (S) entsprechen ([s. unten](#)):

Lfr	PROFIL	Lsp	RoS-
KRM-Inventar Mutter(KRM-Inventar Mutter(KRM-Inventar Mutter(KRM-Inventar Mutter(
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
"F"		"S"	Wie "*2"

23: A Selektorenkombination für (*2) aus (F) MIT (S)

Hier erfolgt die Komputation aus den Selektoren Spanisch mit Französisch für alle Selektoren mit der Komponente „2“ (= 264 Fälle). Die Variablen für ROS- werden also aus 2 Bedingungen eliziert. Der Selektor bestimmt die qualitative Verteilung, hier also die aus Ros*2, wie folgendes Muster belegt:

RoSE2
RoSED2
RoS2

Eine Ausweitung der Grundmenge auf ‚aus 3 Sprachen komponiert‘ verändert die Okkurrenzen, da ja alle Bedingungen erfüllt sein müssen. Anders gesagt: „*2“ muss sich für die Felder FR, IT und SP feststellen lassen. Es ergeben sich 218 Fälle.

Die Reduktion um „S“ erhöht die Zahl der Fälle für FR auf 345, für FR MIT IT auf 274, da eine der Bedingungen aufgehoben wird.

Zur Erinnerung: Die Syntax (F OR Fw) MIT (S OR Sw) für (RoS*2) führt zu 531 Okkurrenzen, weil in diesem Fall immer der größere Wert für (F OR Fw) → 531; (S OR Sw) → 531 den Ausschlag gibt. Die Anzahl der Fälle ist gleich, weil (F OR Fw) und (S OR Sw) jeweils nahezu alle Serien des gesamten KRM-Inventars erfassen.

Die Kombination aus (F OR Fw) mit (S) für (*2) erbringt 393 Serien, die aus (S OR Sw) mit (F) für (*2) 345.

Lfr	PROFIL	Lsp	RoS-
KRM-Inventar Mutter	KRM-Inventar Mutter	KRM-Inventar Mutter	KRM-Inventar Mutter
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
F* Oder *Fw		*S* Oder *Sw*	Wie *2*

24: B Selektion aus (F OR Fw) MIT (S OR Sw) für (*2)

Wie sehr die abhängigen Variablen (z.B in ROS-) von der Grundmenge bestimmt werden, belegen folgende Reihen zu den RoS*-*Zählungen im Zusammenhang mit der Ermittlung der T3-Werte (Opazitätsfälle):

Syntax	FR	IT	PG	SP
Für *-* aus X	1045	1177	1089	1056
Für *-* aus F mit I	862	-	885	881
Für *-* aus I mit P mit S	-	729		
Für *-* aus F mit I mit P mit S	607			
Für *-* aus X OR Xw	1488	1488	1488	1488

25: Zählungen aus gemischt-sprachlichen Grundmengen

Offensichtlich sind die Stärken der T0-, T1- und T2-Kohorten von Sprache zu Sprache nicht sehr ausgeprägt – bei der Grundmenge aus [X OR Xw] sind sie quasi identisch, was sich einmal mehr mit der annähernd gleichen Größe der Grundmengen erklärt.

Sie liegen bei ca. 2,4% und 3,2% für die X-Messung. Die bereinigten T3-Werte; ermittelt aus der Differenz von (*-* minus RoS*-X), weisen hingegen die größten Unterschiede auf. Bemerkenswert ist, dies sei schon hier erwähnt, der niedrige Anteil von [T3 für Französisch](#) in den Transparenzraten. Zu beachten ist, dass Abb. 22 keine gemischt-sprachliche Zählung zugrunde legt.

Die größten Differenzen ergeben sich aus den (X) mit (Y) gebildeten MIT-Grundmengen. Für die Zählungen bedeutet dies:

- Die *-*-Zählungen aus X sowie X mit Y ergeben geringe Unterschiede in den Werten.
- Je variiert die Selektion gestaltet wird, desto geringer die die erzeugte Schnittmenge.
- X OR Xw-Komputationen fassen (wie gesagt) quasi 100% des Gesamts der Datenbank. Die Okkurrenzen für die Selektionen fallen in den einzelnen Sprachen nahezu identisch aus.

Abb. 21 und die folgenden erlauben weitere Einsichten zur möglichen Varianz zwischen den Selektionen der Transparenzklassen:

- Als unabhängige Variable ist, wie angedeutet, die aus den Sprachenselektoren gebildete Grundmenge entscheidend für den Umfang der selektierten Fälle (abhängige Variable).
- Greift ein Selektor komplett seriell identisch, so erlaubt dies keine Abweichung des T-Wertes zwischen den verschiedenen sprachlichen Kohorten. Dies ist der Fall für T4 und T0 in einfachen, nicht kombinierten Grundmengen.
- Zeigt eine Serie hingegen heterogene Zuordnungen bzgl. der Transparenzstufen, so verbindet sich hiermit eine potentielle Varianz der T-Werte. Dies betrifft die T2 und T3-Kohorten (vgl. Abb. 20 mit 21).
- Beinhaltet ein Selektor eine sprachspezifische Messung – wie für die Opazitätsfälle von T3 – so zeigt diese Kohorte deutliche Unterschiede (zur Umrechnung von Opazitäts- in Transparenzfälle: s.u.).

Eine Veranschaulichung der von der Grundmenge (X oder Xw) einerseits und (X) andererseits gezählten Kohorten erlaubt die folgende Graphik für den französischen Anteil des KRM-Repertoires (vgl. für das Italienische [Abb. 19](#)):

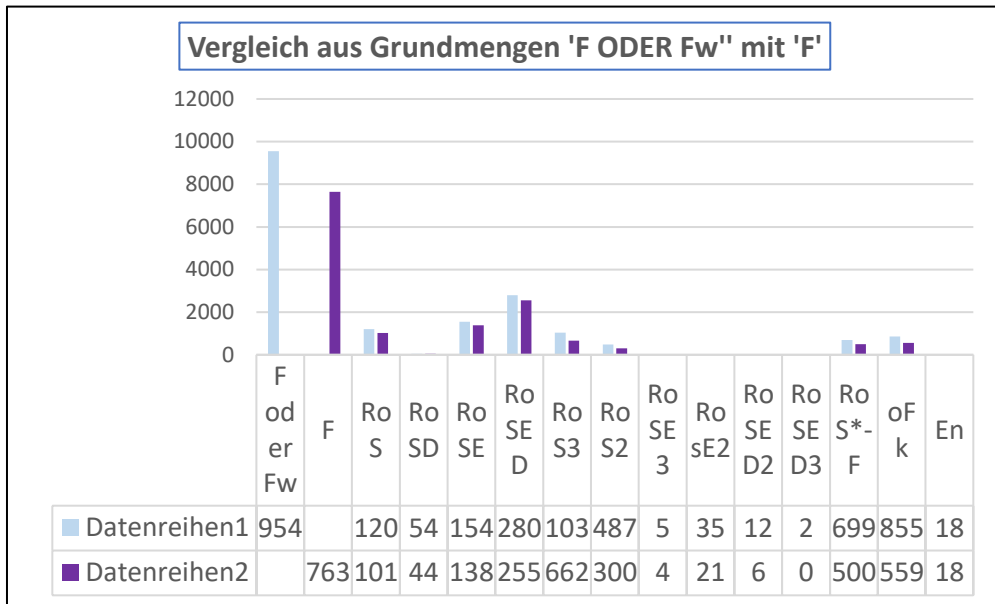


Abbildung 23: Selektionen aus unterschiedlichen Grundmengen zu Französisch

Die Daten der Reihe 1 präsentieren die Ziehungen aus (F OR Fw), die von 2 die aus (F).

Ehe die Daten für gemischt-sprachliche Selektionen vorgestellt werden, erscheint es angeraten, die Varianz der T3-Werte aus unterschiedlichen Serien und Sprachkombinationen für RoS*-X zu erörtern.

8.3 Komputationen zur Transparenzstufe 3: zum Transparenzanteil aus Opaken-Zählungen

Die folgenden Graphiken stellen die Daten der interlingualen Formkongruenzen für T3 dar, d.h. für drei Lemmata aus einer vier Sprachen zählenden (romanischen) Teilsreihe. Gezählt werden die Abweichungen, also opake Lemmata oder Profilwörter über den Selektor (RoS*-X). In Frage kommen im Prinzip alle möglichen Grundmengen; X ist ein Platzhalter für ‚Zielsprache‘. Als Grundmenge wird zunächst deren erweiterte Form herangezogen (etwa für die Kohorte Französisch: [F OR Fw]). Zu unterscheiden ist wiederum zwischen der primären und einer sekundären Zielsprache. Die Kombination lautet daher (F OR Fw) für die primäre und z.B. „I“ für

die sekundäre Zielsprache. Die Syntax (RoS-I) aus (F OR Fw) MIT (I) generiert also alle Fälle für RoS-I in der Mischung beider Grundmengen. Da es sich um Opazitäten handelt, ist der Vergleichbarkeit halber die [Umrechnung in Transparenzfälle](#) angezeigt. Dies erfolgt, wie gesagt, aus der Subtraktion aller RoS*-X-Fälle abzüglich der Opazitäten jener der betroffenen Sprache.

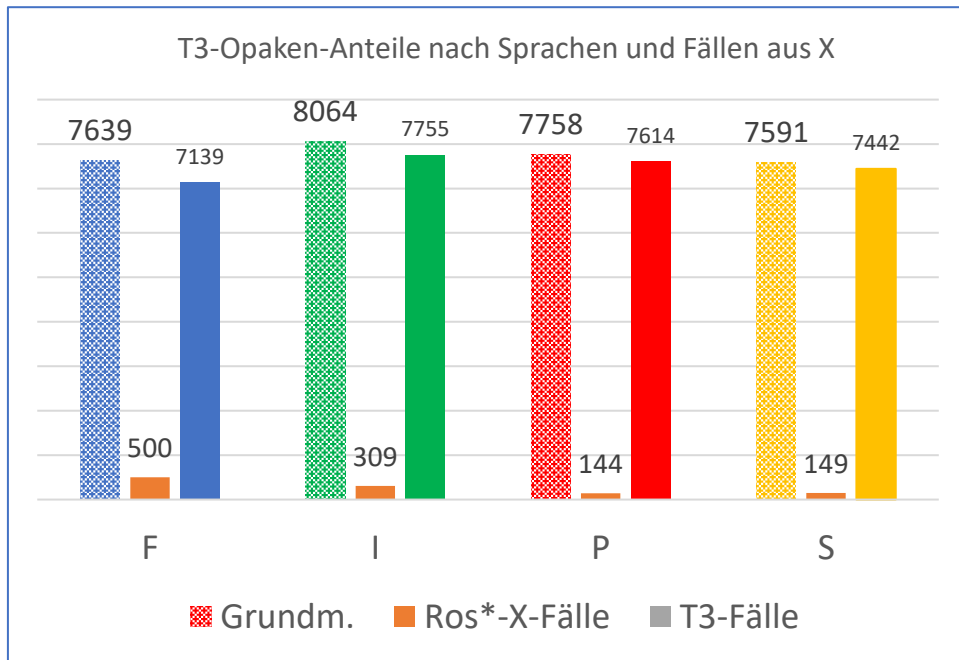


Abbildung 24: Anteil von T3 Transparenzfälle in den Symbolfarben der Zielsprachen³²

Die Grundmengen der einzelnen Sprachen unterscheiden sich deutlich voneinander (8064 für das Italienische, 7591 für das Spanische).

Die Nennung der prozentualen Anteile (Abb. 25) für T3 erleichtert den Vergleich zwischen den einzelnen Sprachen als primäre Zielsprachen. Der Selektor (RoS*-X) charakterisiert den Profilanteil der Sprache X. Der Opakenanteil (aus RoS*-X) ist natürlich größer als die Summe der Profilwörter der einzelnen Sprachen (RoS*-F, RoS*-I usw.). Um den realen opaken Umfang des KRM-Inventars für die einzelnen Zielsprachen zu ermitteln, sind den Anteilen aus den T3-Opazitätswerten jene von T0 hinzuzurechnen (855). Der bereinigte bzw. reale Opaken-Anteil von Französisch beläuft sich so auf 500 + 855= 1355 Serien. Der Opaken-Wert ist nicht mit dem Transparenzwert zu verwechseln.

³² Blau für Französisch, Grün für Italienisch, Rot für Portugiesisch und Orange für Spanisch.

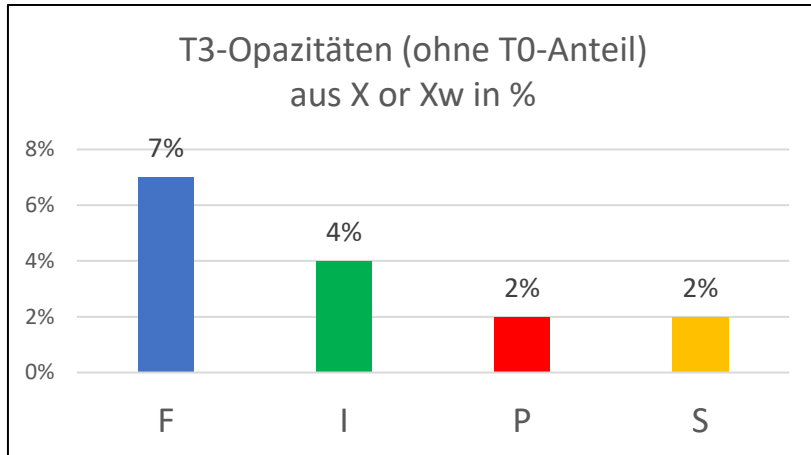


Abbildung 25: T3 mit RoSED-X

Die Werte – die, um den Umfang der Opaken zu berechnen, um den T0-Anteil erweitert werden müssen (10%) – sprechen für sich. Die Zahlen werden bei der Quantifizierung der zwischensprachlichen Transferanteile erneut eine Rolle spielen.

Syntaktisches zur Ermittlung der T3-Kohorte

Die folgenden Ausführungen erweitern am Beispiel der Grundmenge aus (F OR Fw) die Erklärungen zu den Grundlagen der Ermittlung der T3-Kohorte einer primären und einer sekundären Zielsprache.

Nr.	GrdM (F OR Fw)	Zielsprache X	Selektor	
1.	F oder Fw	F	RoS*-F	699
2.	F oder Fw	F	RoS-F	616
3.	F oder Fw	I	RoS*-I	311
4.	F oder Fw	P	RoS*-P	143
5.	F oder Fw	S	RoS*-S	149
6.	F oder Fw	S	RoS-S	117

26: Ermittlung von T3-Serien für FR-kombinierte erweiterte Grundmengen

Zur Grundmenge: Die Selektion (X OR Xw) macht Sinn, da alle verschiedenen RoS*-X-Vertreter erfasst und komputiert werden sollen. Der Selektor (X) würde diese in Teilen ausschließen.

Die Tabelle veranschaulicht, dass RoS-X eine Teilmenge von RoS*-X darstellt (rote Unterlegung). Zur Veranschaulichung werden die RoS-X-Selektoren in der obigen Tabelle (~~durchgestrichen~~) belassen.

Veranschaulichung: Die Formel in Nr. 1 sucht aus der Grundmenge (F OR Fw) alle RoS*-F (GrdM und sekundäre Zielsprache identisch). Die Kombination in Nr. 3 sucht

aus der Grundmenge (F OR Fw) alle RoS*-I (=sekundäre Zielsprache); erfasst werden nur intralingual-formdivergente italienische Lemmata bzw. deren Serien (Typ: *dédain-disprezzo-desdém-desdén*). Würde das Gesamt der sekundär-zielsprachlichen Entsprechungen gesucht, so lautete der Selektor im Feld Lit (I OR Iw). In Kombination mit RoS*-I würden dann 327 Fälle ausgewiesen; mit RoS*-F 616, mit RoS*-P 150. Expliziter gesagt: In der Grundmenge aus (F OR Fw) und (I OR Iw) ließen sich DANN 150 portugiesische Profilwörter identifizieren.

Serien-Beispiele für Selektoren der sekundären Zielsprachen:

- a) RoS*-F: *brouillard-nebbia-nevoa-niebla* (sucht nach opak F für RoS*-Kennung)
RoS*-I: *fra-entre-entre-entre* (sucht nach opak I für RoS*-Kennung)
- b) RoS*-I weitet die Suche von RoS-I auf RoSE-I und RoSED-I aus
- c) RoS*-* - weitet die Suche auf alle Zielsprachen aus (eine opake Form kann (maximal) in einer der Zielsprachen pro Serie auftauchen)

Achtung: RoS*-X erfasst keine RoS2- und RoS3-Markierungen. Lemmata dieser Kennung fallen nicht in die Kategorie T3.

Anwendungen der RoS*-X-Selektoren:

- a) RoS*-X dient zur Identifikation der T3-Opazität in einer bestimmten Zielsprache bei einer gegebenen Grundmenge, vorzugsweise (X OR Xw); in Vergleichen sind die Opazitäts- in Transparenzwerte umzurechnen.
- b) Für eine **primäre Zielsprache** ist die Kombination Sprache(n) der Grundmenge (aus X OR Xw in der Rubrik ROS) zu wählen. Die Selektion antwortet auf die Frage: Wie viele z.B. französische Profilwörter (RoS*-F) lassen sich in der Kohorte aus (F OR Fw) zählen?
- c) Für eine **sekundäre Zielsprache** ist die Kombination RoS*-X mit der Grundmenge (X OR Xw; ≠ Zielsprache in der Rubrik ROS) zu wählen.
- d) der Selektor RoS*-* ist anzuwenden, wenn die Eigenschaften aller T3-Selektoren generell von Interesse sind.

Um den gesamten Opazitätsumfang des KRM-Vokabulars einer einzigen Zielsprache und des Anteils der intralingualen Formdivergenzen zu bestimmen, ist die oben erfolgte Berechnung erneut, diesmal zur Grundmenge X (ohne Reduplikationen), durchzuführen:

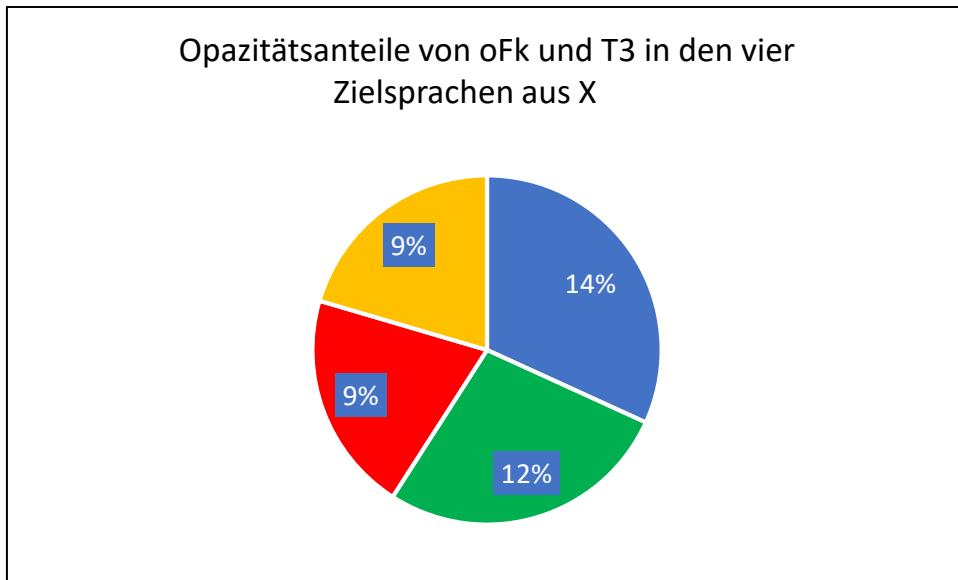


Abbildung 26: Maximale Opaken-Anteile aus monolingualer X-Zählung (in den Symbolfarben der Sprachen)

Auch hier bestätigt sich: Den größten Opakenanteil stellt das Französische, vor dem Italienischen und den beiden iberischen Sprachen.

Unschärfen in gemischter Selektion bei T3: Die Selektion (*-I OR *-S) führt, wie gesagt, zur Auswahl aller Fälle für (*-I) und (*-S). Dabei ist zu beachten, dass innerhalb einer Serie immer nur eine Intransparenz betroffen ist (ansonsten hieße die Kennzeichnung [RoS*2]):

RoS-I	chèque	CHEQUE	CHEQUE	ASSEGNO
RoS-S	CHIEN	CÃO	PERRO	CANE
RoSE-I	COMPARER	COMPARAR	COMPARAR	PARAGONARE
RoSED-S	cravate	GRAVATA	corbata	cravatta
RoS-S	DATE	DATA	FECHA	DATA
RoSE-I	dédain	DESDÉM	DESDÉN	DISPREZZO
RoSE-S	DESSIN	DESENHO	dibujo	DISEGNO

27: RoS*-X-Serien-Typen aus *-I OR *-S in ROS-

Der aus dem komponierten Selektor (*-X OR *-Y) entstehende Wert ist natürlich kleiner als jener aus (X) oder (Y) bezogen auf (*-*). Da der Opazitätswert bzw. der Subtrahend (*-X) kleiner sind, erhöht sich der Transparenzwert. Allerdings umfasst dieser sowohl die (*-X) als auch die (*-Y)-Werte. Um nun den realen Transparenzwert für eine sprachlich gemischte Serie zu ermitteln, ist jeweils sowohl (*-X) als auch (*-

Y) einzeln zu zählen und gesondert auszuweisen. In der I MIT F-Selektion aus Abb. 28 (auf der folgenden Seite) beträgt der T3-Wert für F: 466, der für I: 653.

Der Operator „verschieden von“ (<>) erlaubt es, bestimmte Variablen einer Selektion auszuschließen. Die Syntax zum Befehl ‚Generiere eine Liste aus der gemischt-sprachlichen Grundmenge von (P OR Pw) unter Ausschluss aller (R) :

Rf ▾	Ri ▾	Rp ▾	Rs
R§	M§	R§	R-
R-	M§	R-	R§
R§	M§	R§	R§
R§	M§	R-	R§
R	M§	R-	R§
R-	M-	R§	R
R	M	R§	R§
R	M	R§	R
R§	M§§	R-	R-
R§	M§	R§	R§
R-	M-	R-	R§
R§	M-	R§	R§
R§	M§§	R§	R
R	M§	R-	R
R§	M§	R§	R§
R§	M-	R-	R-
R-	M-	R-	R-
R	M§	R§	R§
R§	M§§	R§	R§

28: Beispiel für eine <>-Selektion

Wie erwartet, zweigt die Rubrik Rp keinen einzigen Fall von (R). (Die Kennung steht für das Frequenzintervall <2000.

Es war vielfach von der Syntax *-* (= RoS*-*) die Rede, ohne dass die Selektoren konkretisiert worden wären. Die entsprechenden-Selektoren können folgende Form annehmen:

RoS-F
RoS-I
RoS-P
RoS-S
RoSE-F
RoSE-I
RoSE-P
RoSE-S
RoSED-F
RoSED-I
RoSED-P
RoSED-S

Abbildung 27: Ausschnitt aus der Selektion RoS-* bzw. *-*

Abb. 28 präsentiert die bereinigten T3-Transparenzwerte für sprachlich-kombinierte Mengen:

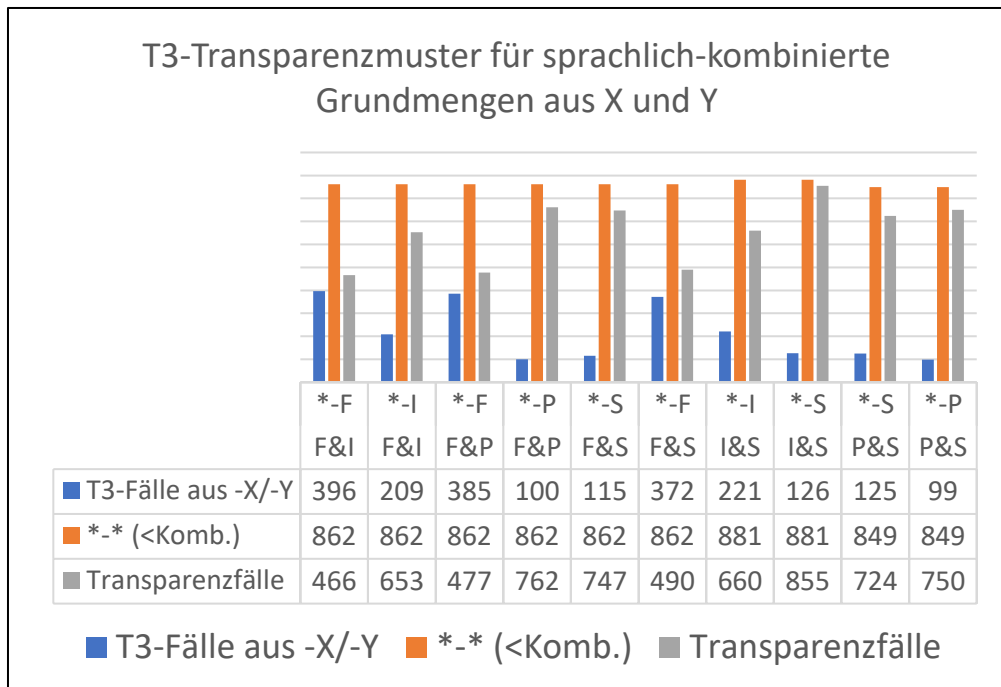


Abbildung 28: Opazität und Transparenz in T3-Selektionen aus X MIT Y

Legende: Die sprachlichen Kombinationen der Grundmengen sind durch & markiert, also F&I. Dabei nimmt die erstgenannte X, die zweitgenannte Y an. Gleiches gilt für die Opazitätsmarkierungen und deren jeweilige Referenzsprache. *-F = 396 bezieht sich auf X, *-I = 209 auf Y. *-* meint den Minuenden, *-X oder *-Y die Subtrahenden. Die Differenz ergibt den Transparenzwert: 466 für F&I aus RoS-F, 653 für F&I aus RoS-I. Zwei Beispiele:

F&I = 399 < *-F in ROS mit F in Lfr und mit I in Lit

F&I = 209 < *-I in ROS mit F in Lfz und mit I in Lit

Sprich: Der T3-Wert aus den französischen Profilwörtern in Kombination mit Italienisch beträgt 399 Fälle; der aus den italienischen Profilwörtern in derselben Kombination 209. Insgesamt sind aus der Grundmenge von Französisch 862 Eintragungen opak, aus der von Italienisch 881 und aus der von Portugiesisch 849. Die verbleibenden T3-Serien sind für RoS*-F mit Italienisch im Umfang von 466 Fällen interlingual transparent, die von RoS*-I im Umfang von 653.

Die blauen Säulen signalisieren Opazitätsfälle, die grauen die Transparenzraten. Die braunen Kolumnen die Gesamtheit der opaken Fälle.

Abb. 29 zeigt neben den Transparenzklassen detailliert die Werte für die Berechnung der T3-Transparenzraten nach ihren sprachlichen Spezifikationen, also nach RoS*-F, RoS'-I usw. differenziert.

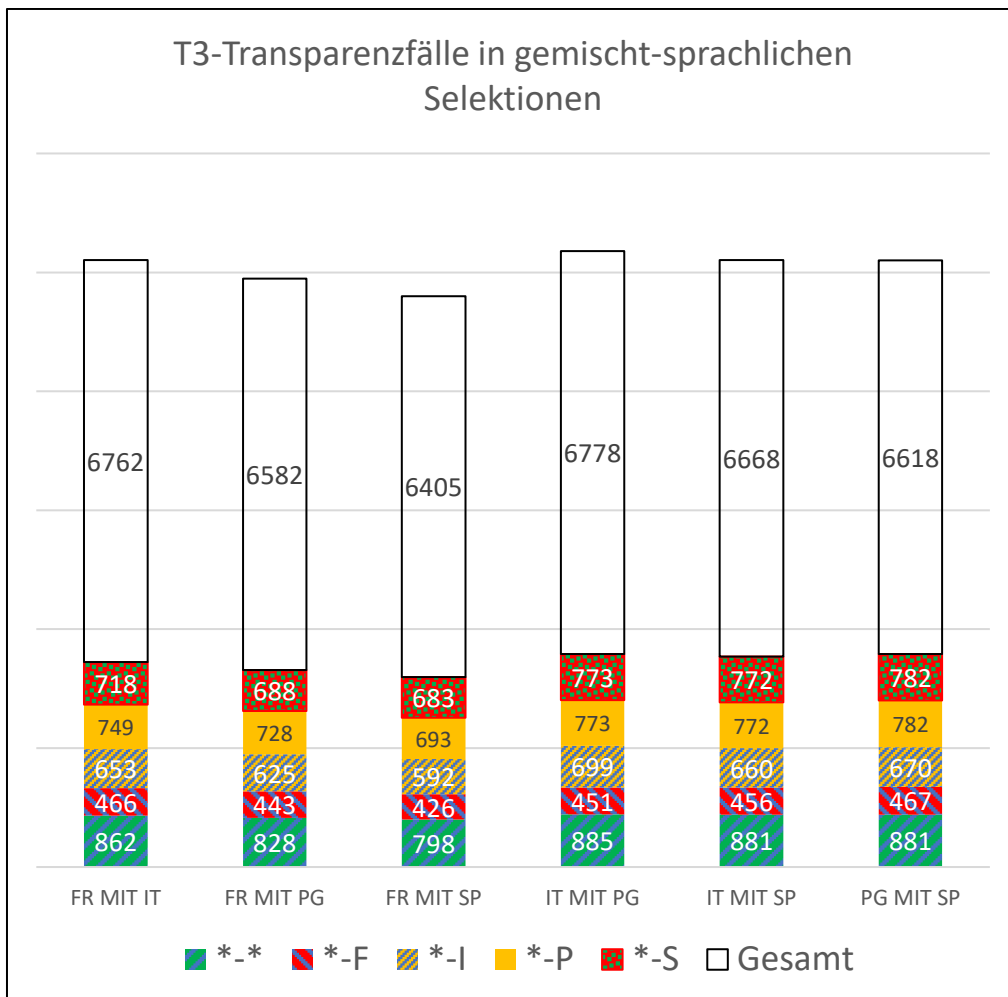


Abbildung 29: Transparenzraten von T-3 für unterschiedliche *-X aus gemischt-sprachlicher Ziehung

Die Graphik – sie spezifiziert das zu den T3-Werten Gesagte nach Profilwörtern für die einzelnen Sprachen – veranschaulicht, warum der T3-Transparenzwert für die unterschiedlichen (RoS*-X) im Prinzip differenziert nach den Variablen von X zu legen ist. Es zeigt sich, dass die Unterschiede zwischen den einzelnen (*-X)-Gruppen von Sprache zu Sprache vergleichsweise wenig zu Buche schlagen. Die Werte signalisieren, dass die Zählungen pro Profilwort nach Sprache zwar unterschiedlich ausfallen können, dass sich die Differenz allerdings nach der entsprechenden Grundmenge richtet: Die Grundmenge aus Italienisch mit Portugiesisch stellt die größte Grundmenge dar und jeder einzelne Wert für den profilwörtlichen Anteil ist größer als der jeder anderen Kolumne – allerdings mit der Ausnahme Portugiesisch mit Spanisch. Der Grund ist, dass die Opazitätsraten zwischen beiden iberischen Sprachen kleiner ist als jene zwischen anderen Sprachen.

Die entsprechenden Prozentwerte erlauben den zwischensprachlichen Vergleich. Dies zeigt z.B. die (*-I)-Selektion: für FR MIT IT: 9,65%, für FR MIT PG: 9,49%, für FR MIT SP: 9,24%; für IT MIT PG: 10,31%; IT MIT SP: 9,89% und PG MIT SP: 10,12%. Der maximale Unterschied beträgt 1,07%. Weitere Vergleiche ergeben durchaus ähnliche Ergebnisse. Dies legt den Schluss nahe, dass es für weitere Vergleiche zu den verschiedenen Transparenzklassen ausreicht, einen einzelnen (*-X) zu berechnen und einen Bias von 1% zu berücksichtigen.

8.4 \$-Markierungen

Obwohl die KRM-Struktur ausreichend Selektoren für die Identifikation unterschiedlicher Transparenzstufen entwickelt hat, kommt es vor, dass innerhalb einer solchen Kategorie selbst eine nochmalige Stufung angedeutet erscheint. Hierzu dient neben der \$-Markierung. Solche Profilmarkierungen (mit „\$“) erlauben die Identifikation von Transparenz- oder Opazitätsrelationen innerhalb gebrochener Serien. Dies betrifft Serien der Reihen „RoS*2“ (531 Serien mit zwei Transparenzpaaren oder 2 opaken Lemmata) und „RoS*3“ (1050 Serien mit 3 verschiedenen Wortstämmen: *blond-biondo-loiro/louro-rubio*; [T1](#)). 1582 umfasst das Gesamt der \$-Serien. Dies entspricht aufgerundet 17% des Inventars.

Profilmarker zur Kennung \$

<i>P</i> -Marker		<i>Bedeutung für interlinguale Transparenz/Opazität</i>
f\$ip\$ps	FR≠IT≠PG~SP	Exklusive Transparenz zwischen PG-SP (nicht zu FR und IT noch zwischen denselben)
f\$ip\$ss	FR≠IT~PG≠SP	Exklusive Transparenz zwischen IT-PG (nicht zu FR und SP noch zwischen denselben)
f\$is\$sp	FR≠IT~SP≠PG	Exklusive Transparenz zwischen IT-SP (nicht zu FR und PG noch zwischen denselben)
fi\$ps\$ss	FR~IT≠PG≠SP	Exklusive Transparenz zwischen FR-IT (nicht zu FR und PG noch zwischen denselben)
fi\$ps	FR~IT≠PG~SP	Transparenz zwischen FR und IT einerseits und PG und SP andererseits
fp\$ip\$ss	FR~PG≠IT≠SP	Transparenz zwischen FR-PG (nicht zu IT, nicht zu SP, noch zwischen denselben)
fp\$ip\$is	FR~PG≠IT~SP	Transparenz zwischen FR-PG einerseits und IT-SP andererseits
fs\$ip\$sp	FR~SP≠IT≠PG	Transparenz zwischen FR-SP (nicht zu IT und nicht zu PG, noch zwischen denselben)
fs\$ip	FR~SP≠IT~PG	Transparenz zwischen FR-SP einerseits sowie zwischen IT-PG andererseits

Um eine konkrete Vorstellung von den Größenordnungen der \$-Markierungen zu vermitteln, erfolgen folgende detaillierte Zählungen.

Romanische \$-Profilierung			\$-Profile nach Sprachen (formdiv. Lemmata „F mit Z2-Profil)			
Nr.	Kennung	Anzahl	FR	IT	PT	SP
1	f\$î\$ps	639	391	461	425	417
2	f\$îp\$S	41	50	47	45	42 ³³
3	f\$îs\$P	80	50	50	55	45
4	fi\$P\$S	176	124	139	123	117
5	fi\$P\$S	423	270	293	315	309
6	fp\$î\$S	55	38	44	36	37
7	fp\$îs	41	22	28	23	24
8	fs\$î\$P	59	37	51	37	42
9	fs\$îp	35	22	27	23	25
		1549				

29: Auszählung der \$-Markierungen

Innerhalb horizontaler Felder zeigen gleiche Farben Transparenzkombinationen an, der Übergang von einer Farbe zur anderen eine Transfersperre. Die Ziehung des gesamtromanischen Profils (weiße Kästchen) erfolgte aus der Mutterdatei, gefiltert in PROFIL nach Kennungen, z.B. „fi\$ps“; die Ziehung der einzelsprachlichen Profile nach z.B. „fi\$ps“ im Feld PROFIL in Kombination mit „F“ in der Rubrik Lfr. Die leichten Abweichungen zwischen den \$-Profilen der Formdivergenzen der Einzelsprachen erklären sich durch die unterschiedlichen Grundmengen für FR, IT usw.

Tab. 29 zählt die Fälle, die in den Zielsprachen insgesamt oder in den einzelnen Zielsprachen für die jeweilige \$-Kennung zutreffen. Insgesamt beträgt der exklusiv romanische Anteil (RoS*), wie erwähnt, 12,6 Prozent aller Serien.

Interpretation: Auffälligkeiten betreffen die Muster 1 und 5. Sie deuten (was nicht überraschen kann) auf sehr hohe Formkonvergenzen zwischen den beiden im KRM berücksichtigten iberoromanischen Sprachen hin. Auch zwischen dem Französischen und dem Italienischen befindet sich ein bemerkenswert hoher Kongruenzanteil (4 und 5). Die Anteile von Französisch mit den beiden weiteren Schwestersprachen sind deutlich geringer. Die gemeinsamen Formen zwischen dem Italienischen und Portugiesischen (2, 9) bzw. dem Italienischen und Spanischen (3, 7) sind ebenfalls deutlich kleiner.

Die folgenden Zählungen betreffen, wie gesagt, ausschließlich den RoS*2 und RoS*3-markierten Selektorenanteil im Feld ROS-.

³³ Die kursiv gesetzten Füllungen entstammen der Selektion F mit S (ohne Wiederholung).

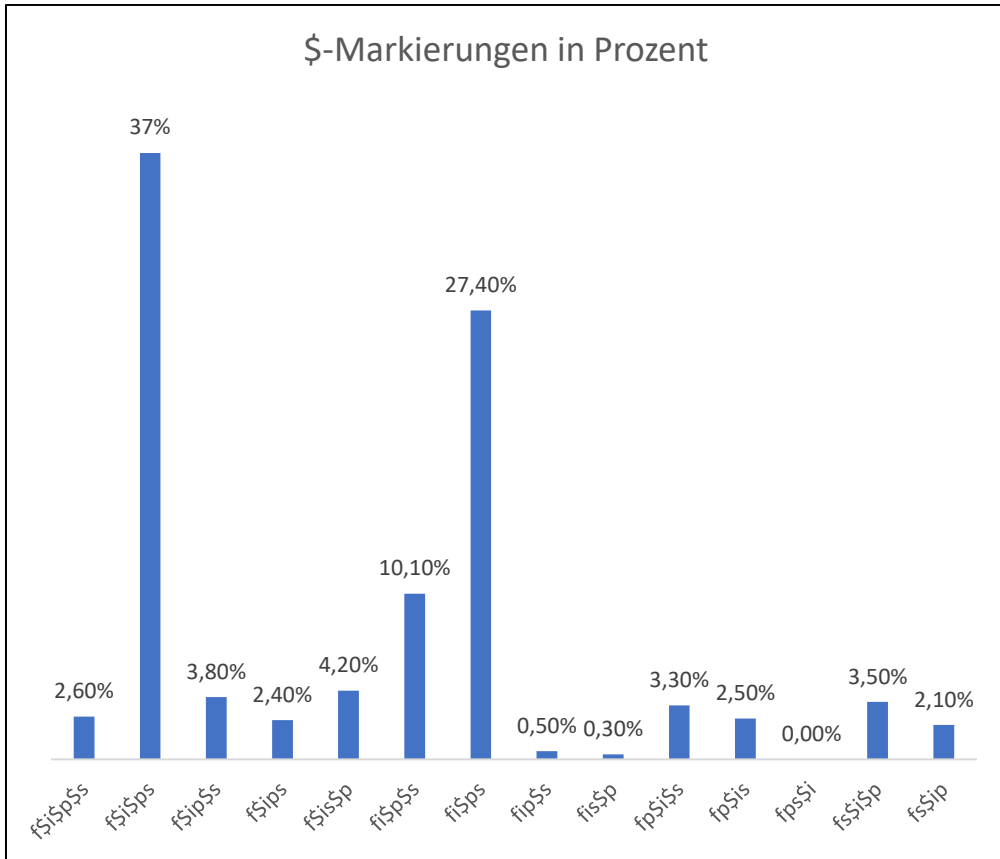


Abbildung 30: Prozentanteile von \$-Markierungen

Bemerkenswert ist der hohe Anteil der portugiesisch-spanischen Konvergenzen von 64,4%, gefolgt von 38,3% französisch-italienischer Formähnlichkeiten. (Alle Werte beziehen sich auf eine Grundmenge von 17% der Gesamtzahl der Serien.)

Abb. 31 zeigt eine zufällig getroffene Auswahl der Selektion. Bemerkenswert sind die Spalten ROS- und PROFIL mit RoS* und *\$*-Variablen:

Fr	Pg	Sp	It	RoS-	PROFIL
MARCHE	CURSO	MARCHA	andamento	RoS3	F\$I\$P
appartenant	pertencente	perteniciente	appartenente	RoS2	FI\$PS
CI-APRÈS	DEPOIS	DESPUÉS	appresso	RoS3	F\$I\$PS
approfondi	exaustivo	exhaustivo	approfondito	RoS2	FI\$PS
ESSUYER	secar	secar	ASCIUGARE	RoS3	F\$I\$PS
ESSAYER	PROVAR	probar	assaggiare	RoS2	FI\$PS
aménagement	ESTRUTURA	ESTRUCTURA	assetto	RoS3	F\$I\$PS
ATTENDRE	ESPERAR	ESPERAR	attendere	RoS2	FI\$PS
balancer	balançar	tirar	dondolare	RoS3	FP\$I\$S
basculer	cair	CAER	CADERE	RoS3	F\$I\$PS
déploiement	desenvolvimento	despliegue	spiegamento	RoS2	F\$IS\$P
désoler	entristecer	entristecer	affliggere	RoS3	F\$I\$PS
détachement	desprendimento	desapego	distacco	RoS3	FI\$P\$S
détenteur	poseedor	tenedor	detentore	RoS3	FI\$P\$S
détresse	perigo	PENA	pericolo	RoS3	F\$IP\$S
biais	obliquo	sesgo	obliquità	RoS3	F\$IP\$S
blâmer	censurar	culpar	biasimare	RoS3	FI\$P\$S
blessure	lesão	herida	ferita	RoS3	F\$IS\$P
bond	SALTO	salto	balzo	RoS3	F\$I\$PS
bonsoir	BOA NOITE	BUENAS TARDES	BUONASERA	RoS3	FI\$P\$S
bordel	confusão	quilombo	bordello	RoS3	FI\$P\$S
bossier	TRABALHAR	TRABAJAR	LAVORARE	RoS3	F\$I\$PS
bouleversement	perturbação	perturbación	sconvolgimento	RoS3	F\$I\$PS
bouleverser	perturbar	perturbar	sconvolgere	RoS3	F\$I\$PS
boulot	TRABALHO	TRABAJO	LAVORO	RoS3	F\$I\$PS
bousculer	empurrar	apurar	urtare	RoS3	F\$I\$PS
brigade	brigada	EQUIPO	squadra	RoS3	FP\$I\$S

Abbildung 31: \$-Markierungen in T2 und T1

Pädagogische Relevanz: Wie die empirische Lernforschung zum interlingualen Identifikationstransfer belegt, ist es für die Realisierung eines solchen Transfers durchaus nicht unerheblich, aus welcher Richtung (Transferbasis → Transfereffekt) bzw. von welcher zuletzt aktivierten Transferbasis aus eine sprachenübergreifende Identifikation versucht wird. Auch vor diesem Hintergrund ist es didaktisch relevant, mögliche Transfersperren bzw. -erleichterungen zu erkennen und zu kennzeichnen. Dies zu leisten ist die Funktion der \$-Markierung.

8.5 FaX-Marker zur Kennzeichnung minimaler morphologischer Distanz

FaX-Markierungen unterstreichen eine *minimal-relative* Formdivergenz der Lemmata im Vergleich zu jenen der romanischen Schwestersprachen unterhalb der mit der RoS*-Kennung bezeichneten morphologischen Distanz. FaX-Markierungen lauten FaF, FaI, FaP oder FaS. Die Kennungen sind im Feld PROFIL verzeichnet.

TON	TONO	TOM	TONO	tone, sound	Farbton, Klangton	FaP
TRAITÉ	TRATTATO	tratado	tratado	treaty	Vertrag	FaF
TREMBLER	TREMARE	TREMER	TEMBLAR	shake	zittern	FaS

30: Beispiele für FaX-Markierungen

Die Zählungen in PROFIL erfolgen aus der Selektion der RoS*-Reihe in ROS-. Der praktische Nutzen der Markierung wird in der Möglichkeit einer zielgenaueren Konstruktion des Inputs von Übungen und Übungsformaten der KRM-Lernapps gesehen.

	ROMANISCH	FR	IT	PG	SP
FaF	204	175	174	164	158
FaI	55	41	49	41	43
FaP	61	50	51	48	49
FaS	32	26	27	25	30
GESAMT	352	292	301	278	80

31: Komputation der FaX-Markierungen³⁴

Wie die FaX-Tabelle zeigt, ist der Anteil mit einer *minimal-relativen* Formdivergenz bei den französischen Lemmata besonders hoch. Begründet ist auch dieses in dem hohen germanischen Adstrat-Anteil man denke allein an die Markierung „fränkisch“ im Feld ETYMA – sowie der starken Formdivergenz zwischen den französischen Erb- und Buchwörtern (Typ *château, cher...* vs. *castillo, caro...*; *mère* vs. *madre, eau* vs. sp. *agua*/it. *acqua, croire* vs. *crédible* it. *credere, credibile* u.a.m. Die Formdistanz verbindet sich oft mit. Die Kenntnis dieser buchwörtlichen Formen begründet einen wichtigen Lernvorteil von Französisch als Brückensprache, weil diese in ganz erheblichem Maße Transparenz zu den Schwestersprachen schaffen. In umgekehrter Richtung lässt sich dies für die anderen romanischen Sprachen einerseits und den erbwörtlichen Anteil des Französischen andererseits nicht sagen (Beispiel: erbwörtlich die schon erwähnte Serie *beaucoup-molto-muito-mucho* und buchwörtlich:

³⁴ Zählung aus X. Die Komputation aus (X OR Xw) ergibt den hier als romanisch bezeichneten Wert.

moulte (>5000)-*molto-muito-mucho*. Vielen Lernern sind diese ‚gelehrten‘ Formen auch aus dem Englischen bekannt oder erschließbar (*maternity/ maternité, much*).

Didaktisch sind die \$- und die FaX-Befunde für eine optimale Selektion des Inputs relevant, weil sie innerhalb der 2- und 3stufigen Transparenzkohorten einerseits jene Lemmata zwischen den primären oder sekundären Zielsprachen identifizieren, die nicht mehr eigens neu erlernt werden, oder aber, andererseits, deren Formen in der Tat neu erlernt werden müssen. Die Steuerung des Unterrichts kann dementsprechend lernerbezogen den Input auswählen und verfahren. Die Auswahl des Vokabulars in den Lernapps erfolgt entsprechend.

Die Zählungen der \$- und FaX-Markierungen betreffen– daran ist zu erinnern – lediglich ein Segment von 1583 Serien (17%). (Der KRM umfasst allerdings 9556.) Um die interlingualen Transparenzanteile auf der Grundlage einzelner Transfersprachen erster Ordnung im Hinblick auf einzelne Zielsprachen zweiter Ordnung zu errechnen, ist prinzipiell von allen formverschiedenen Lemmata aller Zielsprachen, seien sie nun erster oder zweiter Ordnung, auszugehen. Die Quantifizierung dieser Transferformen erfolgt im Weiteren.

9. Methodisches zu sprachlich-kombinierten Komputationen

9.1 Mögliche Sprachkombinationen für den zwischensprachlichen Transfer

Da jeder interlinguale Transfer potentiell [bidirektional](#) greift, sind Doppelungen von Sprachkombinationen auszuschließen. Abb. 32 verdeutlicht nochmals die möglichen Kombinationen nach Sprachen:

Mögliche Transferkombinationen der vier Zielsprachen

FI/IF	FP/PF	FS/SF
IP/PI	IS/SI	
PS/SP		

Abbildung 32: Mögliche Sprachen-Paare und Sprachkombinationen

Die Beschränkung auf die romanischen Sprachen darf nicht darüber hinwegtäuschen, dass vor allem das [Englische](#), wie gezeigt, ein erhebliches Transferpotential (von 4248 Lemmata) bereitstellt. Ein Charakteristikum der englischen Transferbasen ist, dass sie zumeist nicht nur in einer einzigen romanischen Sprache präsent sind: 1548 Serien teilt das Englische exklusiv mit den romanischen Sprachen, 2800 greifen über die Romania hinaus. (Natürlich werden englische Transferbasen nicht ‚aus-sortiert‘.)

9.2 Syntax der Selektoren

Um die Zählungen nachvollziehbarer zu machen, erscheint es vorteilhaft, sich mit der Methode der Such- und Zählweisen weiter vertraut zu machen. Während die bisherigen Überlegungen – sieht man von den \$-Markern ab – vor allem die einzelnen Zielsprachen als Transfer- oder primäre Zielsprachen beschrieben, geht es nun um das Vergleichen der Transparenzstufen in ausverschiedenen Zielsprachen gebildeten Grundmengen. Die folgenden Ausführungen zu [Syntax](#), *Grundmenge* und *Gesamt* beleuchten zunächst die Selektoren, mit deren Hilfe die Zählungen durchgeführt werden. Zur Erinnerung:

Selektoren zu Transparenzstufen *ohne oder mit primärer oder sekundärer zielsprachlicher Spezifizierung*

TRANSPARENZORTUNG	T-Stufen	Syntax / Selektoren
Serielle lexematische Opazität	Transparenz 0:	"oFk"
Transparenz in 2 Lemmata	Transparenz 1:	"RoS3" OR "RosE3" OR "RosED3" ³⁵ bzw. "*3"
Transparenz in 2 Paar Lemmata	Transparenz 2:	"RoS2" OR "RoSE2" OR "RoSED2" bzw. "*2"
Transparenz in 3 Lemmata	Transparenz 3:	"RoS-*" OR "RoSE-*" OR "RosED-*" bzw. "*_*"
Transparenz in 4 Lemmata	Transparenz 4:	"RoS" OR "RosD" OR "RosE" OR "RoSED"

32: Syntax für Ermittlung von Transparenzstufen

Wird ohne Sprachkennung in den Feldern Lf, Li, oder Ls gezählt, so werden Serien ohne jegliche sprachliche Indizierung erfasst. Der Selektor (P OR Pw) im Feld PG filtert hingegen für alle in anderen Feldern genannte Selektoren zur Sprache Portugiesisch. Die Selektion umfasst die intralingual-formdivergenten sowie die reduplizierten Lemmata. Da jede Serie Lemmata aus 6 Sprachen fasst (ohne Latein), erreicht die Anzahl der Menge aus (P OR Pw) oder („F“ oder „F OR Fw“) nahezu gleichermaßen hundert Prozent aller Serien.

Wie es in den Erläuterungen zum Begriff der Grundmenge anklang, ist es oft vorteilhafter, zwischen den Kohorten der jeweiligen intralingual-formdivergenten Lemmata, z.B. aus der Selektion „P“ mit „l“, zu vergleichen, da dies zu

³⁵ Da die Anzahl der Lemmata mit deutscher Formadäquanz ohne englische Entsprechung zu gering ist, um pädagogisch und statistisch relevant zu sein, wird der Selektor „RoSD“ zur Ermittlung on T1, T2 und T3 aus der Zählung genommen.

unterschiedlicheren Komputationsergebnissen führt. Diese Kombination stellt für 6779 Fälle eine Beziehung zwischen den Formdivergenzen des Portugiesischen und des Italienischen her; der Wert erhöht sich auf 7758, wenn die Grundmenge aus (I OR Iw) MIT (P) konstruiert wird. Diese Mengung erfasst allerdings nicht die Formredundanzen aus dem Gesamtinventar des Italienischen. Hierzu bedürfte es natürlich der Selektion „I OR Iw“ mit „P OR Pw“: 9545. Eine Zählung der Daten aus den verschiedenen RoS*-Selektoren selbst bringt weitere Klarheit: RoS*erfasst 8673 von **9555** Serien³⁶, und zwar in folgender Verteilung nach Selektoren und Transparenzklassen:

Selektor	Fälle	T-Klasse
RoS	1203	T4
RoS2	487	T2
RoS3	1040	T1
RoSD	54	T4
RoSD2	0	T2
RoSD3	2	T1
RoSE	1546	T4
RoSE2	35	T2
RoSE3	5	T1
RoSED	2800	T4
RoSED2	12	T2
RoSED3	2	T1
RoSED-F	26	T3
RoSED-I	25	T3
RoSED-P	20	T3
RoSED-S	18	T3
RoSE-F	57	T3
RoSE-I	34	T3
RoSE-P	18	T3
RoSE-S	29	T3
RoS-F	616	T3
RoS-I	324	T3
RoS-P	150	T3
RoS-S	170	T3
Σ1	8673	Z

33: Selektoren der RoS-Reihe mit T-Stufen

³⁶ Zählung aktualisiert am 8.1.2018. Zur Erinnerung: Da die Datenbankpflege eine ständige Aktualisierung erfordert, sind sehr leichte Abweichungen unterhalb von 1 Prozent möglich.

Vermessung des romanischen Kernwortschatzes

Selektor	Fälle	T-Klasse
RoS	1203	T4
RoSD	54	T4
RoSE	1546	T4
RoSED	2800	T4
RoSED-F	26	T3
RoSED-I	25	T3
RoSED-P	20	T3
RoSED-S	18	T3
RoSE-F	57	T3
RoSE-I	34	T3
RoSE-P	18	T3
RoSE-S	29	T3
RoS-F	616	T3
RoS-I	324	T3
RoS-P	150	T3
RoS-S	170	T3
RoS2	487	T2
RoSD2	0	T2
RoSE2	35	T2
RoSED2	12	T2
RoS3	1040	T1
RoSD3	2	T1
RoSE3	5	T1
RoSED3	2	T1
Σ1	8673	Z

Σ1	8673	Z
RoSED	2800	T4
RoSE	1546	T4
RoS	1203	T4
RoS3	1040	T1
RoS-F	616	T3
RoS2	487	T2
RoS-I	324	T3
RoS-S	170	T3
RoS-P	150	T3
RoSE-F	57	T3
RoSD	54	T4
RoSE2	35	T2
RoSE-I	34	T3
RoSE-S	29	T3
RoSED-F	26	T3
RoSED-I	25	T3
RoSED-P	20	T3
RoSED-S	18	T3
RoSE-P	18	T3
RoSED2	12	T2
RoSE3	5	T1
RoSD3	2	T1
RoSED3	2	T1
RoSD2	0	T2
Selektor	Fälle	T-Klasse

34: Sortierung nach T-Stufen und nach Fällen

T3-Selektoren benötigen zur Ermittlung des Transparenzwertes das in Abschnitt 8.3 vorgestellte Verfahren. In gemischt-sprachlicher Sortierung (etwa aus der Grundmenge FR MIT IT) müssen, wie ebenfalls gesagt, für jeden RoS*-X-Typ zwei Werte getrennt ausgewiesen werden (z.B.: RoS*-F und RoS-I).

Nicht erfasst werden in Tab. 35 die Serien der Selektoren:

EN	22
oFk	855
T	5
Σ	882

35: Diskriminierte Selektoren in der T3-Auswahl

Summen aus Σ 1 und Σ 2

Σ 1	8673
Σ 2	882
Σ	9555

36: Kontrolle der Zählungen im Januar 2019

Offensichtlich verarbeiten die verwendeten Selektoren das Gesamt des KRM-Inventars. Abweichungen sind, wie ebenfalls gesagt, in sehr geringem Umfang durch leichte Eingriffe in die Datenbank (Pflege: Korrekturen, Verbesserungen) möglich. Diese Liste wird ständig aktuell gehalten, so dass das jeweilige Bias berechnet werden kann. Korrekturen erfolgen, wenn das Bios eine kritische Größe erreicht hat.

Da mit den Transparenzstufen mögliche zwischensprachliche Transferaktivitäten gezählt werden, ist daran zu erinnern, dass solche Transferprozesse [bidirektional](#) wirken (wie im Abschnitt Latein erläutert). Dies rechtfertigt die Verwendung des Symbols „ \leftrightarrow “. Es ergeben sich für die Komputation von bilingual-zwischensprachlichen Transparenz-Korrespondenzen vier mögliche Kombinationen:

Kombinationsmöglichkeiten von zwischensprachlichen Formkongruenzen und Formdivergenzen

1) A Formdivergenzen aus Sprache X	\leftrightarrow	B Formdivergenzen aus Sprache Y
2) A Formdivergenzen aus Sprache X	\leftrightarrow	C Formdivergenzen u. -kongruenzen aus Y
3) D Formdivergenzen u. -kongruenzen aus X	\leftrightarrow	B Formdivergenzen aus Sprache Y
4) D Formdivergenzen u. -kongruenzen aus X	\leftrightarrow	C Formdivergenzen u. -kongruenzen aus Y

37: Kombinationsmöglichkeiten von Sprachen nach Formkonvergenz u. -divergenz

Natürlich weisen die einzelnen Kohorten (A bis D) jeweils unterschiedliche Mengen von Lemmata auf: Gleiches gilt auch für die Kombinationen $A \leftrightarrow B$, $A \leftrightarrow C$, $D \leftrightarrow B$, $D \leftrightarrow C$. Wir haben es also stets mit unterschiedlichen Grundmengen zu tun. Allerdings, so lässt sich abschätzen, sind deren Unterschiede z.T. minimal. Eine konkrete Komputation zu Tab. 38 zur Kombination von Französisch mit Spanisch:

1)	S ↔ F	6405
2)	F ↔ S OR Sw	7636
3)	F OR Fw ↔ S	7591
4)	F OR Fw ↔ S OR Sw	9546

38: Auszählungen

In Zählkombinationen vom Typ X mit (Y OR Yw) ist stets der X-Wert aussagekräftiger, da (Y OR Yw) immer zu der größeren Vergleichsmenge führt, die nahezu dem Gesamt des KRM-Inventars entspricht. So wäre in 2) der bloße Sw-Wert 9546, d.i. identisch mit 4). Eine Regel lautet: In kombinierten Komputationen bestimmt die kleinere gemeinsame Teilmenge das Selektionsergebnis aus der gemeinsamen Schnittmenge.

Offensichtlich sind nicht nur die Kombinationen von Sprachen und Sprachenpaare vergleichenden Selektionen für die Komputation entscheidend, sondern selbstredend auch die Sprach-Selektoren selbst. Dies betrifft die Konstruktion der Grundmengen und ihrer möglichen Korrespondenzen. Hierzu ein weiteres Beispiel aus der Kombination von Französisch mit Italienisch:

Anzahl der Fälle für...

1	Serien	9555
2	F OR Fw	9549
3	I OR Iw	9550
4	F OR Fw mit I OR Iw	9549
5	F	7639
6	I	8064
7	F mit I OR Iw	7639
8	F OR Fw mit I	8063
10	F mit I	6762

39: Zählungen aus unterschiedlichen Grundmengen

Die sehr geringen Abweichungen insbesondere zwischen 1 und 4 ergeben sich aus den unterschiedlichen Anteilen der Teilmengen an den ausgeblendeten (nicht gezählten) bzw. „statistisch-irrelevanten“ Daten verschiedener Selektionen. Relevant erscheinen die rot gesetzten Zeilen, weil die von ihnen abhängigen Variablen Unterschiede ausweisen werden³⁷.

³⁷ Die Selektion zu den numerisch nahe beieinander liegenden, schwarz gesetzten Selektionen (‚Lemma der Sprache X‘ [‚S‘] OR ‚Lemma der Sprache X wiederholt‘ [‚Sw‘]) ergibt, wie schon erwähnt, stets einen nahezu hundertprozentigen Wert (weil ja jedes Lemma einer jeden Sprache innerhalb einer Serie eine Bedeutungsadäquanz hat). Dies bedeutet, dass sich für mögliche Kombinationen mit nur einem ‚X OR Xw‘-Selektor immer sehr nahe beieinanderliegende Werte ergeben müssen.

10. Komputationen von Transparenzprofilen in der Kombination von zwei und mehreren Sprachen

Das folgende Kapitel misst die interlingualen Transparenzstufen von gemischt-sprachlichen Selektionen für primäre und sekundäre Zielsprachen. Es schließt an Kapitel 8.1 an.

10.1 Grundmengen aus intralingual Formdivergenten und Reduplikationen

Es hat sich mehrfach gezeigt, wie relevant die Komposition der Grundmengen für die Komputation von Transferpotentialen ist. Dies führt an dieser Stelle zu folgender Darstellung: Abb. 33 illustriert die Anteile unterschiedlicher Grundmengen-Typen sowohl aus den intralingual Formdivergenten (X) als auch den intralingualen Reduplikationen (Xw), d.i. quasi das Gesamt aller Eintragungen (X OR Xw) pro Sprache. Beide Typen von Grundmengen sind natürlich Teilmengen des KRM-Inventars.

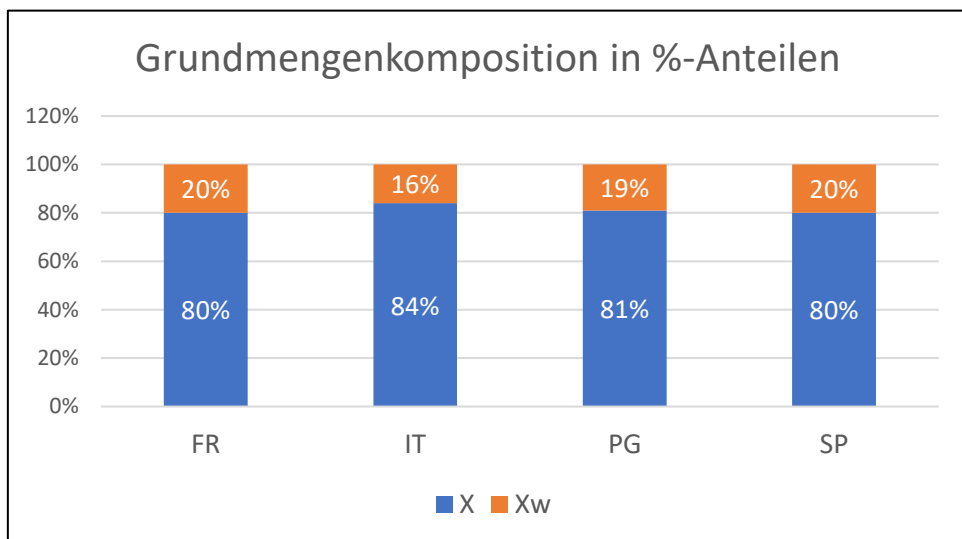


Abbildung 33: Komponentenanteile der erweiterten Grundmengen in %

Offensichtlich liegt der Differenzwert zwischen den intralingual formdivergenten (X) und den Reduplikationen $[(X \text{ OR } Xw) - (X)]$ je nach Sprache und Sprachenvergleich zwischen 16 und 20 Prozent. Der betroffene Unterschied zwischen beiden Gruppen (X OR Xw) bzw. (X) beläuft sich auf durchschnittlich 18,75 Prozent. Dies kann für Überschlagsrechnungen interessant sein.

Für die Grundmenge einer primären Zielsprache (hier aus [I OR Iw]) kombiniert mit einer sekundären (P in Lpt) und der Bestimmung des profilwörtlichen Anteils von (I)

ergibt sich das in Abb. 34 dargestellte Bild. Zur Identifikation der Profiwörter werden also zwei Sprachen, von denen eine mit jener des Grundgesamtts identisch ist (hier IT) herangezogen. Die Transparenzraten für T3 berechnen sich bekanntlich aus der Subtraktion ($[RoS^* - *] - [RoS - I]$) für die Zielsprache Italienisch (=843); für Portugiesisch ist der entsprechende Wert 945. Würde man die gemeinsame Transparenzrate beider Profiwörtergruppen errechnen wollen, so läge sie für IT mit PG bei 700 Okkurrenzen. Zur Ermittlung der T4-T3-Transparenz wäre diese Zahl dem Wert von T4 zuzuschlagen. Dies heißt, dass 74% der Sortierung interlingual hochgradig transparente Lemmata – 4 oder zumindest 3 Lemmata pro Serie sind formkongruent – umfasst; die formkongruenten Anteile aus T2 und T1-Kohorten nicht eingerechnet.

Entsprechende T3-Transparenzwerte für Französisch und Spanisch lassen sich ebenfalls leicht bestimmen.

Welchen praktischen Wert hat diese Aussage? Lerner, die z.B. zum Ausbau ihrer mehrsprachigen Lesekompetenz zuverlässige Auskünfte über Umfang und Art des zu erwerbenden Kernwortschatzes von Italienisch und Portugiesisch erfahren wollen, erhalten Informationen über den Umfang der möglichen interlexikalischen Transferpotenziale. Lehrende können das für Übungen notwendige Material entsprechend auswählen.

Der hohe Anteil von den Lernern bereits bekannten T4 und T3-Lemmata dürfte einen positiven Einfluss auf die Einschätzung ihrer Selbstwirksamkeit bzw. auf ihre Motivation bzw. Resilienz ausüben. Sie gewinnen von vornherein den Eindruck, dass die ‚Hürden‘ zum Erwerb von Lesekompetenz in Italienisch und Portugiesisch ausweislich des geringen T1 mit T0-Wertes (689+540) vergleichsweise gering sind. Dass es bei Verwendung des interkomprehensiven, zunächst auf den Erwerb von Lesekompetenz gerichteten Ansatz nicht bei der ausschließlich rezeptiven Verwendung der Zielsprache bleiben muss, bestätigen zahlreiche dokumentierte Lernerdaten. Lernende gewinnen leicht den Eindruck: „Es ist dumm, mein Lernpotenzial nicht zu nutzen.“

Doch zurück zu den Vergleichszählungen:

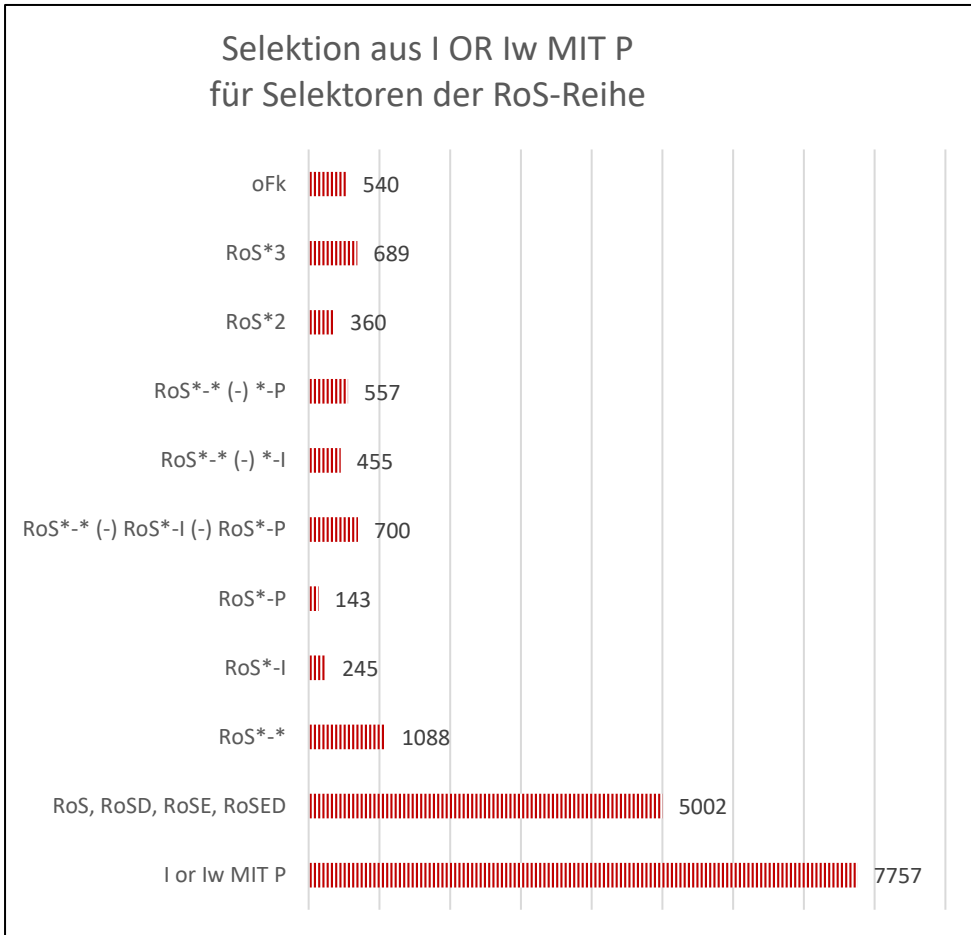


Abbildung 34: Erweiterte Grundmenge mit einer sekundären Zielsprache, Selektoren aus der RoS-Reihe mit T3-Differenzierung für IT und PG

Die Daten machen gewinnen Aussagekraft, wenn man sie mit Zählungen aus gemischt-mehrsprachlichen Ergebnissen aus einfacher Zählung vergleicht. Wechselt man die aus I konstruierte erweiterte Grundmenge zu (P OR Pw), so begegnen folgende Mengen:

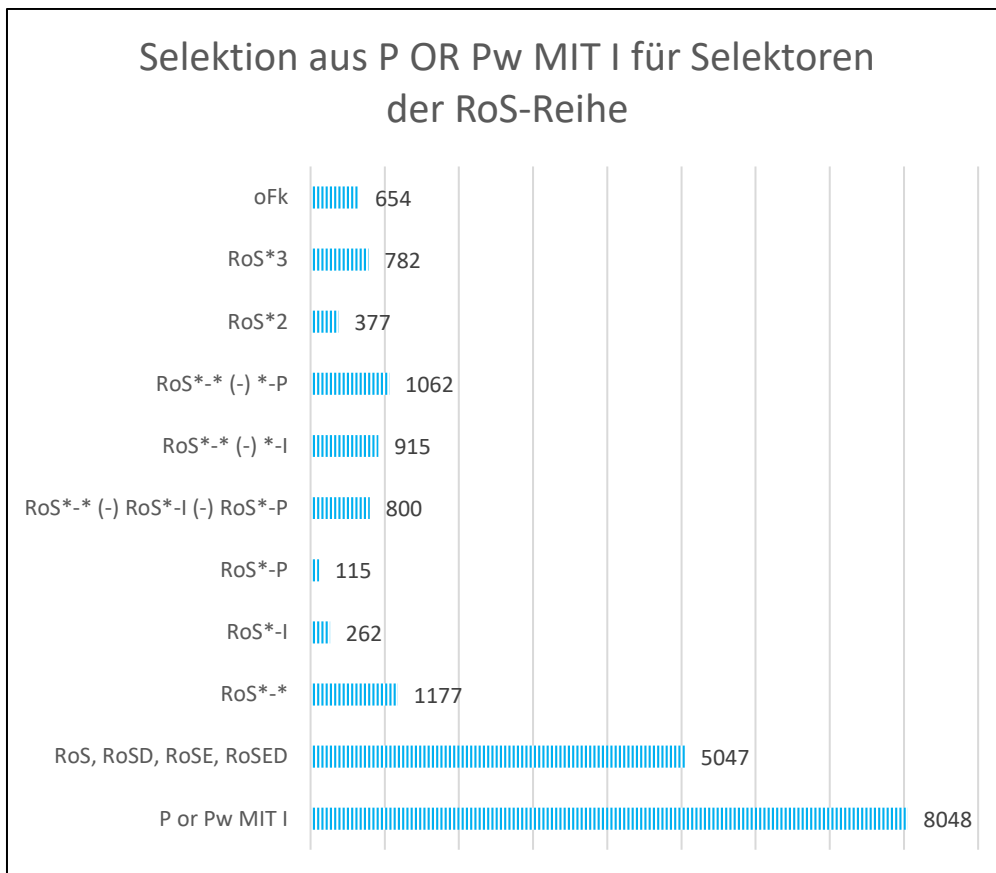


Abbildung 35: Erweiterte Grundmenge mit einer sekundären Zielsprache, Selektoren aus der RoS-Reihe mit T3-Differenzierung für IT und PG

Interpretation: Erwartungsgemäß führt eine größere Grundmenge zu höheren Anteilen der einzelnen Selektionen. Solche Mengen haben natürlich immer den Selektor (X OR Xw) oder in sekundärsprachlicher Markierung (Y OR Yw).

Zur Messung der Transparenzintervalle aus (P OR Pw) MIT (I): T4: 5047 (davon mit Englisch gemeinsame Anteile 3969); T3: 800 (ohne Differenzierung nach einzelsprachlichen Profilwörtern); T2: 377; T1: 782; T0: 654.

Im Verhältnis zu den anderen Selektoren lassen sich die Sprachen benennenden Variablen der Grundmengen, wie angeklungen, als „unabhängige Variablen“ bezeichnen, da z.B. in der Rubrik SP immer nur die Variable S in Gestalt von (S) oder (S OR Sw) und keine andere auftauchen kann; im Feld ROS- hingegen sind variiert Eintragungen und unterschiedliche, einfache und kombinierte Selektoren, einschließlich von Erweiterungen durch Operatoren (wie =, OR, AND, <>), möglich (z.B. ROS RoSD OR RoSE OR ROSED oder RoS*2 oder RoS-S). In Kombination miteinander bilden die Sprachrubriken, wie gesehen, jeweils neue Grundmengen, aus denen

heraus weitere Selektionen komputiert werden. Dieser Zusammenhang erklärt, weshalb selbst seriell einheitliche Ergebnisse erzeugende Selektoren wie „oFk“ zu unterschiedlichen Anteilen am Gesamt der aus unabhängigen Variablen gebildeten Grundmenge führen können.

10.2 Dreisprachige Grundmengen aus gemischt-mehrsprachig einfacher und gemischt-mehrsprachig erweiterter Komposition im Vergleich

Abb. 36 zeigt die Kohorten für die Transparenzstufen t3-sprachenspezifisch 2-wertig differenziert und ansonsten die Selektoren aus einer Grundmenge von 3 Sprachen:

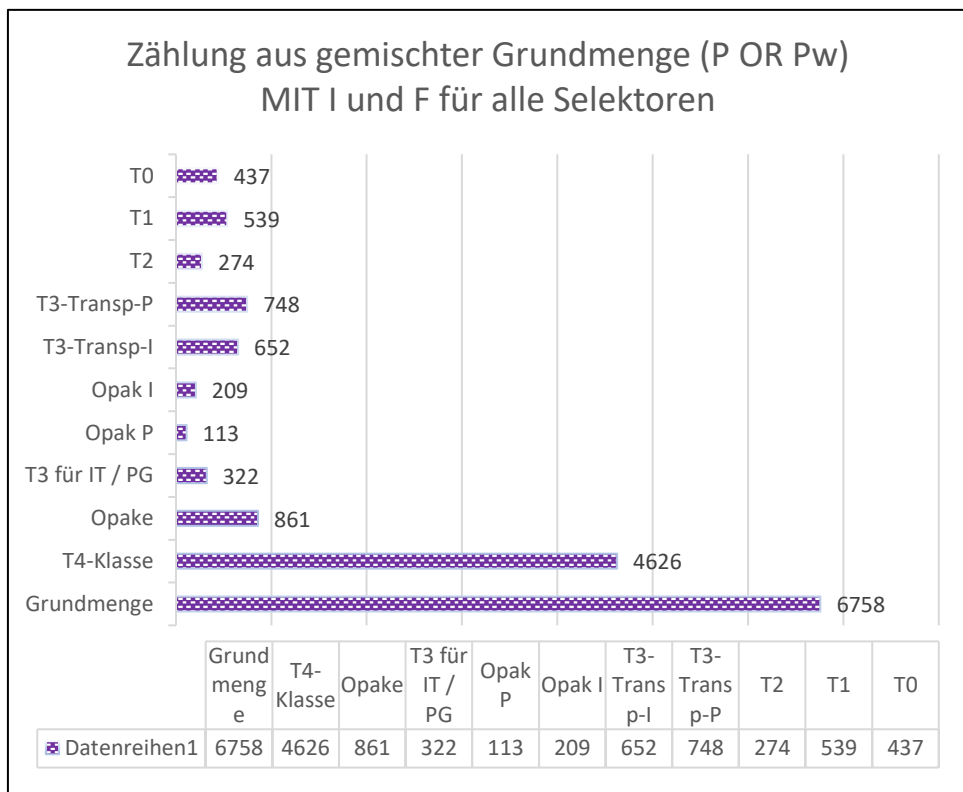


Abbildung 36: Selektoren-Kohorten aus 3sprachiger Grundmenge (P OR Pw) MIT (I) MIT F)

Die Verteilung der RoS-Selektoren (was diese in der Datenziehung erfassen) und der in den Sprachfeldern erfassten macht folgende Tabelle sichtbar:

Vermessung des romanischen Kernwortschatzes

Lfr ▾	Exp ▾	RoS- ▾	Lpt ▾	Lit ▾	Lsp ▾
F	S	RoSE	P	I	S
F	S	RoS	P	I	S
F	S	oFk	P	I	S
F	S	RoSE	P	I	S
F	S	RoS3	Pw	I	S
F	S	RoS3	P	I	S
F	S	RoS-I	Pw	I	S
F	Sw	RoSED	P	I	Sw
F	S	RoSED	P	I	S
F	S	RoSED	P	I	S
F	S	RoSED	P	I	S
F	S	RoSE	P	I	S
F	S	RoSED	P	I	S
F	S	RoSED	P	I	S
F	Sw	RoSE	P	I	Sw
F	S	RoSE	P	I	S

Wie die Tabelle zeigt, liegt die Selektion im Feld Lpt auf (P OR Pw), in Lit auf (I), in Lfr auf (F). Das nicht selektiv-spezifizierte Feld Lsp umfasst sowohl S- als auch Sw-Fälle.

Der Vergleich einer Ziehung aus einer gemischt-mehrsprachig erweiterten Grundmenge zu drei Sprachen mit einer gemischt-mehrsprachig einfachen Selektion (zu drei Sprachen) bietet Abb. 37. Erwartungsgemäß generiert die erweiterte Komposition größere Mengen in den Selektionen.

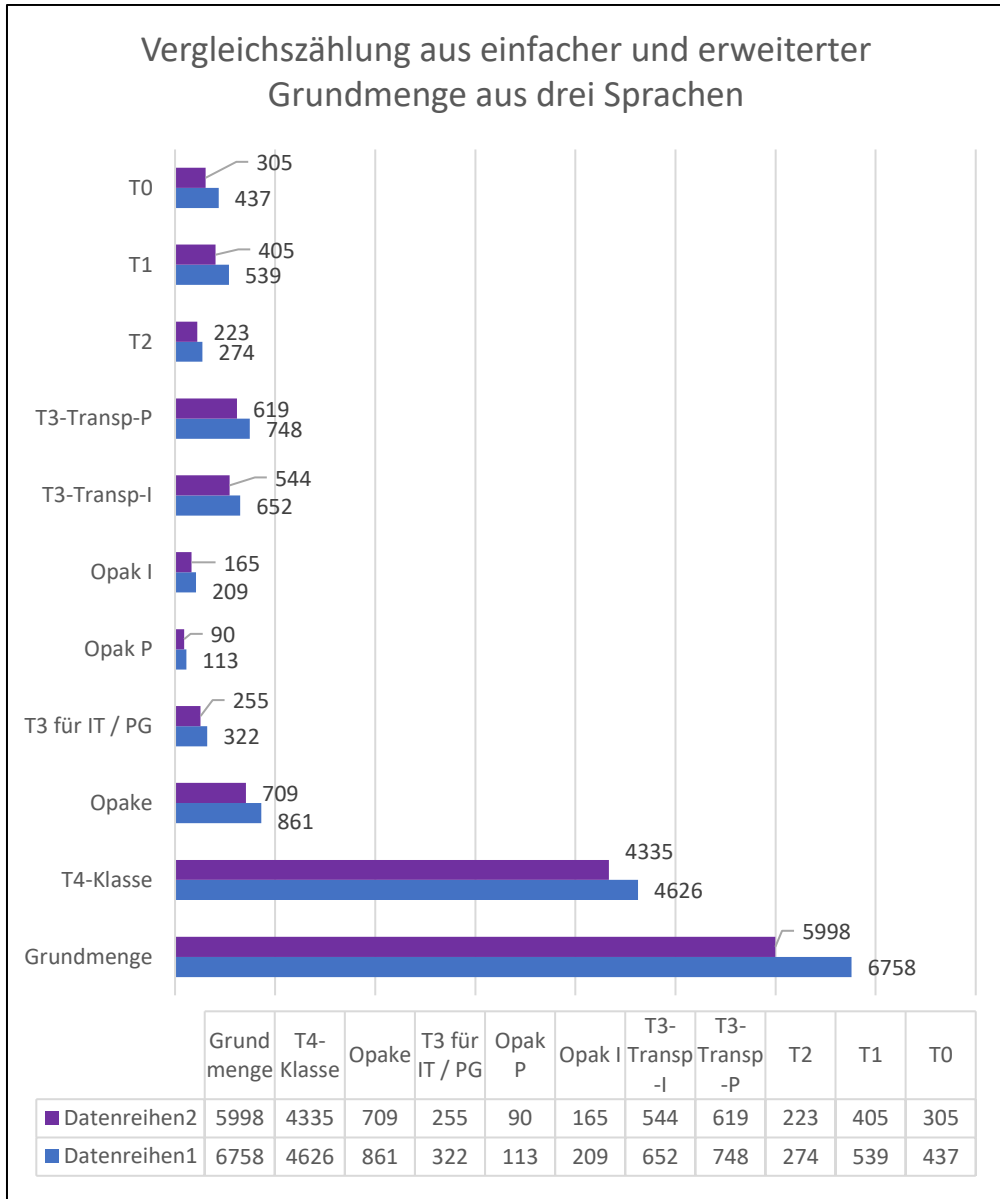


Abbildung 37: Vergleich aus dreisprachigen Grundmengen, gemischt-mehrsprachig, einfach und erweitert

Datenreihe 1 ist aus der Selektion (P OR Pw) MIT (I) MIT (F) (hellblau) komponiert, Datenreihe 2 aus (P) MIT (I) MIT (F) (dunkelblau).

Im folgenden Abschnitt werden die Transparenzprofile in den Blick genommen.

10.3 Transparenzprofile von intralingualen Formdivergenzen in monolingualen und gemischt-bi- oder trilingualen Grundmengen

Die Syntax für die Konstruktion gemischt-sprachlicher Grundmengen soll die Schnittmengen der miteinander verglichenen zwei oder drei Sprachen in Bezug auf die verschiedenen Transparenz- oder Frequenzstufen identifizieren. Dazu benutzt sie die Sprachselektoren (X für die primäre Zielsprache) und (Y für die sekundäre, Z für eine mögliche dritte) in ihren jeweiligen Feldern – also in FR, IT, PG oder SP – gepaart mit der Kombination aus den Selektoren ‚F‘, ‚I‘, ‚P‘ oder ‚S‘. Sind alle Bedingungen erfüllt, heißt dies natürlich für die Zählung, dass nur solche Serien erfasst werden, für die sich in allen betroffenen Feldern ein ‚Gegeben‘, und zwar in den entsprechenden Sprachrubriken als auch der Rubrik der abhängigen Variablen (z.B. RoS in ROS-) – feststellen lässt. Intralinguale Reduplikationen sind durch die Sprachvariable (X) (ohne Xw!), Y (ohne Yw) usw. ausgeschlossen. Anders gesagt: Die Zählung erfasst in dieser Grundmengenlegung keine polysemisch bedingten Dubletten, weder in der primären noch der sekundären Zielsprache.

Wären die Grundmengen aus den unabhängigen (Sprach)variablen gleich, so ergäben die Selektoren seriell für alle (formkongruenten) T4-Lemmata und alle (formdivergenten) T0-Lemmata aller Zielsprachen identische Werte. Doch sind die Grundmengen aus einfach X und einfach Y für die einzelnen Sprachen prinzipiell unterschiedlich stark: Abb. 34 und 35 zeigten durch hohe Übereinstimmungsraten, wie sehr der Variableneffekt (im Modell: [RoS*]) vom Wert der aus den einzelnen Sprachrubriken gebildeten Grundmengen abhängt. In welchem Umfang dies für monolinguale und für sprachlich gemischte Grundmengen gilt, veranschaulichten Abb. 38 und 39:

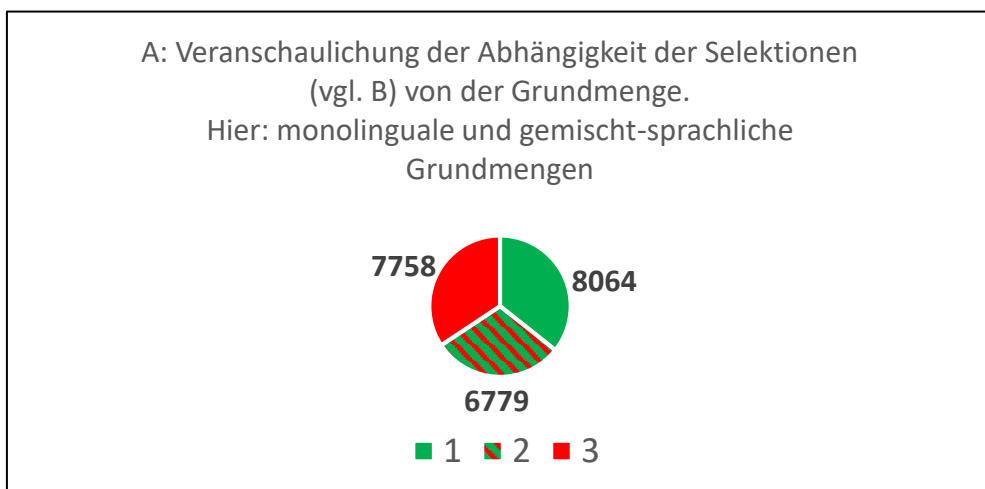


Abbildung 38: A: Abhängigkeit von Variablen

Die aus beiden Grundmengen für Italienisch und Portugiesisch komponierte gemeinsame Menge ist kleiner als die der einzelnen Sprachen.

IT-Grundmenge / PG-Grundmenge > IT MIT PG-Grundmenge.

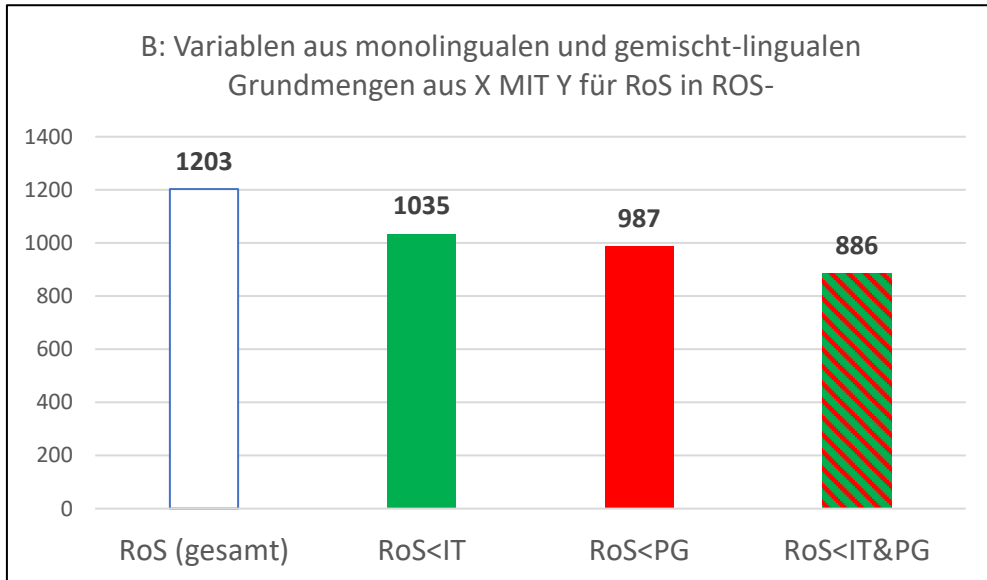


Abbildung 39: B: Variablen und Größenverhältnisse (gepr. 01.05.18)

Ergebnis: IT-Markierungen > PG-Markierungen > IT MIT PG-Markierungen – ein Fazit, das auch andere Selektionen bestätigen: Je größer die Werte der betroffenen Sprachrubriken sind desto größer fällt die entsprechende Variablen-Selektion aus. Dies gilt erwartungsgemäß sowohl für monolingual als auch für bilingual oder trilingual sortierte Ziehungen.

Die folgende Ziehung betrifft sowohl einzelzielsprachliche als auch zwei- und dreisprachliche Mengen

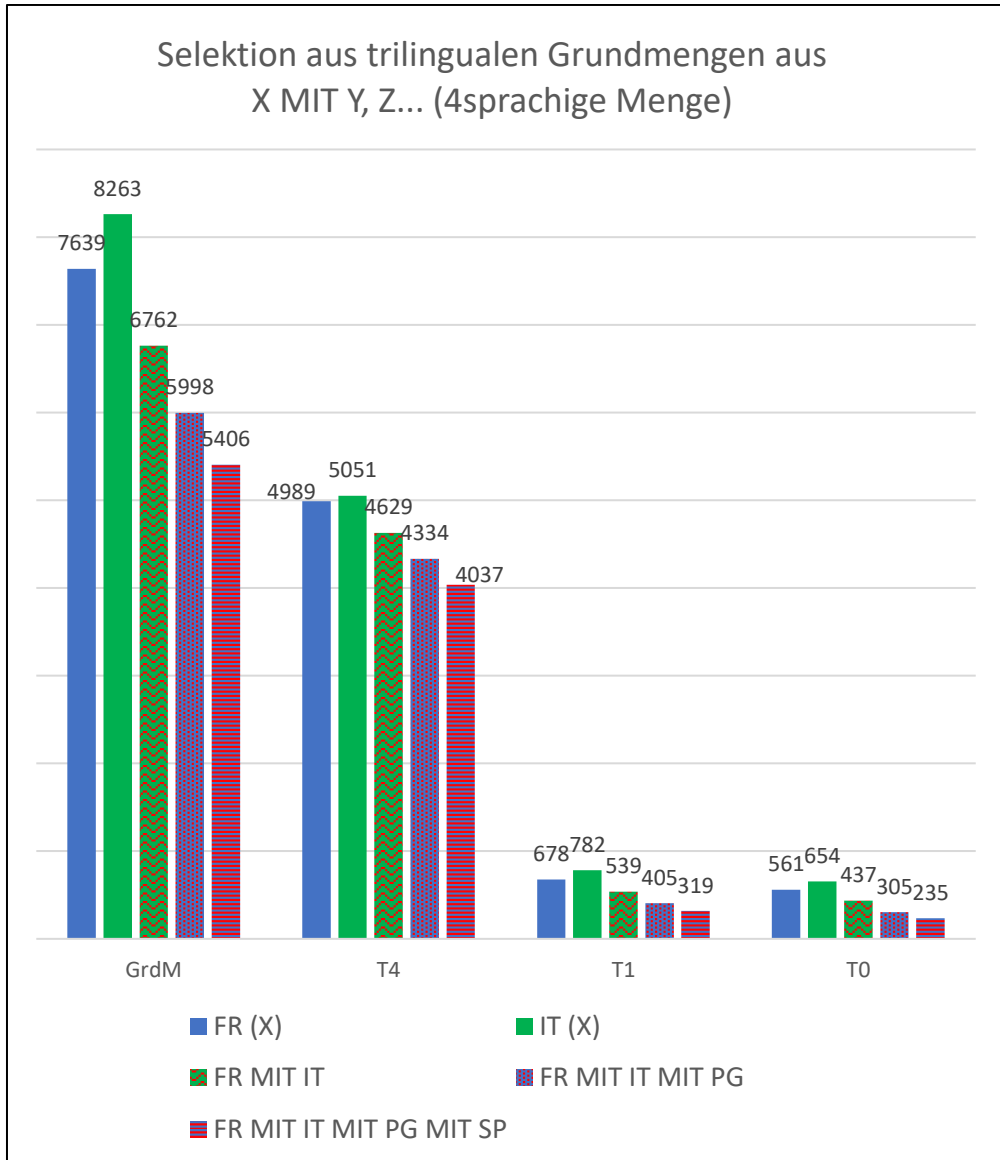


Abbildung 40: Gezielte Selektion aus mono-, gemischt-bi- und -trilingualen Grundmengen

Zur Illustration stellt Abb. 40 drei Transparenzintervalle nach mono-, bi- und trilingual-gemischten Grundmengen nebeneinander. Grob spiegeln die Ergebnisse, wie gesagt, zum einen die Bedeutung der Grundmengengrößen wider. Dies zeigt sowohl der jeweils höhere Wert für Italienisch als auch die ansonsten konstante Deszendenz der anderen Sprachkohorten. Zum anderen bestimmt natürlich auch der Selektor das Ergebnis. Dies zeigt die folgende Graphik deutlicher:

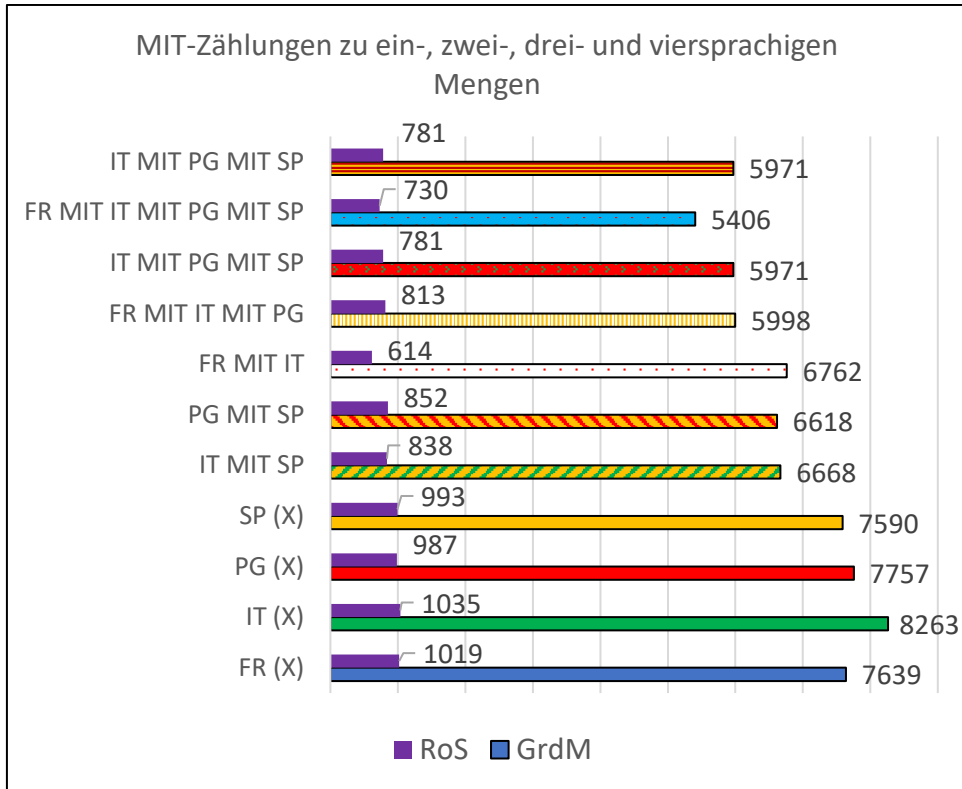


Abbildung 41: RoS-Selektion aus ein- bis viersprachigen Grundmengen

Eine weitere Überprüfung leistet der Selektor (RoS) in ROS-. Auch diese Verteilung dokumentiert die Abhängigkeit der selektierten Kohorte von den Grundmengen, seien sie nun ein- oder viersprachig.

11. Darstellung und Interpretation der gesamten Daten aus bilingualen MIT-Ziehungen

Abb. 42 stellt die lingualen und plurilingualen Profile der romanischen Sprachen für die abhängigen Selektoren-Variablen in den jeweiligen sprachlichen Feldern vor. Alle mehrsprachlichen Zählungen erfolgen allein für (X) und ohne (Xw) für primäre bzw. (Y) ohne (Yw) für sekundäre Zielsprachen. Sie betreffen also nur die intralingual-formdivergenten Lemmata. Abb. 42 präsentiert den vollständigen intralingual-formdivergenten Bestand des zielsprachlichen Inventars aus dieser Grundmengenlegung.

(Die Abbildung umfasst Formdivergenzen in [Kombination](#) mit den romanischen Schwestersprachen bzw. den jeweiligen sekundären Zielsprachen.)

Der Einfachheit halber wird dieser unter dem schon bekannten Begriff der MIT-Gruppe zusammengefasst. (Eine MIT-Grundmenge kann z.B. die Kombination IT MIT PG sein.) Ihre Komposition umfasst mindestens zwei Sprachselektoren.

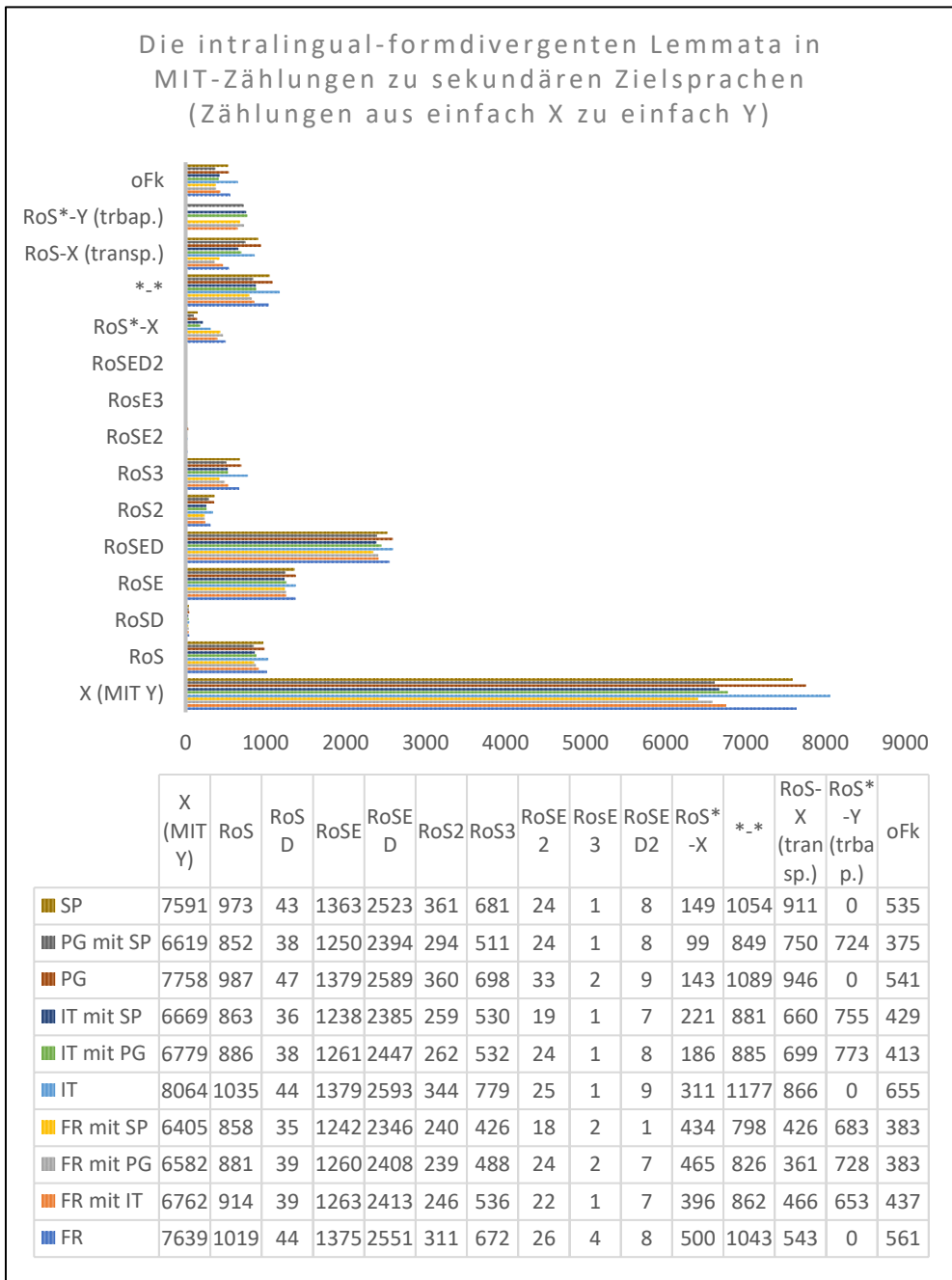


Abbildung 42: MIT-Zählungen (12.12.2017)

Zur leichteren Lesbarkeit von Abb. 42 sei ein Beispiel aus der Reihe IT mit PG erklärt:

Ausgangsmenge ist die Zahl der Serien aus X (IT: 8064): Davon beträgt der IT MIT PG-Anteil 6779. (Der Anteil aller Serien aus I OR Iw wäre 9550; der aus IT MIT PG 9544 [P OR Pw].) Die Zahl der italienischen Formdivergenten in Kombination mit Portugiesisch beträgt, wie gesagt, 6779 Serien, wovon 886 der RoS-Zuordnung, 38 der RoSD-Klasse, 1261 der RoSE- und 2447 der RoSED-Kohorte angehören. Deren Summenwert (4632) umfasst die gesamte T4-Klasse (100prozentige Transparenz). Die T3-Kohorte wird aus der Subtraktion der (*-*)-Fälle (885) abzüglich der (RoS*-I) (186) oder der (RoS*-P)-Fälle (89) ermittelt. D.h. die Zahl der T3-Transparenzen in der MIT-Kombination Italienisch-Portugiesisch beträgt 699 bzw. 773 für den portugiesischen Anteil der Profilwörter. In diesen Serien sind drei zielsprachliche Lemmata füreinander transparent (*enterrer-seppellire-enterrar-enterrar*). Die RoS2-Kohorte (*glisser-glissare-deslizar-deslizar*) {RoS2 OR RoSE2 OR RoSED2} und die RoS*3-Klasse (*bruit-rumore-ruído-ruído*) umfassen die Transparenzstufen T2 und T1 {RoS3 OR RoSD3 OR RoSE3 OR RoSED3}. In der IT-MIT-PG-Kombination beträgt die Gruppe der T0-Transparenzen bzw. Opaken 413 Serien.

Die folgende Graphik präsentiert die Selektorenwerte aus Abb. 42 in Transparenzstufen und Fällen. Der besseren Lesbarkeit halber werden die Werte der RoS-Reihe (RoS, RoSD, RoSE, RoSED) als T4-Transparenz zusammenfasst³⁸.

³⁸ Der Selektor RoS ist streng von RoS* zu unterscheiden! RoS wirft 1205 für die gesamte Datenbank aus, RoS* 8674! „RoS*“ impliziert alle RoS-Erweiterungen, so RoSE, ROSED, RoS-X usw.

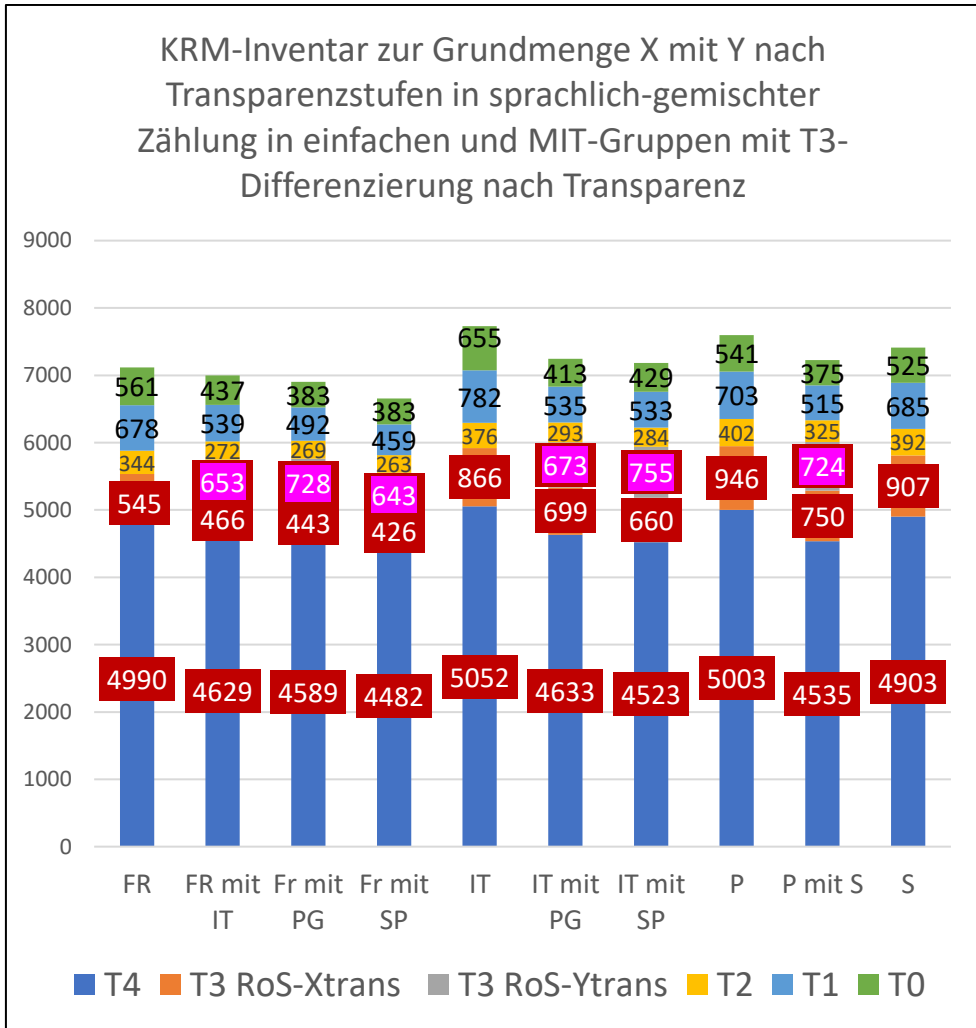


Abbildung 43: Transparenzstufen nach Fällen zu Abb. 39

Legende: In der Excel-Tabelle wurden die T3-Y-Transparenzwerte in Zählungen vom Typ (FR MIT RoS*-F; IT für RoS*-I) = 0 gesetzt. Sie sind hier aus der Graphik genommen. Der T3-X-Wert bezeichnet in MIT-Mengen die in der Kombination zuerst genannte sekundäre Zielsprache, z.B. beträgt die T3-Rate für Französisch in der GrdM FR MIT IT 466 (aus [RoS*-*] (-) [RoS*-F]= 862-396). Die lilafarbig unterlegten Werte zeigen diejenigen der zuletzt genannten Sprache an, d.i. für die MIT-Gruppe FR mit IT natürlich IT (653).

Die Auswahl der MIT-Gruppen folgt dem Schema: F MIT I, F MIT P, F MIT S; I MIT P, I MIT S; P MIT S. Aufgrund der Bidirektionalität der Transfermengen in X zu Y-Relationen ist die korrelierte Auswahl zulässig.

Die Ergebnisse wären in Grundmengen des Typs (X OR Xw) natürlich anders.

Die höchsten Werte erreichen die sprachlich-nicht oder weniger gemischten Selektionen (also einfach FR, IT, PG, SP), und zwar in allen T-Stufen. Der Grund besteht darin, dass die spezifische Selektion, z.B. (RoSED) im Feld ROS- oder PROFIL, für alle aktivierten Sprachen erfüllt werden müssen.

Zur besseren Vergleichbarkeit zeigt die folgende Tabelle die Werte zu Abbildung 39 in Prozent: Nimmt man die sog. MIT-Kolonnen für sich, so fällt wiederum der übergroße Anteil auf, den die T4-Selektion einnimmt: zwischen 58 und 68 Prozent des Gesamtanteils. Auch hier stellen die FR MIT IT- und die PG MIT SP-gemischten Selektionen die stärksten Kohorten. Dem steht ein seriell opaker Teil (T0) von nur 5 bis 6 Prozent gegenüber (für T3 bis zu 10%). Ebenfalls bis zu 8% stark ist unter den sprachlich-gemischten Komputationen der Anteil der Serien, die immerhin eine Transferbasis für eine weitere romanische Sprache aufweisen (T1).

Für die Nutzung des synergetischen Potentials zur Förderung mehrsprachig-romanischer Lesekompetenz und ihren Voraussetzungen an lexikalischen Ressourcen meint dies, dass mit der bloßen Kenntnis eines T4- und T3-Anteils in nur einer einzigen romanischen Sprache bereits 70 bis nahezu 80 Prozent des kernwortschatzlichen Wortschatzes einer weiteren romanischen Sprache erschließbar bzw. Lernern augenscheinlich transparent werden können.

Eine starke Beschleunigung des Ausbaus der interkomprehensiven Lesekompetenz ist vor allen dann in kurzer Zeit realisierbar, wenn schon zu einem möglichst frühen Zeitpunkt der Lernstrecke die opaken Wortschatzanteile vermittelt/erworben werden. Dies betrifft nur 5% bis 7% der Lemmata der jeweiligen Selektion. Bei ca. 7000 formdivergenten Lemmata pro Zielsprache sind dies 350 bis 500 Lexeme³⁹.

³⁹ In X OR Xw-Darstellungen begegnen durchaus seriell vierfach formdivergente (opake) Lemmata, von denen die eine oder andere Form Lernern bekannt sein dürfte: sog. Strukturwörter (im Französischen ist der Terminus der *mots structurels* geläufig), im Deutschen etwa *der, ein, unter, über, von* usw. Solche Wörter stellen den Kern der hochfrequenten Lexis dar.

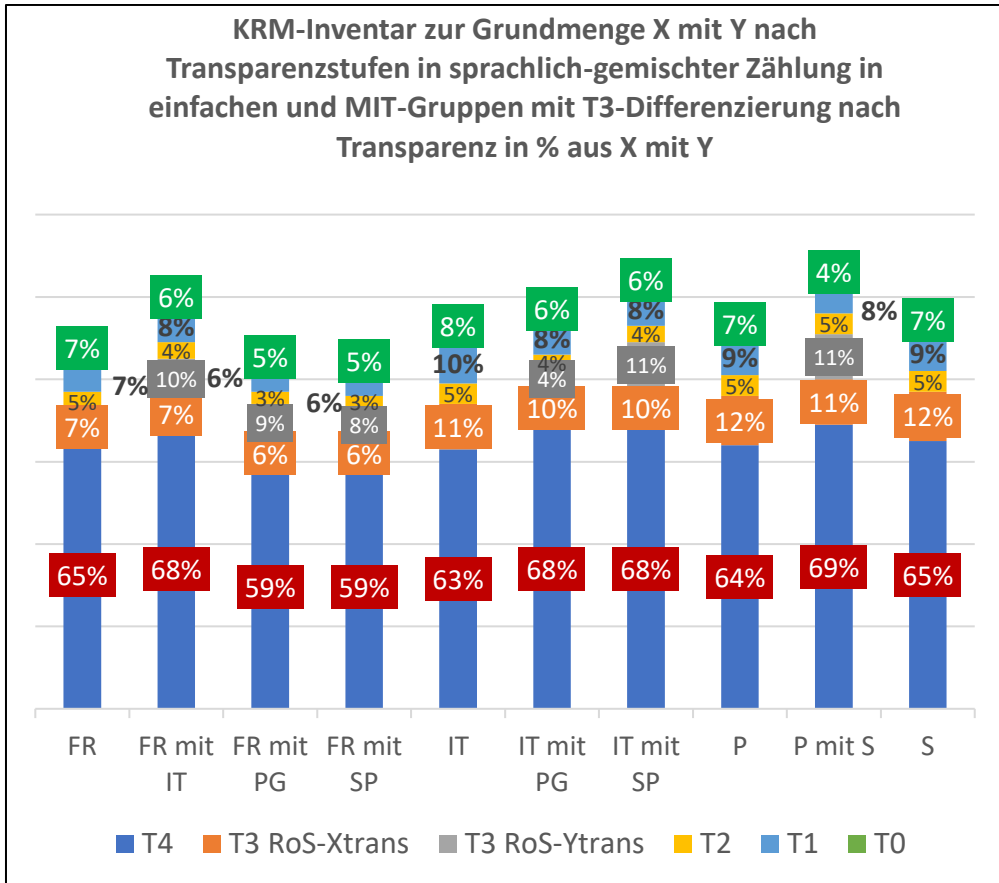


Abbildung 44: die intralingual-divergenten Lemmata in sprachlich-gemischter Kombination in %

Im zwischensprachlichen Vergleich der Transparenzanteile fallen die Unterschiede von Sprache zu Sprache, MIT-Kohorte zu MIT-Kohorte geringer aus als vielleicht erwartet. Nimmt man das Kriterium einer „deutlich überwiegenden Transparenz“, gebildet aus den Anteilen von T4 und T3 (Transparenz für X), so ergeben sich folgende Transparenzraten für sprachlich gemischt gezählte Populationen:

Sprachen-Kombination	Prozent
FR mit IT	75% bzw. 85% ⁴⁰
FR mit PG	65% bzw. 68%
FR mit SP	65% bzw. 73%
IT mit PG	72% bzw. 78%
IT mit SP	78% bzw. 79%
PG mit SP	80% bzw. 80%

40: Transparenzanteil aus T4/T3 in sprachlich-gemischter Selektion

Die unterschiedlichen, durch „bzw.“ getrennten, Werte ergeben sich aus den unterschiedlichen T3-Spezifikationen. Diese erscheinen in Tabelle 40 summiert. Für FR MIT IT lautet also die Rechnung: 68+7+10.

Die Prozentwerte beziffern den Umfang des potentiell hochgradig, nahezu auf den ersten Blick, interlingual transferablen Vokabulars.

11.1 Vermessung morphologischer Kongruenz- und Divergenz-Relationen in gemischt-sprachlichen Serien

Gemischt-sprachlich kombinierte Grundmengen zählen, wie gesagt, aus mehreren – mindestens zwei – Sprachrubriken. Sie korrespondieren damit tendenziell mit der Mehrsprachigkeit der mentalen Lexika von Lernern.

Allerdings sind diese wie gesagt weitaus umfangreicher als die Kernwortschätze:

Ein gebildeter Erwachsener mit Englisch als Muttersprache versteht und verfügt potentiell über wenigstens 50000 Wörter, wobei ein Wort vorläufig als »Wörterbucheintrag« definiert wird. (...) (Aitchison 1997: 8)

Im Hinblick auf die Lesekompetenz ist folgender Hinweis von Interesse:

(Nach Seashore & Eckerson, 1940) kennt ein gebildeter Erwachsener wahrscheinlich mehr als 150000 Wörter und ist in der Lage, 90 Prozent davon aktiv zu nutzen. Die Zahl ist umstritten, weil die Definition von »Wort« problematisch ist. (Aitchison 1997: 6)

Abgesehen von den umstrittenen hohen Zahlen dürfte sicher sein, dass nicht zuletzt der internationale Wortschatz in ganz erheblichem Maße zu diesen beiträgt.

⁴⁰ Der zweitgenannte Wert ist der Ytrans-Wert für die T3-Gruppe, also 8% für FR MIT SP.

Doch zurück zum KRM-Wortschatz und seiner Vermessung.

Im Unterschied Abb. 42 zur Quantifizierung nach Selektoren stellt Abb. 45 die T-Klassen aus zweisprachigen MIT-Zählungen im Überblick dar, und zwar nach Fällen:

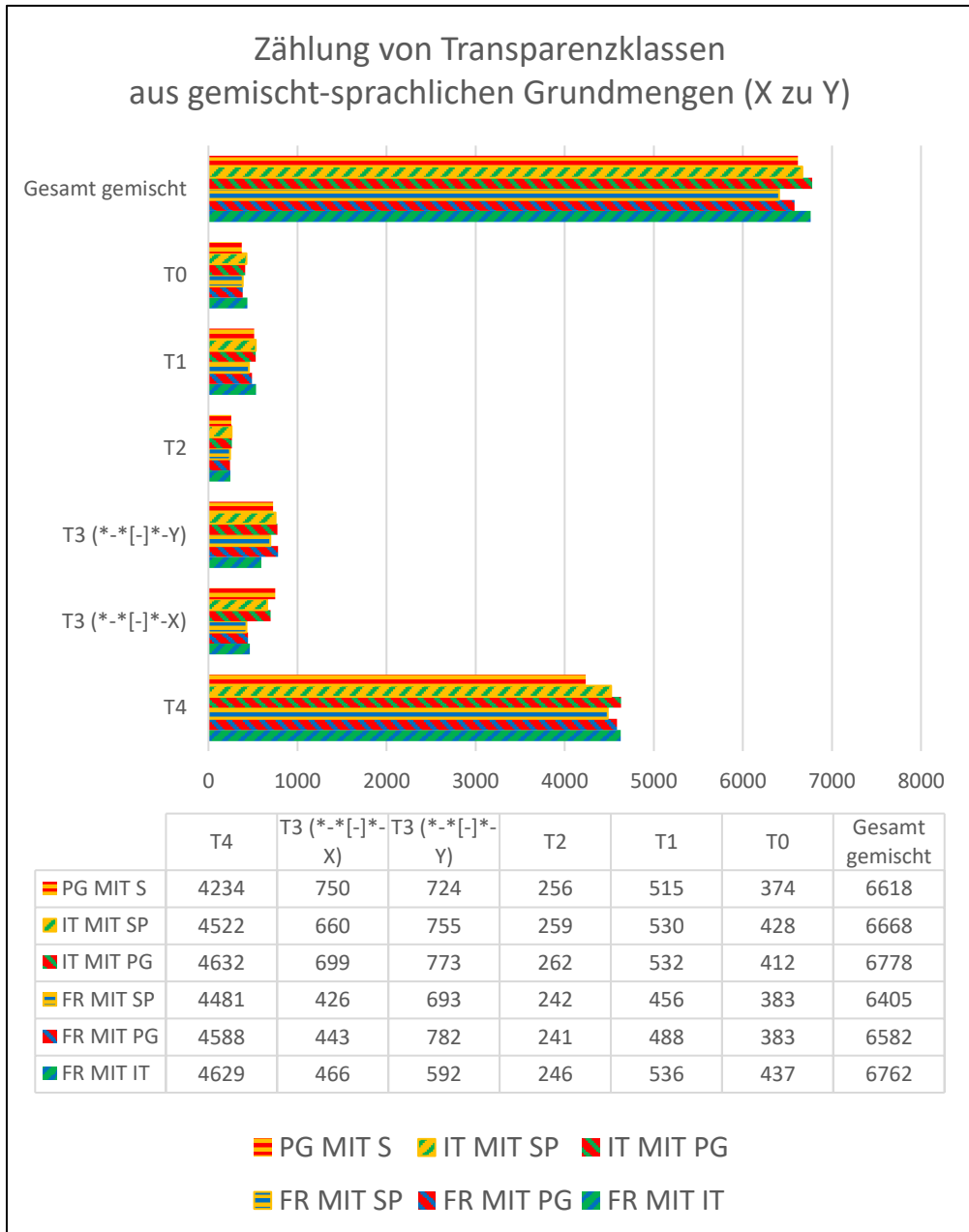


Abbildung 45: Transparenzklassen in gemischt-sprachlichen Grundmengen aus X zu Y

Eine Veränderung der Grundmengen zugunsten der Selektion (X OR Xw MIT Y) dokumentiert Abb. 46. Aus Gründen der Übersichtlichkeit wird diesmal auf die Angabe des T3-Anteils verzichtet:

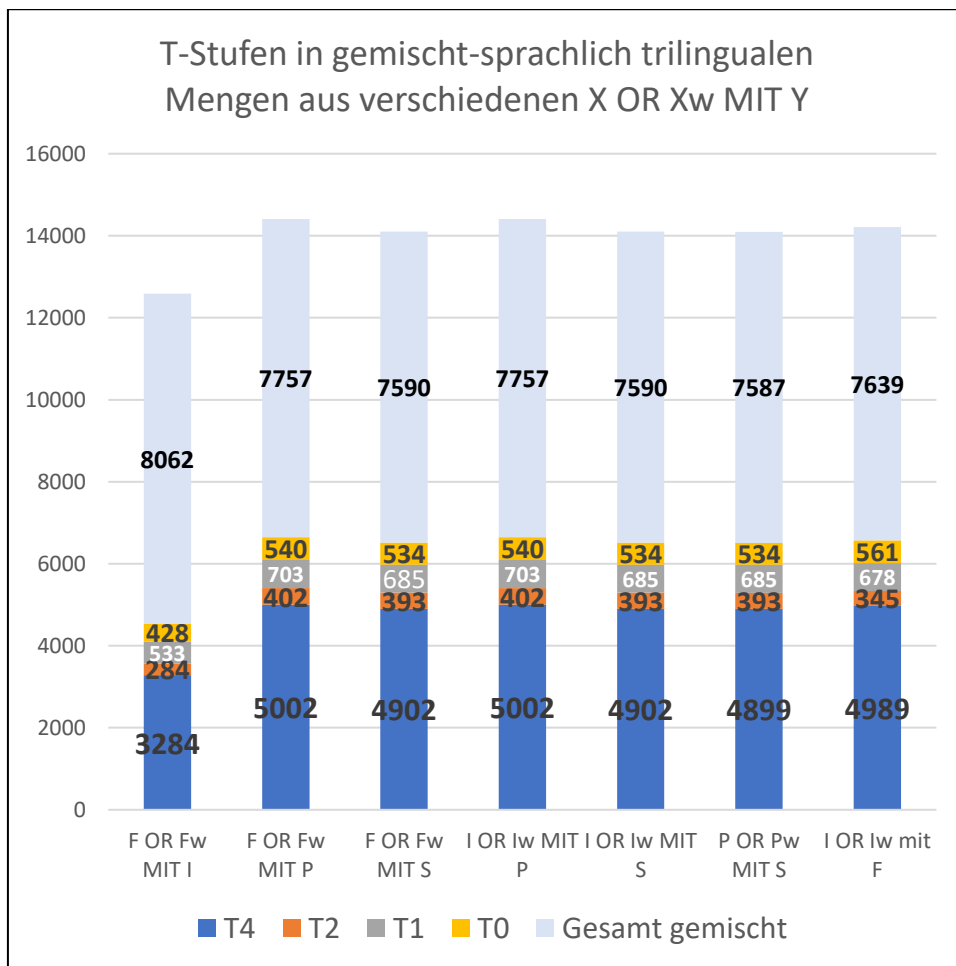


Abbildung 46: Transparenzraten Klassen aus trilingual gemischt-sprachlicher Zählung ohne T3

Die obere Reihe kennzeichnet die Grundmenge (aus z.B. F OR Fw MIT I), die untere die Transparenzklasse. Da die T3-Werte nicht eingesetzt wird, erscheint der Transparenzwert aus exklusiv T4 im Vergleich zu den anderen Abbildungen als zu gering. Würde man die betroffenen T3 aus Abb. 44 addieren und durch ihre Anzahl teilen, so ergäbe sich ein Durchschnittswert von abgerundet 19%. Diese Rechnung ist natürlich aufgrund der unterschiedlichen Grundmengen und der offenen Frage der Anzahl der RoS*-X-Zählungen nicht statthaft! Ansonsten sprechen die Werte in Abb. 46 für sich.

Grundsätzlich macht es in einer Gesamtmenge aus 4 Zielsprachen wenig Sinn Transferbasen zu zählen, wenn alle 4 Zielsprachen als bekannt– d.h. als primäre Zielsprachen bekannt – vorausgesetzt werden. Auch deshalb wurde zu Abb. 46 auf die Komputation der T3-Werte verzichtet. Es genügte ein sehr grober Überblick.,

11.2 Über die Romania hinaus zum europäischen Wortschatz⁴¹

Es wurde darauf hingewiesen, dass Wörter (nicht nur) mit kernwortschatzlicher Präsenz im romanischen, englischen und deutschen Wortschatz mit hoher Wahrscheinlichkeit auch in weiteren Sprachen etabliert sind. Hierzu ein Beispiel zu einer kleinen, willkürlich getroffenen Auswahl von RoSED-Wörtern⁴². Es wurde überprüft, ob sich auf Anhieb eine formkongruente Bedeutungsadäquanz in solchen Sprachen finden lässt, die nicht im Fokus des KRM liegen. Ausgangspunkt waren KRM-Eintragen von Deutsch oder Englisch (RoSE und RoSED). Eine Vorauswahl

⁴¹ Exkurs zur lexikodidaktischen Interpretation: Wirft man einen Blick auf die vertikalen Rubriken von [Abb. 12](#), so fällt auf, dass – auch in der Kombination der romanischen Sprachen der mit dem Englischen und – weniger – dem Deutschen (ohne Englisch) geteilte Kernwortschatz den bei Weitem größten Anteil der Transparenz (T4) stellt (RoSE und RoSED zzgl. des RoSE2 und RoSED2 bzw. des RoSED3 und RoSED3-Anteils). Die lexikologischen Gründe hierzu wurden bereits genannt (vgl. den dem Englischen gewidmeten Abschnitt): Seine besondere internationale Stellung erklärt sich natürlich mit der äußeren Geschichte dieser Sprache bzw. der sich ihrer bedienenden historischen wie heutigen Gesellschaften, Staatswesen und Bevölkerungen: die konstante Integration Britanniens in den lateinisch funktionierenden Verwaltungs- und Militärapparat des Imperium Romanum und dem internationalen Handel in der Sprache Roms; der Sprache der Kirche, der vollen Teilhabe der Insel an der abendländischen, mittelalterlichen und neuzeitlichen *res publica litterarum*, der Wirkungen der Erfindungen und Erfinder (vom *Habeas Corpus Act*, dem Art. 39 der *Magna Carta Libertatum*, der die Rechte des Einzelnen gegenüber dem Staat sichert, John Hobbes, Isaac Newton und, und, und... aber auch die amerikanischen *Federalists*). Zu erwähnen ist selbstredend auch die Entwicklung des *British Empire*, der USA und weiterer anglophoner Zentren mit ihren kaum abschätzbaren und vielfältigen Folgen für die globale Verbreitung des Englischen. Kurz und gut: Der englische Sprachraum hat das kulturelle Gesicht (nicht nur) Europas wesentlich mitgeprägt. In englischem Gewand wurden viele Wörter an die europäischen Sprachen und über diese hinaus weitergereicht. Neue Entwicklungen erforderten neue Nomenklaturen. Natürlich lassen sich diese und andere Phänomene in vergleichbarem Umfang mit Fug und Recht auch von mancher anderen europäischen Sprache sagen, und natürlich greifen diese knappen Bemerkungen viel zu kurz.

Bereits erwähnt wurden die lexikologischen Kategorien, die sich auch mit dem Anteil des Englischen am europäischen Kernwortschatz verbinden: Kultismen, *migrazione dotta*, Szientismen, Wörter der modernen Zeit, neolateinische Kompositionsweisen, der wissenschaftliche Fachwortschatz, z.B. der Medizin (Beispiel: *Mikrowellen-Metastasenablation*, *ablation de métastases par micro-ondes*, *microwave ablation of metastases*) usw.

⁴² Einen umfassenden Überblick bietet das Standardwerk von Schulz & Basler (1973ff).

wurde selbstverständlich nicht getroffen^{43/44}. Natürlich erhebt die erstellte Liste keinerlei weiteren Anspruch als die unvollständige augenscheinliche Überprüfung der internationalen Verbreitung des europäischen Wortschatzes. Die Bescheidenheit des Anspruchs erklärt, weshalb Bemerkungen zur lexikographischen Praxis der benutzten Wörterbücher unterbleiben. Es wäre indes leicht möglich, die impliziten ‚Behauptungen‘ dieser Aussage durch einschlägige lexikologische Studien zu erhärten (vgl. Meißner 1993b; Munske & Kirkness 1996; Braun et al. 2003). Das Verfahren erscheint gangbar, weil die Liste auch in der Unterrichtspraxis verwendbar sein soll.

Wie Tabelle 41 zeigt, wurde zwischen europäischen und exotischen Sprachen unterschieden.

<i>KRM-Lemma</i>	<i>Bedeutungsadäquanzen mit identischen morphologischen Elementen in anderen europäischen Sprachen als den romanischen, Deutsch und Englisch (zufällig erstellte Liste)</i>	<i>Bedeutungsadäquanzen mit identischen morphologischen Elementen in exotischen nicht-europäischen Sprachen</i>
<i>accent</i>	poln. akcent	-/-
<i>attività</i>	poln. aktywność	-/-
<i>appareil</i>	poln. aparatura, aparat	-/-
<i>ônibus</i>	(автóбус)	-/-
<i>blusa</i>	poln. bluzka	-/-
<i>broche</i>	Брóшка, poln. broszka	-/-
<i>civilización</i>	russ. цивилизáция, poln. cywilizacja	-/-
<i>consommer</i>	slovak. konzumovať; schwed. (att) konsumera	-/-
<i>discusión</i>	poln. dyskusja	-/-
<i>formare</i>	schwed. att forma; poln. formować	-/-
<i>lancer</i>	poln. lansować	-/-
<i>passagem</i>	russ. пассáж, poln./slowak. <u>p</u> asas	-/-
<i>projet</i>	poln. projekt, проéкт	-/-
<i>religion</i>	russ. религия,	-/-

⁴³ Gesucht wurde in den Beständen verschiedener Online-Wörterbücher (<https://defi.dict.cc/>; <https://dict.leo.org/englisch-deutsch/>; <http://www.angkor-planet.com/dico>; http://www.lexilogos.com/turc_dictionnaire.htm u.a.m.).

⁴⁴ Das Ergebnis wäre anders ausgefallen, wenn man ausgehend von den Entlehnungen europäischer Wörter durch z.B. asiatische Sprachen gesucht hätte (vgl. Werner 2003; insbes. S. 193ff). Offensichtlich bleiben Wörter, die in der eigenen Tradition des Landes lebten, in den Kernwortschatzen präsent, so etwa die Entsprechungen von *Polizei*, *Grammatik*, *Mathematik*, *Genie*, *Enthusiasmus* usw. im *Chinesischen*. Eine jüngere Entwicklung betrifft die starke Beeinflussung durch das Englische im Rahmen eines sog. westlichen Lebensstils.

<i>spettacolo</i>	poln. spektakl	-/-
<i>Zucker</i>	poln. cukier, сахар	jap. やっつける ([yattsukeru])
<i>Theater</i>	russ. театр, poln. teatr, schwed. teater	-/-
<i>extremely</i>	russ. Экстремально, schwed. extremt	-/-
<i>risco</i>	poln. ryzyko, russ. риск	-/-
<i>relation</i>	Poln. relacji	-/-
<i>estilo</i>	russ. стиль, poln. styl	-/-
<i>to adapt</i>	schwed. att adaptera, russ. адаптировать	-/-
<i>constitution</i>	russ. Конституция, poln. konstytucja	-/-
<i>hypothèse</i>	bulg. russ. гипотеза	-/-
<i>Zigarette</i>	russ. сигарета, türk. sigaret	-/-
<i>genius</i>	ungar. zseni, norw. geni, russ. гений	

41: Formkongruenzen über die Romania hinaus... (Beispiele)

Was signalisiert Tabelle 41? Sie rechtfertigt die Hypothese, dass mindestens 46% des romanischen Kernwortschatzes Formkongruenz mit europäischen Sprachen teilt, die nicht Teil des KRM-Korpus sind⁴⁵. Diese Transferbasen bilden eine zentrale lexikalische Grundlagenkomponente für die Erzeugung von Synergien einer umfassenden europäischen, weit über die Romania hinausreichenden Mehrsprachigkeit, bestehend aus zumindest jeweils einer germanischen, einer romanischen und einer slawischen Sprache neben und mit Englisch. Eine solche Annahme unterstützt die Grundthese des EuroCom-Modells (zu den Grenzen: Berthele et al. 2011). Nimmt man die romanischen Sprachen als Zielsprachen, so unterstreicht in umgekehrter Richtung die Bidirektionalität potentieller Transferrelationen die ‚leichte‘ Erlernbarkeit eines Großteils des KRM-Wortschatzes z.B. durch deutsch-, tschechisch-, englischsprachige Lerner, wobei die romanophonen weitere Vorteile verbuchen können⁴⁶.

⁴⁵ Der Wert ergibt sich aus den 4348 RoSE- oder RoSED-Markierungen bezogen auf die Serienzahl von 9558.

⁴⁶ Die in diesem Aufsatz noch nicht verwendete Liste der Interligalexe dokumentiert zudem, wie sehr Wörterlernen auch immer das Erlernen von Wortbildungselementen – lexematische Ligalexe, prä- und suffigale Ligalexe, Kombinationsregeln – bedeutet. Dies alles umreißt u.a., wie sehr Sprachenbewusstheit durch metalexikalisches und einfaches laien-lexikologisches Wissen gefördert werden kann. Dessen wesentliche Komponenten sind in morphologischer Sicht: die Segmentierungsfähigkeit (De- und Rekomposition) von Wörtern, in semantischer Sicht: die Zurückführung von kookkurrentiellen Bedeutungen auf Kernbedeutungen und umgekehrt die Extrapolation von Kernbedeutungen zu konkreten Bedeutungen in Ko-Texten unterschiedlicher Sprachen, in Bezug zur „Wortbewusstheit“: die Öffnung zu historischen Aspekten um Etymologien, Wortgeschichten und mentalen Szenarien in Verbindung mit den Eigenschaften von Wörtern.

12. Ergebnisse und Diskussion

Vorab erfolgt einleitend die Darstellung der Ergebnisse, gemessen an den Zielen der Studie.

12.1 Was zu erbringen war...

Das Ziel der vorliegenden lexikostatistischen Studie bestand in der Erhellung bzw. der detaillierten Quantifizierung des romanischen Kernwortschatzes, komponiert aus vier wichtigen, als Erst-, Zweit- und Fremdsprachen weit verbreiteten romanischen Sprachen. Dem didaktischen Ansatz gemäß lag der Fokus auf der potentiellen Lernökonomie – Stichwort: mögliche Synergien nutzen, um den Weg zur (zunächst?) rezeptiven romanischen Mehrsprachigkeit zu erleichtern. Dies betrifft rein materiell vor allem den Wortschatz, besser: sowohl die einem Lerner oder einer Lernerin schon bekannten Wortschätze als auch das mit einer neuen Sprache zu erwerbende neue Vokabular.

Unbestreitbar ist – solange eine diversifizierte und abgestufte Mehrsprachigkeit ein Lernziel ist (Meißner 2019) – die Analyse des mehrsprachlichen Inputs und der Sprachcurricula von entscheidender pädagogischer Wichtigkeit. Hier liegt die Verbindung des KRM zu den Arbeiten der didaktisch orientierten Lexikographie und ihrer Absicht, den Input nach gebrauchorientierten Kriterien zu erfassen und pädagogisch aufzubereiten. Daher haben hier seit über hundert Jahren sog. Minimalwortschätze für einzelne Sprachen ihren Platz in der Fremdsprachendidaktik. Doch was in puncto Lernökonomie für einzelne Sprachen gilt, gilt erst recht für auf Mehrsprachigkeit zielende Ansätze. Deshalb beschreibt der KRM die Lemmata der vier Zielsprachen nach Frequenzrang-Intervallen, deren Komposition in der Literatur vielfach beschrieben wurden (u.a. Kühn 1990; Zöfgen 1994; 2006; Tschirner 2005; Siepmann 2006; Perkuhn et al. 2012).

Qualitativ liegen der plurilingualen Lernökonomie (und ihrer Nutzung) lexikalische Transferbasen zugrunde, die lernerseitig über einen sprachenübergreifenden Identifikationstransfer aktiviert werden (können). Dies erklärt die Relevanz der ‚Qualifikatoren‘ von interlingualer Transparenz und Opazität. Sprache, erst recht Sprachen, und vor allem die Semantik funktionieren aufgrund einer ‚Verdichtungsleistung‘, in der sich die *langue* im Sinne Saussures in *parole* umformt. Auf diesem Weg wird die lexikographierte Bedeutung eines Wortes zu einer okkurrentiellen konkreten Bedeutung der Rede (andere sprechen von *Meinung*) oder eines bestimmten Ko-textes und Kontextes. Psychologisch gesehen ist dies ein Weg von der Mehrdeutigkeit der Sprache zur Konkretion – (möglichst) der Eindeutigkeit – des Sprechens. Auf der Seite der Rezeption setzt dies korrelierende Operationen voraus:

die Konstruktion und einer sprachlichen Mitteilung im Sinne des vom ‚Sender‘ mitgeteilten Sinns voraus. Gilt dies grundsätzlich für jeden Akt monolingualer Kommunikation, so erst recht, wenn Mehrsprachigkeit beteiligt ist.

Der zweite, bereits angesprochene ‚Qualifikator‘ ist die pädagogische Nützlichkeit, als deren wichtiges Merkmal die Anwesenheit eines Wortes im zentralen Wortschatz einer Sprache gesehen wird. – Beide Qualifikatoren, interlinguale Transparenz und prinzipielle Transferierbarkeit des romanischen Vokabulars einerseits sowie seine Nützlichkeit im Sinne der lexikalischen Frequenzforschung andererseits, wurden in der vorliegenden Arbeit detailliert beleuchtet.

Hierzu mussten lexikostatische Wege und Beschreibungsmodi entwickelt werden, die es erlaubten, die Offenheit des Wortschatzes – in der englischen Literatur oft treffend als *fuzziness* beschrieben – auf greifbare und damit zählbare Parameter zu bringen. Dies verlangte, solche Selektoren zusammenzustellen, die eine zielführende Quantifizierung des mehrsprachlichen KRM-Inventars nach den ins Auge gefassten lexikologischen und pädagogischen Kriterien ermöglichten.

In der KRM-Datenbank leisten dies unterschiedliche Schreibkonventionen und Markierungen (Selektoren).

Zugleich begrenzt der KRM die Reichweite seiner Daten auf den sog. Kernwortschatz (*Core Vocabulary, lexique central oder fondamentale* ⁴⁷), also auf das zentralste Segment des sog. zentralen Wortschatzes ([Abb. 1](#)). Auch wenn die Angabe des Frequenzrangs einer Vokabel für deren Erlernen ziemlich irrelevant ist, so bleibt sie doch für die Auswahl des lexikalischen Inputs ein mitentscheidender Faktor.

Die hier beschriebenen Schritte (Meißner 2016; 2018) waren bereits bei der Entwicklung der Datenbank getan worden. Die vorliegende Komputation stellt implizit eine Überprüfung der Gangbarkeit der konstruierten lexikostatistischen Methode als auch der Zuverlässigkeit der Datenbank dar (die von der des lingualen Inputs zu trennen ist). Insofern trägt die Zählung auch den Charakter eines Audits.

Dies wirft zuvorderst die Frage nach der Zuverlässigkeit der Daten auf. Auch dieser Frage wurde in den genannten Studien nachgegangen. Hier sei nur daran erinnert, dass die Daten das KRM-Inventar zählen, welches auf der einschlägigen IT-basierten lexikalischen Frequenzforschung der Computerlexikographie basiert (Davies 2006: 2ff.; sodann: <https://corpus.byu.edu/>). An der Zuverlässigkeit kann kein Zweifel bestehen.

⁴⁷ Die Bezeichnungen greifen nicht immer eindeutig.

Die Originalität des KRM gegenüber den als „Quellen“ benutzten einzelsprachlichen Wortlisten besteht in der Mehrsprachigkeit der Darstellung, d.i. der Auswahl und Zuordnung der verschiedensprachlichen Lemmata in bedeutungsadäquaten Serien und ihrer didaktischen Kennzeichnung, so dass vielfältige Selektionen innerhalb des Repertoires möglich sind. Die Apps richten sich an diesen möglichen Selektionen aus.

Die Zählungen machen für sich genommen die Originalität des KRM-Korpus sichtbar. Die Selektoren und ihre Anwendungen haben auch eine besondere Bedeutung für die Selektion der Inhalte der Lernapps, indem sie zu Aussagen über Transferabilität, Operabilität und deren Umfang sowie zu Frequenzrängen bzw. Nützlichkeit führen.

Ein weiterer, wichtiger Grund für die vorliegende Analyse (und ihrer Publikation) ergibt sich daraus, dass der Inhalt der Datenbank mit einem Copyright belegt ist und die Daten selbst in toto der Öffentlichkeit nicht vollständig zugänglich sind. Die Studie versucht, diesen für die fachwissenschaftliche Öffentlichkeit zweifellos schweren Nachteil abzumildern, indem sie mögliche lexikologische, lexiko-statistische und lexikodidaktische Fragestellungen unter Heranziehung der KRM-Daten behandelt.

12.2 Didaktische Rahmung

Die quantitativen Analysen der Formkonvergenzen und Formdivergenzen des KRM in Bezug auf interlinguale Transparenzstufen bzw. Transferbasen generierten erwartungsgemäß belastbare Aussagen zu Art und Umfang des Transferpotenzials bestimmter Sprachen, Sprachenpaare und -kombinationen zwischen sog. Brücken- und Zielsprachen. Schon diese Kennzeichnung weist den KRM als ein Instrument des interkomprehensiven bzw. mehrsprachigkeitsdidaktischen Ansatzes aus.

Mit der Identifikation und Frequenzierung potentieller interlingualer Transfermengen innerhalb des KRM-Inventars antworten die Befunde auf das Desiderat einer empirisch begründeten Kenntnis der materiellen Grundlagen der romanischen und z.T. europäischen Mehrsprachigkeit. Die Kenntnis hierüber ist, wie angedeutet, eine Voraussetzung für die Identifikation des Umfangs der interlingual möglichen Lernökonomie bei der pädagogischen Zuwegung zur Mehrsprachigkeit. Dies erklärt, weshalb diese Frage immer wieder im Zusammenhang mit einer optimalen Folge von Schulfremdsprachen aufgegriffen wurde: Seitdem Thorndike & Woodworth (1901) sich des Begriffs *Transfer* bedienten, besitzt die Frage einer optimalen Reihenfolge und Anordnung der Lerngegenstände in der Lernpsychologie Relevanz. Im Kontext des Fremdsprachenlernens betrifft dies die Sprachenfolge.

Nicht nur dies zeigt, weshalb das Projekt des KRM über die Interkomprehensionsdidaktik hinausgreift. Es erfasst durchaus auch die Praxen des heutigen Unterrichts

im Allgemeinen, und zwar (nicht nur) romanischer Sprachen⁴⁸. In diesem Sinne schließen die Ausführungen an eine umfangreiche deutschsprachige und internationale Forschung und Fachliteratur zum interkomprehensionsbasierten Lehren und Lernen fremder Sprachen an. Ein weiterer Grund, weshalb die KRM-Analysen von übergreifendem Interesse sind, folgt aus den Ergebnissen der Mehrspracherwerbsforschung. Studien zum mehrsprachlichen mentalen Lexikon zeigen, dass Wörter (und die weiteren Bestände) unterschiedlicher Sprachen interagieren. Dies äußert sich nicht nur in Interferenzen („negativem Transfer“), deren Bedeutung jahrzehntelang überschätzt wurde, sondern vor allem in positivem Transfer und handfesten Lernerleichterungen (Meißner 1999; Reinfried 1999).

Die Relevanz des Themas ist nicht nur für das sprachen- und fächerübergreifend- vernetzende Lernen augenfällig, wie es z.B. im Konstrukt des Gesamtsprachen- curriculums entgegentritt (Hufeisen & Lutjeharms 2005; Hufeisen 2006). Auch der Englisch- (Jakisch 2015) und der DaF-Unterricht (Marx 2010) haben inzwischen das plurilinguale Vorwissen der Lerner und deren eigene Sprachlernerfahrungen als Faktoren des Sprachenwachstums entdeckt und integrieren dieses in die didaktische Steuerung. Vor diesem Hintergrund sei an den eingangs zitierten Grundgedanken der Interkomprehensionsdidaktik erinnert, wie ihn in der Lernpsychologie David Ausubel formulierte (1968: vi): "If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor in influencing learning is what the learner already knows." Anders gesagt: Ein Unterricht, der diese Faktoren ignoriert oder beiseitelässt, reduziert das Lernpotential der Lerner⁴⁹! Zugleich reduziert er – wie zahlreiche einschlägige empirische Arbeiten ausweisen – deren Möglichkeit, das Erlebnis von Selbstwirksamkeit im Umgang mit fremden Sprachen zu erfahren. Dass dies eine Voraussetzung der Sprachlernkompetenz überhaupt ist, sei hier in Erinnerung gerufen.

⁴⁸ Sie sind sowohl für die didaktische Lexikographie der romanischen Sprachen, des Englischen des Deutschen als auch für die Inputanalyse des Unterrichts romanischer Sprachen – nicht zuletzt im Rahmen von Mehrsprachigkeit – von Interesse. Adressiert sind sowohl deutschsprachige als auch anderssprachige Lerner einer ersten oder einer weiteren romanischen Sprache, darunter auch romanische Muttersprachler.

⁴⁹ Dass der rasche Erwerb von Lesekompetenz in einer neuen Sprache qua interkomprehensionsdidaktischer Verfahren i.d.R. zu einer deutlichen Positivierung von Selbstwirksamkeitserfahrungen auf dem Feld des Fremdsprachenlernens führt, zeigen übereinstimmend eine Reihe von empirischen Studien.

12.3 Was konkret wozu zu vermessen war

Aufgabe war die Vermessung des den vier größten romanischen Sprachen gemeinsamen Kernwortschatzes nach pädagogisch relevanten Kriterien⁵⁰. Die Voraussetzung bestand, wie erwähnt, in der Kompilation des romanischen Kernwortschatzes (Frequenzrang <5000) und seiner Präsentation in bedeutungsadäquaten mehrsprachlichen Serien mit der Kennzeichnung zahlreicher pädagogisch relevanter Marker. Dies ermöglichte ein pädagogisch sinnvolles, u.a. nach Sprachen kombinierbares elektronisches Suchen und eine variierbare Präsentation des Materials. Was nun speziell dessen Vermessung betrifft, so waren vor allem zwei Kriterien maßgeblich, und zwar die sog. Frequenzrang-Intervalle und der Grad interlingualer Formkongruenz/Opazität – beide bezogen auf die vier romanischen Zielsprachen⁵¹. Eine weitere Voraussetzung bestand in der Entwicklung einer angemessenen lexikostatistischen und lexikodidaktischen Methode. Dieser Schritt nahm erhebliche Zeit in Anspruch.

Die Erstellung des belastbaren, sechs bzw. sieben Sprachen umfassenden KRM-Inventars und seine Aufbereitung in elektronischen Lernapplikationen (Lern-Apps) erfolgte im Wesentlichen zwischen 2015 bis 2018. Deren weitgehende Fertigstellung war Voraussetzung für eine Vermessung des interlingualen Transferpotentials des KRM-Repertoires.

Im Detail zu entwickeln waren Verfahren, die Folgendes erlaubten:

- Kategorisierung und Vermessung des gesamten KRM-Inventars nach drei bzw. vier (für Italienisch) [Frequenzrang-Intervallen](#) mithilfe entsprechender Selektoren (R, R\$, M\$, R- bzw. M-).
- Identifikation und Quantifizierung von [grund- und kernwortschatzlichen Anteilen](#).
- Französisch und Englisch als Brückensprachen für Italienisch und Spanisch.
- Bewertung der Schulfremdsprache Latein als Quelle lexikalischer Transferbasen zur auf der Grundlage der KRM-Daten⁵².

⁵⁰ Dessen Beschreibung erfolgte außerhalb dieses Aufsatzes. Seine Entwicklung ist jedoch eine Vorbedingung für alle Zählungen (vgl. hierzu: im Detail Meißner 2016; auch Abel 2006).

⁵¹ Die Beschreibung dieses Schrittes erfolgte ebenfalls außerhalb dieses Aufsatzes. Seine Entwicklung ist eine Vorbedingung für alle Zählungen.

⁵² Die Beschreibung erscheint in einer lexikostatistischen Arbeit angemessen, weil die Frage des interlingualen, vor allem interlexikalischen, Transferpotentials von Latein für das Erlernen von romanischen Sprachen in der Diskussion um die Vor- oder Nachteile von Lateinunterricht eine zentrale Rolle spielt.

- ?Kategorisierung und Vermessung des gesamten KRM-Inventars nach fünf Transparenzstufen (T4, T3, T2, T1, T0) mithilfe entsprechender Selektoren (s. [Abschnitt 3.5](#), [Abb. 4](#)).
- Messbarkeit von weiteren lexikodidaktischen Markierungen (hier nur zum Teil berücksichtigt), z.B. für Zahl der Etyma pro Serie und der semantischen Distanz von Lemmata einer Serie u.a.m..
- Grundsätzliche Trennung von einzelziel- und gemischt-sprachlichen Grundmengen.
- Trennung nach sog. primären und sekundären (romanischen) Zielsprachen und Ermöglichung von Zählungen in der Kombination von zwei oder mehr Sprachen.
- Intraserielle Trennung von einzelsprachlich-formdivergenten und einzelsprachlich-formkonvergenten Lemmata, so dass zwischen Wortformen und ihren Reduplikationen (z.B. in Fällen von Polysemie) unterschieden werden kann.
- Interserielle Trennung von formdivergenten und formkonvergenten Bedeutungsadäquanzen, so dass sich grundsätzlich zwei Möglichkeiten zur Legung der aus zwei oder mehr Sprachen komponierten Grundmengen ergeben: aus (X) oder aus (X OR Xw) mit (Y) oder (Y OR Yw).
- Quantifizierung der Transparenzstufenanteile einzelzielsprachlich und nach einzelsprachlichen Kombinationen mit der Erweiterungsmöglichkeit um Deutsch und Englisch aus dem Variablen des Selektors (X) oder (X OR Xw) komponierten Grundmengen.
- Die Quantifizierung der Transparenzstufenanteile in der Kombination von mehreren Sprachen.
- Identifikation und Komputation von intra- (und interlingualen) Transferpotentialen innerhalb gemischt-formdivergenter Serien durch [\\$-Marker](#) ([Abb. 24](#)). Dieser Schritt antwortet auf die Frage: Welche Lemmata stellen in gemischt formdivergenten Serien füreinander Transferbasen bereit?
- Identifikation und Zählung einzelzielsprachlicher Formabweichungen innerhalb ansonsten serieller Formkongruenz (RoS*-X) der Transparenzklasse T3.
- Nochmals zu T3: Für RoS*-X-Fälle [Berechnung von Transparenz- aus Opazitätsdaten](#) ([Abb.24](#)).
- Identifikation und Komputation von morphosemantisch minimal distanten Lemmata mithilfe der FaX-Markierung ([FaX-Marker](#)).
- Einstufung semantischer Distanzen in Fällen einer ‚nahen‘ Polysemie und ihre Quantifizierung durch einen Distanzkoeffizienten und durch die Zahl der Etyma pro Serie (hier nicht gezählt).
- Quantifizierung der [englischen](#) und deutschen Transferbasenpotenziale im romanischen Kernwortschatz ([Abb. 10](#)).

- Gezielt kombinierte Zählungen romanischer Lemmata mit Englisch als Brückensprache.
- Quantitative Extrapolation des über die Romania hinausreichenden, gemeinsamen europäischen Kernwortschatzes ([Tab. 41](#)).
- Kombiniert-sekundärsprachliche Zählungen aus gemischt- primärsprachlichen Grundmengen.
- Herstellung eines ‚raschen Überblicks‘ zu den Anteilen der verwandten Selektoren am Gesamt des KRM-Datenbestandes.
- Da bisher noch keine voll umfängliche Erfassung der Interligalexe vorliegt, bleiben die Ergebnisse dieser möglichen Erhebung und Analyse an dieser Stelle ausgespart⁵³.

Überwiegend wurden die oben zusammengestellten Beschreibungsziele, wie in diesem Beitrag in vielen lexikostatistischen Detailaussagen festgehalten, erreicht. Hierzu lassen sich folgende Punkte anführen:

12.4 Lexikodidaktische und -statistische Gesamtcharakteristik

Ungefiltert und in gesamtromanischer Betrachtung beläuft sich der interlingual seriell-komplett formkongruente Anteil des kernwortschatzlichen Vokabulars auf mindestens 60 Prozent ([Abb. 6](#), [Abb. 44](#)). Eine drei Zielsprachen umfassende Transparenz liegt (zusätzlich) bei 16 Prozent vor, so dass sich sagen lässt: Der interlinguale Transparenzanteil der kernwortschatzlichen romanischen Lemmata liegt bei 75 bis 76% für drei oder vier Sprachen. Jeweils paarige Formkongruenzen innerhalb der vier Sprachen erfassenden Zählung lassen sich im Umfang von 6% vorfinden. Man kann diese Menge je nach der Kombination der miteinander verglichenen Sprachen entweder als „eher transparent“ oder als „eher opak“ bezeichnen. Die voll opaken Serien belaufen sich auf 9%, zu denen jene Serien hinzuzurechnen sind, die lediglich zwei Sprachen durch eine gemeinsame morphologische Transferbasis miteinander verbinden können. Diese T1-Werte umfassen immerhin 11 Prozent, so dass die Transparenzrate 9% plus 11% umfasst. Die Ziehung dieser Zahlen erfolgte aus einem Gesamt von 9557 Serien.

Exkurs: Der transparente Wortschatz der europäischen Sprachen ist selbstredend viel größer als der romanische Kernwortschatz (Zimmer 1990: 75ff; [Abb. 1](#); hierzu: Meißner 1989). Dies beziffern die hohen Transparenzraten, die das Deutsche und vor allem das [Englische](#) mit dem romanischen Wortschatz teilen. Die Formkongruenz zwischen Englisch, Deutsch und den romanischen Sprachen beläuft sich auf nicht

⁵³ Es ist geplant, die Rolle der Interligalexe in einem separaten Beitrag eingehend zu beleuchten.

weniger als 52 bis 55% in formdivergenten Grundmengen. Im Gesamt des KRM-Inventars verringert sich der Anteil an Formkonvergenzen, den das Deutsche bzw. die Kombination Deutsch plus Englisch mit dem romanischen Kernwortschatz teilen, auf 30%. Die 0,6% des exklusiv Deutsch mit Romanisch geteilten Anteils fallen nicht ins Gewicht. Dies ist das Potential der formseitigen lexikalischen Ressourcen, mit dem der Englischunterricht das Erlernen einer romanischen Sprache bzw. von Lesekompetenz in einer oder mehreren romanischen Sprachen erleichtern kann⁵⁴. Der Vorteil betrifft tendenziell nicht nur romanischen, sondern – allerdings mehr oder weniger eingeschränkt – auch die anderen Sprachen der europäischen Koine.

In der fachdidaktischen Diskussion um die Vorzüglichkeit dieser oder jener romanischen Sprache als optimale Brückensprache für das Erlernen der Schwestersprachen spielte immer wieder die Frage nach Art und Umfang der gelieferten Transferbasen eine Rolle. Sie betrifft den Kernwortschatz in besonderer Weise. Aufschlussreich sind in diesem Zusammenhang zunächst die T4- und T3-Raten für die einzelnen primären (vgl. [Abb. 26](#)) in Kombination mit den sekundären Zielsprachen. Wie die KRM-Komputationen belegen, liegen die Transferpotenziale der einzelnen romanischen Sprachen für ihre romanischen Schwestern sehr eng beieinander. Allein das Französische weist eine deutlich höhere Anzahl an Profiwörtern auf als die südlichen Nachbarn. Sein Wortschatz ist daher weniger transparent. Dies belegt gerade die T3-Kohorte. Sie erreicht den höchsten Wert an opaken Wörtern aller romanischen Zielsprachen, und zwar 1553 (699 Fälle aus [RoS*-F] und 855 aus [oFk]; der Gesamtwert für Spanisch wäre 868; in gesamtserieller Zählung). Ein weiteres Faktum kommt hinzu, und zwar der Lautstand der französischen Erbwörter⁵⁵: Das Französische kennt nicht nur eine stärkere Modifikation im Bereich des Vokalismus (PATREM > *père, padre*; CANEM > *chien, it. cane, sp. can* usw.) gegenüber dem Latein und den romanischen Schwestersprachen, französische Wörter sind auch im Vergleich zu ihren romanischen Entsprechungen zumeist ‚kürzer‘ (*eau* ~ *agua, cher* ~ *caro, tailler* ~ *tagliare* u.a.m.)⁵⁶.

⁵⁴ Eine konsequente Berücksichtigung der Frequenzrang-Intervalle der disambiguierenden Sprachen (Deutsch und Englisch) erfolgte nicht.

⁵⁵ Das Französische zeigt hier – was lexikologisch versierte Leserinnen und Leser nicht überraschen kann – den höchsten Wert von 1553 (699 Fälle aus [RoS-F] und 855 aus [oFk]; der Gesamtwert für Spanisch wäre 868).

⁵⁶ Morphologisch distant zu den romanischen Sprachen sind natürlich die erbwörtlichen Formen (*eau, père, chien, beaucoup* usw.). Für diese bieten die Schwestersprachen in der Regel ihrerseits überwiegend keine für den linguistischen Laien erkennbare Transferbasen; umgekehrt jedoch tut dies das Französische mit seinem buchwörtlichen Bestand durchaus: *aqueux, paternel/patron, canin, moult/multitude* (*agua, água, acqua; padre/padrone; cane, can, cão; moltitudine, multitud, multidão*). Jamets (2007) Arbeit zeigt, dass ein möglicher

Für das romanisch-interkomprehensive Potential des Französischen bedeute dies – wie man auf den ersten Blick meinen könnte – einen Nachteil. Dieser Eindruck erhärtet sich aber nur solange, wie man das Französische als „lernensunwerte“ Sprache außerhalb der Analyse hält. (Es dürfte allerdings kaum möglich sein, hierfür nachvollziehbare Gründe zu finden.) All diese Argumente sind jenen von Horst G. Klein (2002) hinzuzufügen, der das Französische als optimale Brückensprache für die romanischen Schwesteridiome bezeichnet.

Ansonsten unterscheidet sich – angesichts der Summe aus T4- plus T3-Serien – in kaum nennenswertem Umfang voneinander. Dies kann freilich für im KRM nicht erfasste romanische Sprachen anders ein. Dabei ist vor allem an das Rumänische zu denken. Die hohe Gleichwertigkeit ergibt sich aus der Bidirektionalität von Transferbasen und Transfereffekten (im Zusammenhang mit dem hohen Anteil an teilweise oder vollständig formkongruenten Serien).

Identifikationstransfer zwischen Italienisch und Französisch von Italophonen dann nur schwer realisiert werden kann, wenn die zielsprachliche Gestalt (Französisch) gegenüber der italienischen Transferbase ‚verkürzt‘ ist.

Transferpotenziale zwischen den romanischen Zielsprachen innerhalb erweiterter oder einfacher zweisprachlicher Mengen.

	F OR Fw MIT P OR Pw	F OR Fw MIT I OR Iw	F OR Fw MIT S OR Sw	F OR Fw MIT P	F OR Fw MIT S	F OR Fw MIT I	F OR Fw F MIT I	F OR Fw F MIT P	F OR Fw F MIT S
GESAMT	9545	9549	9546	7757	7590	8062	6762	6582	6405
T4	4345	4347	5604	3967	3885	3971	4630	4589	4482
T3	1487	1487	1487	1088	1055	1176	861	827	797
T2	531	531	531	402	393	377	274	270	264
T1	1050	1050	1050	703	685	782	539	492	459
T0	854	854	854	540	534	654	437	383	383
T4%	46%	46%	46%	51%	51%	49%	68%	70%	70%
T3%	23%	23%	23%	14%	14%	16%	13%	13%	12%
T2%	6%	6%	6%	5%	5%	5%	4%	4%	4%
T1%	11%	11%	11%	9%	9%	10%	8%	7%	7%
T0%	10%	10%	10%	7%	7%	8%	6%	6%	6%

Abbildung 47: Synopse der Transferpotenziale

12.5 Welche methodischen Konsequenzen sind aus der quantitativen Analyse des KRM zu ziehen?

Im Sinne Ausubels (1968) sollte der Fremdsprachenunterricht den Lernern möglichst frühzeitig bewusstmachen⁵⁷, dass ihnen ihr eigenes lernrelevantes Vorwissen sehr lernwirksame Strategien und Ressourcen bietet. Der internationale Kernwortschatz ist dabei ein wesentlicher Faktor. Denn er verbindet das, was Lerner schon aus der einen oder anderen Sprache wissen, mit dem, was sie in der neuen Zielsprache erst noch lernen müssen. Dies bedeutet, dass die neuen, im mentalen Lexikon von Lernern noch ungefestigten sprachlichen Schemata an ihren längst solidifizierten Korrespondenzen aus den Nachbarsprachen andocken können: Auf diese Weise greift das Erlernen von neuem Vokabular fester, weiter und schneller, als es in einzelsprachlicher Semantisierung der Fall ist. Fazit: ‚Vernetzen‘ und ‚Vergleichen‘ von Wörtern; Formen und Strukturen als Ausdruck von Sprachen- und Sprachlernbewusstheit sind daher wichtige Orientierungen für eine mehrsprachige Wortschatzarbeit (Meißner 2000; Neveling 2006). Grundlegend erscheinen auf der unterrichtsmethodischen Ebene drei bzw. vier Lernziele:

1. Die Kompetenz, intermorphologische Korrespondenzen zu identifizieren, etwa nach dem Muster der *Sieben Siebe* des *EuroComRom* (Klein & Stegmann 1999).
2. Zu dieser Kompetenz gehört auch die Fähigkeit zur Identifikation von minimalen morphologischen Analogien (\$-Markierungen und FaX-Gruppe).
3. Möglichst frühzeitige Reduktion des für Lerner interlingual opaken Wortschatzes im Falle des Erlernens von Lesefertigkeit in einer neuen Sprache, z.B. von Italienisch nach Französisch oder Englisch. Dies wird durch eine gezielte und frühzeitige Präsentation dieses lexikalischen Segments (T1 und T0 und – zielsprachenspezifisch – die Opaken aus T3) erreicht.
4. Wie Forschungen zu den Wirkungen von Übungsformaten belegen, unterstützen die Formate selbst die Ausbildung entsprechender Suchroutinen zur Einleitung einer Transferoperation. Die Unterrichtspraxis kennt derlei Formate z.T. seit jeher (z.B. Interlineardarstellungen in zwei oder gar mehr Sprachen).

Es liegt im Wesen des Kernwortschatzes einer Sprachfamilie, dass sich mit seinem Erwerb Ressourcen verbinden, die positive Effekte auf das Erlernen von mehreren Sprachen verbinden (vor allem wenn diese ein und derselben Familie angehören). Mehr als für andere Sprachfamilien gilt dies für die romanische. Denn ihr Wortschatz

⁵⁷ Vor allem Morkötter (2016) und zuvor Doyé (2009: 26ff) zeigen eindrucksvoll, wie Kinder mit Hilfe des Sprachenvergleichs disambiguieren.

ist eine bedeutende Teilmenge des europäischen Kernwortschatzes bzw. Wortschatzes überhaupt. Sucht man nach den Anteilen eines slawischen, germanischen und romanischen Sprachen gemeinsamen Wortschatzes, so findet man vor allem romanisches, zumeist ‚gelehrtes‘, Sprachmaterial. In der Tat lässt sich von den slawischen als auch den germanischen Sprachen sagen, dass ein Großteil ihrer vor allem gelehrten Register aus Wortstämmen und Wortbildungselementen besteht, die auch im KRM begegnen (Tafel et al. 2009: 50ff.). Dies gilt in noch weitaus größerem Maße für den intra- und extrazentralen Wortschatz der Wissenschaften mit seinen lateinischen und griechischen Kompositionselementen.

Schon früh wurde in Beispielen dargelegt (u.a. Abel 1971), wie mehrsprachige Lesekompetenz im Unterricht einer romanischen Sprache gefördert werden kann. Diese Möglichkeit zu nutzen, ist heutzutage wichtiger denn je, weil Mehrsprachigkeit und Mehrkulturalität, der bereichernde Umgang mit diesen in der heutigen vielsprachigen Lebenswelt, die Fähigkeit der fremd- und mehrsprachlichen Weiterbildung und die Offenheit gegenüber Sprachen und Kulturen zu den Kompetenzen gehören, die für ein erfolgreiches Leben in einer globalisierten Welt wichtige Voraussetzungen sind; wie zahlreiche Dokumente der Europäischen Union und des sprachenpolitischen *decision making* betonen. Dabei ist daran zu erinnern, dass Mehrsprachigkeits- und Interkomprehensionsdidaktik Ansätze sind, die den herkömmlichen Unterricht ergänzen, nicht ersetzen.

Angesichts des mit guten Gründen feststellbaren Methodensynkretismus der fremdsprachendidaktischen Steuerung und erst recht der Praxen besteht der Gegensatz zwischen überkommenen und neuen Methoden nicht zuletzt in den traditionellen, lern- und lehrseitigen Stereotypen zu Lehren und Lernen; er ist viel weniger in der Eigenschaft des Lernstoffes selbst begründet, dessen Auswahl und methodische Präsentation der fachlich begründeten pädagogischen Urteilskraft der Lehrerinnen und Lehrer in Bezug auf ihre Lerngruppen folgt.

13. Anwendungen (Beispiele)

Das unbewusste, aber latent-bekanntes lernrelevante Wissen wecken, und das zum erfolgreichen Weiterlernen notwendige unbekanntes Wissen den Lernern zugänglich zu machen, erfordert einen entsprechenden Input. Die folgenden Beispiele stehen für diese Absicht. Die Sprachen sind willkürlich genannt, d.h., dass die Anwendungen in anderen Sprachenkombinationen als die genannten erfolgen können.

Das erste Beispiel erfasst alle opaken Formen für die Zielsprache Italienisch. Die Syntax lautet (oFk or Ros*-I in ROS- MIT IT=I). Die Zählung erbringt 945 Fälle.

423	CHAUFFER	riscaldare		oFk	I
9628	correspondance	riscontro		RoS-I	I
7739	recouvrement (de l'imp	riscoSSIONE (delle impost		oFk	I
2126	TOUCHER	riscuotere		oFk	I
3780	réentendre	RISENTIRE		oFk	I
2817	épargner	risparmiare		oFk	I
4737	épargne	risparmio		oFk	I
9638	retirer	ritrattare		RoS-I	I
3784	TOURNER	RIVOLGERE		oFk	I
3785	CHOSE	ROBA		RoS-I	I
9642	contrat	rogito		oFk	I
9647	rouge à lèvres	rossetto		oFk	I
2950	ROULER	rotolare		RoSED-I	I
1834	RENVERSER	rovesciare		oFk	I
9652	gâché	rovinato		oFk	I
7573	désordonné	rovistato		oFk	I
1906	ROULER	ruotare		oFk	I
4385	essai	saggio	FaF	RoS-I	I

Abbildung 48: Ausschnitt aus der Opaken-Liste für Italienisch

Die Erweiterung der Grundmenge (IT = I OR Iw) würde die Okkurrenzen auf 1181 erhöhen. Eingeschlossen wären in diesem Fall die Reduplikationen. Bereits hier ist daran zu erinnern, dass Reduplikationen (Wiederholungen) grundsätzlich einer mnemotechnischen Strategie entsprechen. Dies gilt besonders, wenn sie in neuen Ko-Texten auftauchen, was aufgrund der seriellen Darstellung in der KRM-Datenbank und den Lernapplikationen gegeben ist.

Natürlich ist die Auswahl der Zielsprache Italienisch willkürlich gewählt. Ebenso gut könnte auch Spanisch oder Portugiesisch Zielsprache sein. Da die insgesamt 945 Lemmata die undurchsichtigen Formen des Italienischen umfassen und die serielle Darstellung zugleich bekannte Formen einer anderen Kontrast- und Brückensprache – hier Französisch – darstellen, darf die Darstellungsform als ‚lernwirksam‘ bezeichnet werden (sofern dieses Segment frühzeitig in das Sprachcurriculum implementiert wird). Die Zusammenstellung liefert das Material für weitere Übungsformate, die hier nicht erläutert werden.

Das zweite Selektionsbeispiel erfasst einen kleinen Ausschnitt englischer Transferbasen für die Zielsprache Italienisch. Sie Syntax ist: (RoSE OR RoSED OR RosE in ROS-MIT IT=I).

7977	fisica	RoSED	physics
8908	fisiologia	RoSED	physiology
8911	fisioterapista	RoSED	physiotherapist
8907	fisioterapia	RoSED	physiotherapy
9414	pianoforte	RoSED	piano
7466	pittoresco	RoSED	picturesque
316	PEZZO	RoSE-F	piece
7429	parcella	RoSED	piece of property, parcel
6225	perforare	RoSE-F	pierce, perforate
8674	cumulo	RoSED	pile, cumulation
8125	pellegrinaggio	RoSE	pilgrimage, journey
6251	pilota	RoSED	pilot
6252	pilotare	RoSE	pilot, fly, drive
7463	pino	RoSED	pine (tree)
7585	rosato	RoSED	pink, rosy

Abbildung 49: Ausschnitt aus der Selektion: englische Transferbasen für Italienisch

Insgesamt bringt es die Zahl der vom Englischen aus möglichen interlingualen Transferoperationen auf (alle RoSE* oder RoSED*-markierten) 4190 Serien.

Ist man an bestimmten Sprachkombinationen mit Englisch interessiert, so ist es möglich, Lemmata nach konkreten Merkmalen, z.B. RoS*-S, aus der jeweiligen Selektion herauszufiltern. Hier ein Ausschnitt der portugiesischen Profilwörter und ihrer Serien, gezogen aus einer nur das Englische betreffenden Grundmenge.

Eine solche Ziehung ist natürlich interessant, wenn es darum geht, Lerner für ihre englischen Kenntnisse und deren Transferpotenzial für die Zielsprache Portugiesisch zu sensibilisieren.

Fr	Pg	Sp	It	RoS-	En
BIEN	BENS	BIEN	BENE	RoS-P	property
BORD	BEIRA	BORDE	BORDO	RoS-P	limit, border
clameur	BERRO	clamor	clamore	RoSE-P	clamour
gomme	BORRACHA	GOMA (PER BORRAR)	gomma (per cancellare)	RoS-P	rubber
CUIR	CABEDAL	CUERO	CUOIO	RoS-P	leather
pipe	CACHIMBO	pipa	pipa	RoS-P	pipe
omnibus	CAMIONETA	ómnibus	AUTOBUS	RoS-P	omnibus
PAUVRE	COITADO	POBRE	POVERO	RoS-P	poor
RÉPARATION	CONCERTO	reparo	riparazione	RoS-P	repair
CERTAINEMENT	DECERTO	CIERTO (EN ~ MODO)	CERTO (DI ~)	RoS-P	certainly
défaire	DESEMBRULHAR	deshacer	disfare	RoS-P	unpack
PAQUET	EMBRULHO	PAQUETE	PACCHETTO	RoS-P	packet
ROUGE	ENCARNADO	ROJO	ROSSO	RoS-P	red
APPUYER	ENCOSTAR	APOYAR	APPOGGIARE	RoS-P	lean, touch
adorn	ENFEITAR	adornar	adornare	RoS-P	adorn, decorate
TISSU	FAZENDA	tejido	tessuto	RoS-P	cloth, material
FORME	FEITIO	FORMA	FORMA	RoS-P	shape, form
four	FOGÃO	horno	fornello	RoSE-P	hearth, oven, furnace

Abbildung 50: portugiesische Profilwörter im Kontrast mit englischen Transferbasen

Je nachdem, ob das Lehrziel ‚Förderung der Lesekompetenz in mehreren romanischen Sprachen‘ beträgt oder aber nur ‚in einer bestimmten Zielsprache‘, kann der Kontrast entsprechend gelegt werden.

Die folgende Tabelle präsentiert einen Ausschnitt englischer Transferbasen für romanische Zielsprachen generell. Wird z.B. Schülern eine solche Liste gegeben, verbunden mit einer Aufgabe, die entsprechenden Lemmata zu unterstreichen, die bereits kennen, so sensibilisiert sie dies besonders stark dazu, ihre Englischkenntnisse stark in ihre Suchprozesse zur Initiierung eines Transfers einzubeziehen. Die Erwähnung der Selektoren (RoSE, RoSED usw.) wirkt motivierend, weil sie die Effizienz der erfassten Lexeme für die Mehrsprachenlernen signalisiert.

Vermessung des romanischen Kernwortschatzes

116	RoSED	ANGLAIS	INGLESE	inglês	INGLÉS	English
9041	RoSE	ingrédient	ingrediente	ingrediente	ingrediente	ingredient
3615	RoSE	initial	INIZIALE	INICIAL	INICIAL	initial
5910	RoSE	initialement	inizialmente	inicialmente	inicialmente	initially
3616	RoSE	initier	INIZIARE	INICIAR	INICIAR	initiate, begin
4002	RoSED	initiative	INIZIATIVA	INICIATIVA	iniciativa	initiative
902	RoSE	ENNEMI	nemico	INIMIGO	ENEMIGO	enemy
3041	RoSE	ENNEMI	nemico	INIMIGO	ENEMIGO	inimical
7303	RoSED	injection	iniezione	injecção	inyección	injection
5911	RoSED	injecter	iniettare	injectar	inyectar	inject
5913	RoSE	injustice	ingiustizia	injustiça	injusticia	injustice, unfairness
5912	RoSE	injuste	ingiusto	injusto	injusto	unjust
5914	RoSE	innocence	innocenza	inocência	inocencia	innocence
3127	RoSE	innocent	innocente	inocente	INOCENTE	innocent
5915	RoSED	innovation	innovazione	inovação	innovación	innovation
8037	RoSED	innovateur	innovatore	inovador	innovador	innovative
8038	RoSE	inscrit	iscritto	inscrito	inscrito	enrolled, inscribed
1162	RoSED	INSECTE	insetto	INSECTO	INSECTO	insect
7283	RoSE	insécurité	insicurezza	insegurança	inseguridad	insecurity
7284	RoSE-F	sûr (pas ~)	insicuro	inseguro	inseguro	unsure, unsafe
5921	RoSED	insensible	insensibile	insensível	insensible	insensitive
7285	RoSE	inséparable	inseparabile	inseparável	inseparable	inseparable
3621	RoSED	insérer	INSERIRE	inserir	INSERTAR	insert, include
5922	RoSE	insignifiant	insignificante	insignificante	insignificante	insignificant
7286	RoSE	insinuer	insinuare	insinuar	insinuar	hint, insinuate
9050	RoSE	insistance	insistenza	insistência	insistencia	insistence

Abbildung 51: Ausschnitt englischer Transferbasen für romanische Sprachen

Natürlich ist es auch möglich, nach bestimmten Buchstabenkombinationen innerhalb von Wörtern zu suchen. Dies empfiehlt sich besonders für die Identifikation regelhafter morphologischer Korrespondenzen, wie sie die *Sieben Siebe* des EuroComRom auflisten (Klein & Stegmann 1999).

Ein Beispiel: Eine der hervorstechendsten morphologischen Kontrastmerkmale des Italienischen mit seinen romanischen Schwestern betrifft die Korrespondenzregel

,Konsonant + L ↔ Konsonant + l' (Typ sp. *flor* ~ it. *fiore*). Die Syntax lautet im Feld FR: ([*bl* OR *cl* OR *chl* OR *dl* OR *fl*] MIT [*bi* OR *ci* OR *chi* OR *di* OR *fi* usw.]) im Feld IT. Die Selektion zeigt 166 Fälle an. Natürlich betrifft die Auswahl nicht alle Okkurrenzen der Konsonant-L~Konsonant-l-Korrespondenzen, sondern nur die der Strecke *bl* bis *fl*, so dass zur vollständigen Erfassung der interlingualen Oppositionen eine weiter fassende Selektion erforderlich ist, und zwar für die Buchstabenkombination *gi*/*ghi* bis *zi* für das Feld IT. In unserem Beispiel erfolgt diese Vervollständigung aus Gründen der Platzökonomie nicht. (Vollständige Selektionen sind jedoch selbstverständlich verfügbar.)

Wie die Auswahl der Ziehung belegt, greift sie über die genannte Korrespondenzregel hinaus (*bicyclette-bicicletta*). Daher wird manuell eine zusätzliche Kennzeichnung im Feld SONDERFÄLLE durchgeführt. Die mit „x“ gekennzeichneten Sonderfälle trennen die Regelfälle von anderen Lemmata mit dem Nexus [*L]. Die gefundenen Okkurrenzen werden sodann zusammengefasst und alphabetisch im Feld IT sortiert.

ID	Fr	It	Pg	Sp	SONDERFÄLLE
78	AIMABLE	AMABILE	AMÁVEL	AMABLE	
279	BICYCLETTE	BICICLETTA	BICICLETA	BICICLETA	
288	BLANC	BIANCO	BRANCO	BLANCO	x
292	BLOND	biondo	LOIRO	RUBIO	x
386	CERCLE	CERCHIO	CÍRCULO	CÍRCULO	x
466	CLAIR	CHIARO	CLARO	CLARO	x
468	CLÉ	CHIAVE	CHAVE	LLAVE	x
470	CLOU	chiodo	PREGO	CLAVO	x
472	CLORE	CHIUDERE	CERRAR	CERRAR	x
473	clôture	recinto	CERCA	CERCA	
617	COUVERCLE	coperchio	TAMPA	TAPA	x
682	DÉCLARER	DICHIARARE	DECLARAR	DECLARAR	x
1146	IMPOSSIBLE	IMPOSSIBILE	IMPOSSÍVEL	IMPOSIBLE	
772	DIABLE	DIABOLO	diabo	DIABLO	
850	ÉCLAIRER	CHIARIRE	esclarecer	ACLARAR	x

Abbildung 52: Phonologische Korrespondenz FR = Konsonant+L ~ IT = Konsonant+l

Hier nun die gefundenen Okkurrenzen aus der eingeschränkten Selektion:

BIANCO, biondo, CERCHIO, CHIARO, CHIAVE, chiodo, CHIUDERE, coperchio, DICHIARARE, CHIARIRE, FIORE, soffiare, BIANCHERIA, CHIUSO, fiorire, fiamma, CHIARAMENTE, DICHIARAZIONE, INDUBBIAMENTE, CHIAREZZA, CHIAVE, BIANCO, biasimare, CHIUDERE, bibbia, chiarificare, chiarire, schiavo, fiume, gonfiare, cerchio, sfiorare, richiamo, schiarimento, chiarimento, chiudibile, fianco, fiorentino, fiorino, gonfiore

Abbildung 53: Alle FR = Konsonant+L ~ IT = Konsonant+I-Korrespondenzen aus KRM

Die bisherige Auswahl müsste also weiterfassen, um vollständig zu sein (vgl. allein *glace* ~ *ghiaccio*, *Angleterre* ~ *Inghilterra* u.a.m.). Abb. 54 zeigt einen Ausschnitt aus der Strecke fr. [PL] mit it. [PI]:

PLUS (EN ~), (DE ~)	PIÙ (DI ~)	MAIS (DE ~)	ADEMÁS
déplaire	DISPIACERE	disagradar	desagradar
EMPLOYER	IMPIEGARE	EMPREGAR	EMPLEAR
PLACE	PIAZZA	PRAÇA	PLAZA
PLACER	piazzare	COLOCAR	COLOCAR
PLAINE	PIANURA	planície	LLANO
PLAIRE	PIACERE	GOSTAR	GUSTAR
PLAISIR	PIACERE	PRAZER	PLACER
PLAN	PIANO	PLANO	PLANO
PLANTE	PIANTA	PLANTA	PLANTA
PLANTER	PIANTARE	PLANTAR	PLANTAR
PLAT	PIATTO	PRATO	PLATO
PLAT	PIANO	CHATO	LLANO
PLEIN	PIENO	CHEIO	LLENO
PLEURER	PIANGERE	CHORAR	LLORAR
PLEUVOIR	PIOVERE	CHOVER	LLOVER
PLIER	PIEGARE	DOBRAR	DOBLAR
PLUIE	PIOGGIA	CHUVA	LLUVIA
PLUS	PIÙ	MAIS	MÁS

Abbildung 54: Ausschnitt aus der Selektion FL="*pl*" MIT IT="*pi*"

Die folgende Beobachtung illustriert, wie zweisprachige Buchstabenkorrespondenzen erweitert werden können zu weiteren Korrespondenzregeln, hier zur Regel:

FR	IT	PG	SP
L	*I*	*R*	*L*
L	*I*	*CH*	*LL*

Beispiele: *plaire-piacere-prazer-placer*, *plat-piatto-prato-plato* bzw. *plat-piano-chato-llano*, *plein-pieno-cheio-lleno*, *pleuvoir-piovere-chover-llover*. Lerner erkennen in solcher Darstellung spätestens auf den zweiten Blick die phonologischen Oppositionsregularitäten zwischen den Zielsprachen. – Solche Illustrationen sind im Zusammenhang mit der Förderung einer mehrsprachlichen Lesekompetenz aufschlussreich und lernrelevant. Sie befördern offensichtlich auf die Fähigkeit, selbstständig sprachenvergleichend zu lernen.

Bei der Identifikation von interlingualen Buchstabenkombinationen empfiehlt es sich, entweder mehrere Vergleichssprachen heranzuziehen oder zumindest diejenige Sprachenkombination zu wählen, die mit der Zielsprache die meisten Übereinstimmungen hat. Und selbstredend sind – wenn möglich und sinnvoll – die Sprachen der Schülerinnen und Schüler bzw. der Lerner zu berücksichtigen.

Das Französische unterscheidet sich von den Schwestersprachen u.a. durch den initialen Anlaut [ch] /ʃ/- im Vergleich zum romanisch /k/ (Typ: fr. *chien*, it. *cane*). Dies ist z.T. eine Sekundärfolge des Lautwechsels von lat. A zu [e,i] in offener Silbe (MATER > *mère*, PRATUM > *pré*). Im Italienischen wird dies durch [c^{+e:i}] /tʃ/ (*ciliegia*), im Spanischen durch [c^{+e:i}] bzw. /θ/ (*cereza*) gefasst. Doch kennt das Französische natürlich auch [ch] vor dunklen Vokalen, z.B. *chat* (*gatto*), *chanter* (pg./sp. *cantar*), *choux* (pg. *couve*), *chute* (*caduta*, *queda*, *caída*).

Eine erste Selektion (*ch* in FR MIT *c* in IT) weist 323 Okkurrenzen aus. (Genug Material zur Konstruktion von Übungen und Aufgaben.) Wiederum erfolgt eine manuelle Auswahl der „Sonderfälle“ mit der Kennzeichnung „x“ oder „y“ zur Markierung von „Ausnahmen“ oder pädagogisch zu erklärenden Fällen. Die Selektion gibt Formen aus, die zumeist auf den buchwörtlichen Erbteil des französischen Wortschatzes zurückgehen, die dem Lautwandel von lat. [C] /k/ zu fr. [ch] /ʃ/ nicht unterlagen.

Vermessung des romanischen Kernwortschatzes

ID	Fr	It	Pg	Sp	SC
27	ACHETER	COMPRIARE	COMPRAR	COMPRAR	
306	BOUCHE	BOCCA	BOCA	BOCA	x
307	BOUCHER	MACELLAIO	açougueiro	carnicero	
320	BRANCHE	BRANCA	RAMO	rama	x
321	BRANCHE	BRANCA	SECTOR	sector	x
330	broche	broche	broche	broche	y
341	CACHER	NASCONDERE	ESCONDER	ESCONDER	
359	char	CARRO	CARRO	CARRO	x
392	CHACUN	ciascuno	cada umo	cada uno	x
394	CHAÎNE	catena	CADEIA	CADENA	x
396	CHALEUR	CALORE	CALOR	CALOR	x
397	CHAMBRE	CAMERA	QUARTO	cámara	x
398	CHAMP	CAMPO	CAMPO	CAMPO	x
399	CHANCE	OCCASIONE	OPORTUNIDADE	oportunidad	+++
400	change	CAMBIO	câmbio	cambio	x
401	CHANGEMENT	CAMBIAMENTO	câmbio	cambio	x
402	CHANGER	CAMBIARE	CAMBIAR	CAMBIAR	+++
403	CHANSON	CANZONE	CANÇÃO	CANCIÓN	x
405	CHANT	CANTO	CANTO	CANTAR	x
406	CHANTER	CANTARE	CANTAR	CANTAR	x
407	CHAPEAU	CAPPELLO	CHAPÉU	SOMBRERO	x
409	CHARBON	CARBONE	CARVÃO	carbón	x
410	CHARGE	CARICA	CARGA	CARGO	x
411	CHARGE	CARICO	CARGA	carga	x
412	CHARGER	INCARICARE	ENCARREGAR	ENCARGAR	
413	charité	CARITÀ	caridade	CARIDAD	x

Abbildung 55: Phonologische Korrespondenzen für FR = *ch* und IT ⁵⁸“c“

Eine Eigenheit des Portugiesischen, die Lerner immer überrascht, betrifft sog. „Verkürzungen“ im Vergleich zu den Brückensprachen. Zu denken ist an pg. *ter* ~ sp. *tener*, fr. *tenir*; *honra* ~ *honor*, *honneur*, en. *honor/honour* usw. Aus der empirischen Interkomprehensionsforschung zum Hörverstehen ist bekannt, dass Verkürzungen dem interlingualen Identifikationstransfer erhebliche Hindernisse entgegensetzen

⁵⁸ Die Kennzeichnung +++ dient anderen, in dieser Publikation nicht erwähnten Zwecken.

(Jamet 2007). Ähnliches gilt auch für das Wiedererkennen von orthographisch deutlichen Unterschieden zwischen Sprachen.

Die „FaP“-Markierung in PROFIL filtert, weniger ausgeprägt als RoS-X, in Kombination mit manuell durchgeführten x-Markierungen in der Rubrik SONDERFÄLLE solche Verkürzungen (x) oder aber erhebliche Formabweichungen (xx).

PG	SP	PROFIL	SFÄLLE
alheio	OTRO (d'~)	FaP	xx
AVALIAÇÃO	valutación	FaP	xx
AVALIAR	EVALUAR	FaP	x
CADEIA	chirona	FaP	xx
CONTEÚDO	contenido	FaP	x
DESCER	DESCENDER	FaP	x
DISPOR	DISPONER	FaP	x
fio	HILO	FaP	x
GERALMENTE	generalmente	FaP	x
GRÃO	GRANO	FaP	x
herdar	HEREDAR	FaP	x
HONRA	HONOR	FaP	x
LÃ	LANA	FaP	x
LUA	LUNA	FaP	x
maciço	masivo	FaP	xx
MÃO	MANO	FaP	x
MEIO	MEDIO	FaP	x
MOEDA	MONEDA	FaP	x
NEM	NI	FaP	xx
NOMEAR	NOMBRAR	FaP	xx
OBTER	OBTENER	FaP	x
PAIRAR	planear	FaP	xx
PESSOA	PERSONA	FAP	x
PÓ	POLVO	FaP	x
POSSUIR	POSEER	FaP	xx
POVO	PUEBLO	FaP	x
RAIVA	RABIA	FaP	xx
recuar	recular	FaP	x
reter	retener	FaP	x

Vermessung des romanischen Kernwortschatzes

PG	SP	PROFIL	SFÄLLE
saudar	SALUDAR	FaP	x
SAÚDE	SALUD	FaP	x
SEIO	seno	FaP	x
SÓ	SÓLO	FaP	x
SOM	SONIDO	FaP	x
SOMENTE	solamente	FaP	x
TOM	TONO	FaP	x
VIR	VENIR	FaP	x

Abbildung 56: Minimale Formabweichungen von Portugiesisch

Die Tabelle zeigt, dass das Material der FaX-Kohorten in Übungen nicht fehlen sollte. Übungen sind lernwirksamer, wenn die zu internalisierende Regel gehäuft durch entsprechende Formen präsent und sichtbar ist. Die KRM-Architektur erlaubt eine rasche Zusammenstellung entsprechender Fälle zu FaF, FaL, FaP oder FaS.

Die iberischen Sprachen liefern füreinander die zahlreichsten und augenfälligsten Transferbasen. Wo aber unterscheiden sie sich voneinander? Die Selektoren (*P*\$*S*) in PROFIL MIT (oFk) in ROS- geben in zweistufiger Suche die Antwort des KRM. Sie betrifft 1194 Adäquanzen: 340 Fälle für die (*P*\$*S*)-Selektion und 854 für die (oFk aus PG=P OR Pw MIT SP=S OR Sw)-Sortierung. Zur Veranschaulichung seien wenige Serienausschnitte genannt wie: *vassoura ~ escoba, parvo ~ tonto, logo ~ pronto, açougueiro ~ carniceiro, aquecer ~ calentar máu/má ~ mal*. Schon diese kleine Auswahl zeigt an, dass formkonvergente Adäquanzen auch paarig auftreten (*até logo~hasta luego* über den Selektor („luego *“), die in anderen Sortierungen durchaus genannt werden:

Fr	It	Pg	Sp
PUIS	POI	DEPOIS	LUEGO
CONSÉQUENT (PAR ~)	CONSEGUENZA (DI ~)	CONSEGUINTE (POR ~)	LUEGO
BIENTÔT (À ~)	PRESTO (A ~)	LOGO (ATÉ ~)	LUEGO (HASTA ~)

Abbildung 57: até logo/hasta luego

In ähnlicher Weise sind auch die 1043 RoS*2 und RoS*3-Okkurrenzen zwischen Portugiesisch und Spanisch auszumachen ("*\$PS" OR "\$PS"). Die Auflistung zeigt Lernern einmal mehr, weshalb der Erwerb von portugiesischer Lesekompetenz auf der Grundlage von Spanischkenntnissen so einfach ist. (Entsprechende Listen sagen Entsprechendes zu anderen Sprachkombinationen.)

Der Einsatzort solcher ‚Evidenzlisten‘ sind z.B. sog. Schnupperkurse, hier für die iberischen Sprachen; oder aber Übungen zur Festigung des Wissens um interlinguale Bedeutungsadäquanzen generell.

PRATA	PLATA	FI\$PS
ARRANCAR	ARRANCAR	F\$I\$PS
PARAR	PARAR	F\$I\$PS
CHEGADA	LLEGADA	FI\$PS
CHEGAR	LLEGAR	FI\$PS
SENTAR	asentar	F\$I\$PS
BASTANTE	BASTANTE	FI\$PS
ATAR	ATAR	F\$I\$PS
ESPERAR	ESPERAR	FI\$PS

Abbildung 58: Ausschnitt aus Formkonvergenzen Portugiesisch/Spanisch

arrestation	arresto	detenção	detención	FI\$PS
habillement	abbigliamento	VESTIDO	vestido	FI\$PS
VÉRIFIER	VERIFICARE	apurar	apurar	FI\$PS
olive	OLIVA	azeitona	aceituna	FI\$PS
AIDER	AIUTARE	auxiliar	auxiliar	FI\$PS
paternité	paternità	autoria	autoría	FI\$PS
PRÉCÉDER	precedere	anteceder	anteceder	FI\$PS
RETARD (EN ~)	RITARDO (IN ~)	atrasado	atrasado	FI\$PS
travail (de ~)	lavorativo	trabalho (de ~)	laboral	FP\$IS
nécessiteux	indigente	necessitado	indigente	FP\$IS
rocher	scoglio	rochedo	escollo	FP\$IS
praticabilité	praticità	praticabilidade	practicidad	FP\$IS
renforcer	potenziare	reforçar	potenciar	FP\$IS
roman	novella	romance	novela	FP\$IS
zoologiste	zoologo	zoologista	zoólogo	FP\$IS
avance	ANTICIPO	avanço	anticipo	FP\$IS

Abbildung 59: Sprachlich-paarige Selektion

Der Einsatz des Computers in Schulklassen hat sich bekanntlich bewährt. Dies führt zur Konstruktion elektronischer Übungen, die Schüler selbst gestalten. Hierzu nutzen sie ausgewählte Bereiche des KRM-Inventars.

In der Klasse können solch vergleichende Selektionen angewandt werden, um die Aufmerksamkeit auf konkrete Merkmale der Lemmata zu lenken. In Abb. 59 zur Serie fr. *préceder* (...) pg. *anteceder* (...) werden semantische Fragen angestoßen, denen Schüler nachgehen können: Bedeuten die Vorsilben *pré-* und *ante-* wirklich dasselbe? Sind sie eventuell austauschbar? Das online-Wörterbuch LEO weist den portugiesischen und spanischen Nachbarbegriffen gleichermaßen die Bedeutung ‚vorausgehen‘, ‚vorhergehen‘ zu; Davies *Corpus del Español* nennt zur Form sp. *preceder* 84 Fundstellen, u.a.: „...*debe preceder todo discurso...*“, „...*preceder al acompañamiento...*“, „... *suele preceder a las* [Nomen]; *anteceder* bringt es auf nur einen einzigen Beleg: „*¿puede anteceder la revolución costumbrista al otro tipo de revolución?*“ Portugiesisch *preceder* bringt es in Davies *Corpus do Português* auf 24 Belege: „...*preceder a satisfação...*“, „...*preceder o próprio acesso...*“, „...*a separação que deveria preceder a esse conflito...*“; für *anteceder* nennt CP 13 Fundstellen: „... *A anteceder as fotografias estão um texto de Mário Soares...*“, *Nada como uma boa refeição a anteceder o sono...*“ Auch weitere Vergleiche mithilfe z.B. portugiesisch-französischer und portugiesisch-englischer Wörterbücher disambiguieren nicht hinreichend, sondern bestätigen die sehr weitgehende Bedeutungsgleichheit von pg. *anteceder* und *preceder*. Offensichtlich ist *preceder* im portugiesischen Wortschatz zentraler positioniert als *anteceder*; im spanischen Lexikon ohnehin. Zum Englischen und Französischen begegnen Lerner den Adäquanzen *to precede*, *to come before*, *to antecede* und *préceder*. Fazit: Zur produktiven Verwendung werden Lerner die Formadäquanzen von *preceder* vorziehen; sie haben gelernt, dass *anteceder* viel seltener vorkommt und gegenüber *preceder* nur unscharf abgrenzt. Die so aktivierte Sprachenbewusstheit sollte zu Fragen an sprachkompetente Personen veranlassen. Und last but not least: Ihre Aufmerksamkeit für die Konkurrenz von *proced-* und *anteced-*Formen ist geweckt.

Mentale Prozesse wie die hier geschilderten befördern Sprachenbewusstheit, ohne die eine wirksame Fehlerprophylaxe nicht möglich ist. Hierzu gehört auch die Kompetenz, Wörterbücher und Konkordanzen für den eigenen Spracherwerb einsetzen zu können. Wie man sieht, sollte die Nutzung des KRM nicht unbedingt ‚für sich bleiben‘, sondern im Verein mit anderen Lehrermitteln eingesetzt werden. Paarige Formkongruenzen laden zu solchen Vergleichen geradezu ein. So sind die KRM-Apps auch ein Mittel, die Nutzung anderer lexikodidaktisch sinnvoller Korpora zu unterstützen. Dies alles hat viel mit der Fähigkeit des autonomen Lernens zu tun. Dass dies für viele, hier nicht erwähnten Selektionen gilt, verdient keine eigene Erwähnung.

Mögen die KRM-Datenbank und ihre Lernapps ihrem Auftrag gemäß dazu beitragen, den Weg in eine zumindest rezeptive europäische Mehrsprachigkeit zu verkürzen, indem sie zur rechten Auswahl lexikalischen Materials veranlassen. In diesem Sinne bleibt zu hoffen, dass die vorliegende Studie zu mehr Bewusstheit für Fragen der Integration des lexikalischen Inputs – seinen Formen und ihres dia-frequentiellen Status in den romanischen Zielsprachen – in Lehrmaterialien im Zusammenhang mit (nicht nur) der romanischen Mehrsprachigkeit beiträgt.

Literatur

- Abel, Fritz (1971): Die Vermittlung passiver Spanisch- und Italienischkenntnisse im Rahmen des Französischunterrichts. In: *Die Neueren Sprachen* 70, 355-359.
- Abel, Fritz (2006): Gemeinsamkeiten im Häufigkeitswortschatz des Französischen, Italienischen und Spanischen. Bericht über ein Experiment. In: Martinez et al. (Hrsg.), 37-52.
- Aitchison, Jean (1997): *Wörter im Kopf. Eine Einführung in das mentale Lexikon*. Tübingen: Niemeyer (aus dem Englischen [1994] von Martina Wiese).
- Androulakis, Georges. Beckmann, Christine. Blondin, Christiane. Fagnant, Annick. Goffin, Chistelle. Górecka, Joanna. Martin, Romain. Mattar, Cathérine. Meißner, Franz-Joseph. Mosorka, Edyta. Nowicka, Agnieszka. Pelt, Monika. Schröder-Sura, Anna. Skrivánek, Hanna. Wilczyńska, Weronika & Wojciechowska, Bernadeta: *Pour le multilinguisme: Exploiter à l'école la diversité des contextes européens. Résultats d'une étude internationale*. Liège: Editions de l'Université de Liège. <http://www1.uni-giessen.de/rom-didaktik/Multilingualism/download/facette1%20integral.pdf>
- Ausubel, David (1968): *Educational Psychology. A Cognitive View*. New York: Harcourt Brace & World.
- Bandura, Albert (1997): *Self-Efficacy. The exercise of control*. New York: Freeman.
- Baqué, Lorraine & Estrada, Marta (1999): L'intercompréhension orale entre langues romanes. In: Martínez, Sierra. González, Fermín & Hernández, Carmen (eds.): *Las lenguas en la Europa Comunitaria III. Diálogos Hispánicos* 23, Amsterdam: Rodopi, 269-282.
- Bär, Marcus (2009): *Förderung von Mehrsprachigkeit und Lernkompetenz. Fall studien zu Interkomprehensionsunterricht mit Schülern der Klassen 8 bis 10*. Tübingen: Narr.
- Bär, Marcus (2010): Motivation durch Interkomprehensionsunterricht empirisch geprüft. In: Doyé & Meißner (Hrsg.), 281-290.
- Berthele, Raphael. Colliander, Peter. Duke, Janet. Hufeisen, Britta. Lutjeharms, Madeline. Marx, Nicole. Möller, Robert & Zeevaert, Ludger (2011): Zu den Grenzen des EuroCom-Konzeptes für EuroComGerm - Zwischenfazit. In: Pöckl, Wolfgang. Ohnheiser, Ingeborg & Sandrini, Peter (Hrsg.): *Translation, Sprachvariation, Mehrsprachigkeit. Festschrift für Lew Zybatow zum 60. Geburtstag*. Frankfurt a.M., 483-598.
- Braun, Peter. Schaeder, Burkhard & Volmert, Johannes (Hrsg.) (2003): *Internationalismen. Studien zur interlingualen Lexikologie und Lexikographie II*. Tübingen: Niemeyer.
- Carver, Ronald P. (1994): Percentages of unknown words in text as a function of the relative difficulty of the text: Implications for instruction. In: *Journal of Reading Behavior* 26, 413-437.

- Christ, Herbert (1980): Ein Vorschlag zur Veränderung der Schulsprachenpolitik: Die Homburger Empfehlungen. In: *Neusprachliche Mitteilungen* 33, 97-99.
- Dauzat, Albert. Dubois, Jean & Mitterand, Henri (1971): *Nouveau dictionnaire étymologique. (2^e édition revue et corrigée)*. Paris: Larousse.
- Davies, Mark (2006): *A Frequency Dictionary of Spanish. Core vocabulary for learners*. New York & London: Routledge.
- Davies, Mark & Gardner, Dee (2010): *Frequency Dictionary of Contemporary American English. Word sketches, collocates, and thematic lists*. New York: Routledge.
- Davies, Mark (o.J.): *O Corpus do Português*.
<http://www.corpusdoportugues.org/>
- Davies, Mark (o.J.): *El Corpus del Español*.
<http://www.corpusdelespanol.org/>
- De Mauro, Tullio. Mancini, Federico. Vedovelli, Massimo & Voghera, Miriam (1993): *Lessico de frequenza dell'italiano parlato*. Prefazione di Pierluigi Ridolfi. s.l.: ETASLIBRI.
- Doyé, Peter (2009): *Didaktik der bilingualen Vorschulerziehung. Dargestellt am Beispiel der schulischen Einrichtungen in Berlin und Wolfsburg*. Tübingen: Narr.
- Doyé, Peter & Meißner, Franz-Joseph (Hrsg.) (2010): *Lernerautonomie durch Interkomprehension/Promoting Learner Autonomy Through Intercomprehension/L'autonomisation de l'apprenant par l'intercompréhension*. Tübingen: Narr.
- Eaton, Helen S. (1940): *An English-German-French-Spanish Word Frequency Dictionary. A Correlation of the first six thousand words in four single language frequency lists*. New York: Dover Publications.
- Fäcke, Christiane & Meißner, Franz-Joseph (Hrsg.) (2019): *Handbuch Mehrsprachigkeits- und Mehrkulturalitätsdidaktik*. Tübingen: Narr.
- Freund, Stefan. Schröttel, Wolfram & Thier, J. (2003): *Die 1000 häufigsten Wörter im Query-Corpus*. In:
http://www.carolinum.net/files/caronet/files/redaktion_all/Fachschaften/Latein/Materialien/Query%20Tabelle.pdf
- Gnutzmann, Claus (2005): Globalisation, plurilingualism and English as a Lingua Franca (ELF). Has English as a Foreign Language become obsolete? In: Meißner, Franz-Joseph (Koord.): »Neokommunikativer« Fremdsprachenunterricht. In: *Fremdsprachen Lehren und Lernen* 34, 15-26.
- Grimm, Wilhelm & Grimm, Jacob (1952): *Deutsches Wörterbuch*. (Nachdruck): München: dtv.
- Haag, Ludwig & Stern, Elsbeth (2003): In search of the benefits of learning Latin. In: *Journal of Educational Psychology* 95, 174-178.
- Hu, Hsueh-Chao Marcella & Nation, Paul (2000) : Unknown vocabulary density and reading comprehension. In: *Reading in a Foreign Language* 13, 403-430.

- Hufeisen, Britta (2006): DaFnE, EuroComGerm, EaG Forschungsbeiträge für die Entwicklung eines allgemeinen und doch (noch) idealtypischen Gesamtsprachencurriculum. In: Martinez et al. (Hrsg.), 111-123.
- Hufeisen, Britta (2008): Gesamtsprachencurriculum, curriculare Mehrsprachigkeit und Mehrsprachigkeitsdidaktik-Utopie – Allheilmittel oder Schreckgespenst aller AnglistInnen und EnglischlehrerInnen? In: Bausch, Karl-Richard. Burwitz-Melzer, Eva. Königs, Frank G. & Krumm, Hans-Jürgen (Hrsg.): *Fremdsprachenlernen erforschen: sprachspezifisch oder sprachenübergreifend? Arbeitspapiere der 28. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts*. Tübingen: Narr, 97-106.
- Hufeisen, Britta & Lutjeharms, Madeline (Hrsg.) (2005): *Gesamtsprachencurriculum, Integrierte Sprachendidaktik, Common Curriculum. Theoretische Überlegungen und Beispiele der Umsetzung*. Tübingen: Narr.
- Jakisch, Jenny (2015): *Mehrsprachigkeit und Englischunterricht. Fachdidaktische Perspektiven, schulpraktische Sichtweisen*. Frankfurt a.M.: Lang.
- Jamet, Marie-Christine (2007): *A l'écoute du français. La compréhension orale dans le cadre de l'intercompréhension des langues romanes*. Tübingen: Narr.
- Jessner, Ulrike. Török, Valentina & Pellegrini Claudia (2018): The role of Latin in multilingual learners' strategy. In: *Language Education and Multilingualism. The Langscape Journal* 1/2018, 85-102.
- Jones, Randall L. & Tschirner, Erwin (2006): *Frequency Dictionary of German. Core vocabulary for learners*. New York: Routledge.
- Klein, Horst G. (2008): L'anglais, base possible de l'intercompréhension romane. In: *Revue de didactologie des langues-cultures et de lexiculurologie* 149, 119-128.
- Klein, Horst G. (2002): Das Französische: die optimale Brücke zum Leseverstehen romanischer Sprachen. In: *französisch heute* 33, 34-46.
- Klein, Horst G. & Reissner, Christina (2006): *Basismodul Englisch. Englisch als Brückensprache in der romanischen Interkomprehension*. Aachen: Shaker.
- Klein, Horst G. & Stegmann, Tilbert D. (1999): *EurocomRom. Die sieben Siebe. Romanische Sprachen sofort lesen können*. Aachen: Shaker.
- Krashen, Stephen D. (1995): The Input Hypothesis: Some Current Issues. In: Bruschi, Wilfried & Stiller, Hugo (Hrsg.): *Lust auf Sprachen: Schlüssel zu Europa – Tor zur Welt. Beiträge zum Internationalen Fremdsprachenkongress in Hamburg 1994*. Hamburg: Petersen, 15-30.
- Kühn, Peter (1990): Das Grundwortschatzwörterbuch. In: Hausmann, Franz Josef. Reichmann, Oskar. Wiegand, Herbert Ernst & Zgusta, Ladislav (Hrsg.): *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Berlin & New York: Gruyter, 1353-1364.
- Laufer, Batia (1997): The lexical plight in second language reading. Words you don't know, words you think you know, and words you can't guess. In: Coady, James &

- Huckin, Thomas (eds.): *Second Language Vocabulary Acquisition*. Cambridge: C.U.P., 20-34.
- Le Besnerais, Martine (2011): Les langues romanes sont-elles un atout pour l'apprentissage de l'anglais ? In: Meißner et al. (eds.), 35-48.
- Lonsdale, Deryle & Le Bras, Yvon (2009): *Frequency Dictionary of French. Core Vocabulary for Learners*. New York: Routledge.
- Lutjeharms, Madeline (2002a): Lesestrategien und Interkomprehension in Sprachfamilien. In: Kischel, Gerhard (Hrsg.): *EuroCom - Mehrsprachiges Europa durch Interkomprehension in Sprachfamilien. Tagungsband des internationalen Fachkongresses zum Europäischen Jahr der Sprachen 2001. Hagen, 9.-10. November 2001*. Aachen: Shaker, 124-140.
- Lutjeharms, Madeline (2002b): Wortrepräsentation bei Mehrsprachigen und Wortschatzerwerb in einer Fremdsprache. In: Hufeisen, Britta & Marx, Nicole (Hrsg.): *Beim Schwedischlernen sind Englisch und Deutsch ganz hilfreich. L3-Spezifika und Forschungen zum multiplen Spracherwerb im Jahre 2002*. Frankfurt: Peter Lang, 47-62.
- Lutjeharms, Madeline (2007): Processing levels in foreign-language reading. In: ten Thije, Jan D. & Zeevaert, Ludger (eds.): *Receptive Multilingualism*. Amsterdam: John Benjamins, 265-284.
- Martinez, Hélène. Reinfried, Marcus & Bär, Marcus (Hrsg.) (2006): *Mehrsprachigkeitsdidaktik gestern, heute und morgen. Festschrift für Franz-Joseph Meißner*. Tübingen: Narr.
- Martinez, Hélène (2010): Plurilingüismo, intercomprensión y autonomización: el papel de la tercera lengua en el desarrollo de la autonomía. In: Doyé & Meißner (eds.), 146-160.
- Marx, Nicole (2010): eag and Multilingualism Pedagogy. An Empirical Study of Students' Learning Processes on the Internet Platform English after German. In: Doyé & Meißner (Hrsg.), 225-236.
- Meißner, Franz-Joseph (1978). *Wortgeschichtliche Untersuchungen im Umfeld von französisch enthousiasme und génie*. Genf: Droz.
- Meißner, Franz-Joseph (1989): Grundwortschatz und Sprachenfolge. Eine statistische Quantifizierung zum lexikalischen Transfer: Französisch/Englisch – Englisch/Französisch, Spanisch, Italienisch. In: *französisch heute* 28, 377-387.
- Meißner, Franz-Joseph (1993a): Interlexis – ein europäisches Register und die Mehrsprachigkeitsdidaktik. In: *Die Neueren Sprachen* 92, 532-554.
- Meißner, Franz-Joseph (1993b): Anmerkungen zur Interlexikologie aus romanistischer Sicht. In: *Muttersprache* 103, 113-130.
- Meißner, Franz-Joseph (1996): Eurolexis und Fremdsprachendidaktik. In: Munske & Kirkness (Hrsg.), 284-305).
- Meißner, Franz-Joseph (1999): Das mentale Lexikon aus der Sicht der Mehrsprachigkeitsdidaktik. In: *Grenzgänge* 6, H. 12, 62-80.

- Meißner, Franz-Joseph (2000): Zwischensprachliche Netzwerke. Mehrsprachigkeitsdidaktische Überlegungen zur Wortschatzarbeit. In: *französisch heute* 31, 55-67.
- Meißner, Franz-Joseph (2003): Altsprachlicher Unterricht und Fremdsprachenunterricht. In: Bausch, Karl-Richard. Christ, Herbert & Krumm, Hans-Jürgen (Hrsg.): *Handbuch Fremdsprachenunterricht*. (4. völlig neu bearbeitete Auflage). Tübingen und Basel: A. Francke Verlag, 151-157.
- Meißner, Franz-Joseph (2010): Empirisch geprüft: Interkomprehensionskompetenz, Vorwissen, Spracherlebnis, Lernerautonomisierung. In: Doyé & Meißner (Hrsg.), 193-224.
- Meißner, Franz-Joseph (2012): A la recherche des origines d'une didactique du plurilinguisme. In: Colombat, Bernard. Fournier, Jean-Marie & Raby, Valérie (coord.): *Vers une histoire générale de la grammaire française. Matériaux et perspectives. Actes du colloque international de Paris (HTL/SHEL 27-29/1/2011)*. Paris: Champion, 533-550.
- Meißner, Franz-Joseph (2014): *En fildelare i Sala döms av tingsrätten att betala 4,3 miljoner kronor...* Sprachtransfer, Lerntransfer, Intakeoptimierung beim ersten Leseversuch in Schwedisch. In: Hoffmann, Sabine & Stork, Antje (Hrsg.): *Lernorientierte Fremdsprachenforschung und –didaktik. Festschrift für Frank G. Königs*. Tübingen: Narr, 71-84.
- Meißner, Franz-Joseph (2016): *Der Kernwortschatz der romanischen Mehrsprachigkeit (KRM). Didaktische, lexikologische, lexikographische Überlegungen zu Erstellung, Präsentation, Anwendungen einer elektronischen Mehrsprachigkeitswortliste und von Lernapps zur romanischen Mehrsprachigkeit. (GiF:on 7)* (http://geb.uni-giessen.de/geb/volltexte/2016/11950/pdf/GiFon_7.pdf).
- Meißner, Franz-Joseph (2018): The 'Core Vocabulary of Romance Plurilingualism' (the CVRP-Project). In: Ambrosch-Baroua, Tina & Kropp, Amina & Müller-Lancé, Johannes (Hrsg.): *Mehrsprachigkeit und Ökonomie*. München: Open Access LMU, 91-106. (<https://epub.ub.uni-muenchen.de/40299/>).
- Meißner, Franz-Joseph (2019): Mehrsprachigkeitsdidaktik. In: Fäcke & Meißner (Hrsg.), (im Druck).
- Meißner, Franz-Joseph. Beckmann, Christine & Schröder-Sura, Anna (2008): *Mehrsprachigkeit fördern. Vielfalt und Reichtum in der Schule nutzen. Zwei deutsche Stichproben einer internationalen Studie zu den Klassen 5 und 9 zu Sprachen und Fremdsprachenunterricht*. Tübingen: Narr. (Auch: <http://www1.uni-giessen.de/rom-didaktik/Multilingualism/download/facette1%20deutschland.pdf>)
- Meißner, Franz-Joseph & Burk, Heike (2001): Hörverstehen in einer unbekannt romanischen Fremdsprache und methodische Implikationen für den Tertiärsprachenerwerb. In: *Zeitschrift für Fremdsprachenforschung* 12 (1), 63-102.
- Meißner, Franz-Joseph. Capucho, Filomena. Degache, Christian. Martins, Adriana, Spiță (eds.) (2011): *Intercomprehension. Learning – teaching – research/*

- Apprentissage – enseignement – recherche/Lernen – Lehren – Forschung.*
Tübingen: Narr.
- Meyer-Lübke, Wilhelm (1911): *Romanisches etymologisches Wörterbuch.* Heidelberg: Carl Winter's.
- Muljačić, Žarko (1979): Per un dizionario storico degli europeismi (DSE). In: Mair, Walter N. & Sallager, Edgar (Hrsg.): *Sprachtheorie und Sprachpraxis Henri Vernay zu seinem 60. Geburtstag*, Tübingen: Narr, 279-286.
- Müller-Lancé, Johannes (2002): La corrélation entre la ressemblance morphologique des mots et la probabilité du transfert interlinguistique. In: Kischel, Gerhard (Hrsg.), *EuroCom - Mehrsprachiges Europa durch Interkomprehension in Sprachfamilien. Tagungsband des Internationalen Fachkongresses, Hagen, 9.-10. November 2001.* Hagen: FernUniversität/Gesamthochschule (auch unter Aachen: Shaker), 141-159.
- Müller-Lancé, Johannes (2003): *Der Wortschatz romanischer Sprachen im Tertiärsprachenerwerb. Lernerstrategien am Beispiel des spanischen, Italienischen und Katalanischen.* Tübingen: Stauffenberg.
- Munske, Horst Haider & Kirkness, Alan (Hrsg.) (1996): *Eurolatein. Das griechische und lateinische Erbe in den europäischen Sprachen.* Tübingen: Niemeyer.
- Murillo Puyal, Julio (2011): Contribución al estudio de la intercomprensión oral en la comunicación exolingüe: la percepción por catalanohablantes y francófonos de la modalidad interrogativa en español. In: Meißner et al. (eds.), 204-218,
- Neveling, Christiane (2004): *Wörterlernen mit Wörternetzen. Eine Untersuchung zu Wörternetzen als Lernstrategie und als Forschungsverfahren.* Tübingen: Narr.
- Neveling, Christine (2006): Leichter Französisch lernen durch Latein? In: *französisch heute* 37, 36-47.
- Ortner, Tuulia M. (2011): Der Latein-Effekt. Schult Lateinunterricht die kognitiven Fähigkeiten? In: *Pegasus-Online Zeitschrift XI*, 69-81.
- Phillipson, Robert & Skutnabb-Kangas, Tove (1999): Englishisation: one dimension of globalisation. In: *AILA-Review* 13, 19-36.
- Perkuhn, Rainer. Keibel, Holger & Kupietz, Marc (2012): *Korpuslinguistik.* Tübingen: Narr-Attempto.
- Price, Glanville (1988): *Die französische Sprache. Von den Anfängen bis zur Gegenwart.* Tübingen: UTB.
- Quian, David (2002): Investigating the relationship between vocabulary knowledge and academic reading performance. In: *Language Learning* 52, 513-536.
- Reinfried, Marcus (1999): Innerromanischer Sprachtransfer. In: *Grenzgänge. Beiträge zu einer modernen Romanistik* 6, H. 12, 96-125.
- Schulz, Hans & Basler, Otto (1973-1983): *Deutsches Fremdwörterbuch*, begründet von H. Schulz, fortgeführt von O. Basler, weitergeführt im Institut für Deutsche Sprache, Berlin/New York, 6 Bde.

- Siepmann, Dirk (2006): Was der Fremdsprachenunterricht von der Korpuslinguistik erwarten darf. In: Jung, Udo O.H. (Hrsg.): *Praktische Handreichung für Fremdsprachenlehrer*. In Zusammenarbeit mit Heidrun Jung. 4. vollständig neu bearbeitete Auflage. Frankfurt a.M.: Peter Lang, 321-330.
- Swanborn, M. & de Glopper, K. (1999): Incidental word learning while reading. A meta-analysis. In: *Review of Educational Research* 69, 261-285.
- Tafel, Karin. Durić, Rašid. Lemmen, Radka. Olshevska, Anna & Przyborowska-Stolz, Agata (2009): *Slavische Interkomprehension - eine Einführung*. Tübingen: Narr & Francke Attempto.
- Thorndike, Edward L. & Woodworth, Robert S. (1901): The influence of improvement in one mental function upon the efficiency of other functions. In: *Psychological Review* 8: 247-261, 384-395, 553-564.
auch: <http://psychclassics.yorku.ca/Thorndike/Transfer/transfer1.htm>
- Tschirner, Erwin (2005): Korpora, Häufigkeitslisten, Wortschatzerwerb. In: Heine, Antje. Henning, Mathilde & Tschirner, Erwin (Hrsg.): *Deutsch als Fremdsprache. Konturen und Perspektiven eines Faches. Festschrift für Barbara Wotjak zum 65. Geburtstag*. München: iudicium, 133-152.
- Wandruszka, Mario (1979): *Die Mehrsprachigkeit des Menschen*. München: Piper.
- Wenzel, Veronika (Hrsg.) (2014): *Fachdidaktik Niederländisch*. Münster: LIT.
- Werner, Angelika (2003): Internationalismen im Japanischen. Oder: Ein Computer ist ein computer ist ein konpyu-ta. In: Braun et al. (Hrsg.), 183-201.
- Zimmer, Rudolf (1990): *Äquivalenzen zwischen Deutsch und Französisch. Theorie - Korpus - Indizes. Ein Kontextwörterbuch*. Tübingen: Niemeyer.
- Zöfgen, Ekkehard (2006): Wortschatzdidaktik in chresto-lexikographischem Gewand. In: *französisch heute* 37, 164-179.
- Zöfgen, Ekkehard (1994): *Lernerwörterbücher in Theorie und Praxis. Ein Beitrag zur Metalexikographie mit besonderer Berücksichtigung des Französischen*. Tübingen: Niemeyer.

Index

- \$-Markierung** 24, 33, 72, 75
 ‚Evidenzlisten‘ 127
 <>-Selektion 68
 Abweichende Zählungseffekte 18
 Adressaten 12
 Audit 106
 Augenscheinliche Formdivergenz 27
Bamberger Query Corpus 48
 Bedeutung 105
 Bedeutungsadäquanz 10, 11, 19, 28,
 40, 82, 102
 bidirektional 16
 Bildungswortschatz 10, 20, 45, 49
 Brückensprache 16, 20
 Buchwort 123
 Copyright 12, 107
Core Vocabulary 106
 Dispersion 14
 Disponibilität 14
 eag 39, 133
 einfache Selektion Lx 16
 Englisch 102
 Erbwörter 112
 erweiterte Selektion (X OR Xw) 16
Etyma 7, 9, 24, 26, 27, 30, 31, 32, 48,
 110
 FaX-Markierung 76, 77
Formdivergenz 9, 17
Formkongruenz 9, 17, 22
 Französisch 113
 Frequenz 14
 Frequenzrang 11, 19, 27, 28, 36, 40,
 52, 54, 55, 105, 106, 109
fuzziness 106
 gebrochene Serien 72
Gesamt 78
 Gesamtsprachencurriculums 108
 Grundmenge 16, 78
 Grundmengen 16, 85
 Grundwortschatz 27, 28, 36, 52, 56,
 133
 Interligalexe 10, 17, 111
 interlinguale Transparenz 105
 Kernbedeutungen 104
 Kernwortschatz
 europäischer 116
 Klett’sche Grundwortschatzlisten 27
 Konvention der Komputation 17
 ko-textliche Bedeutung 104
 lexikalisches Wissen Erwachsener in
 der Muttersprache 99
 Ligalexe 17
 Majuskel 27
 Mehrsprachenerwerbsforschung 108
 Mehrsprachenlexikographie
 didaktische 14
 Mehrsprachigkeit 106
 über die Romania hinausreichende
 104
 Minimalwortschätze 37, 105
 mehrsprachige 15
 MIT-Grundmenge 94
 MIT-Gruppen 60
 MIT-Ziehungen 93
 Motivation 84
 nach Sprachen getrennte Zählung 21
 negativer Transfer 108
 Opakenanteil 64
 Opazität 105
 Operatoren 86
 Originalität des KRM 107
 Polymorphie 11
 Polysemie 11, 16, 57, 58, 110
 positivem Transfer 108
 präpädagogische Analyse 14
 primäre Transfersprache 16
 Profilanteil 64
 Profilwörter 63

- qualitative Verteilung 60
- Reduplikationen 57
- regelmäßige morphologische
Korrespondenzen 120
- RoS*-X 66
- sekundäre Zielsprache 20
- Selbstwirksamkeit 84
- Serialisierung 9, 11, 22
- Serie** 22
- Sprachenbewusstheit 104
- Sprachenfolge 14, 37, 39, 107
- Sprachkennung 78
- Sprachkombinationen 21, 24, 77,
118
- Sprachlernkompetenz 108
- statistisch-irrelevante Daten 82
- Suchroutinen 115
- Symbolfarben 64
- T4-T3-Transparenz 84
- Transfer 20, 28, 44, 49, 77, 78, 133,
136
- Transferabilisierbarkeit 46
- Transferbasis 20, 44
- Transferpotential 60
- Transferprozesse
bidirektional 81
- Transfers* 107
- Transparenzstufe 28
- Übungsformate 115
- Umfang des transferablen Vokabulars
99
- unabhängige Variable 62, 86
- Wortähnlichkeiten 17
- Wortschatz
europäischer 103, 111
- Zielsprache* 16
- Zuverlässigkeit der Daten 106

Franz-Joseph Meißner, Emeritus im Fach Didaktik der romanischen Sprachen an der Gießener Justus Liebig-Universität, ist einer der führenden Vertreter auf dem Gebiet der romanistischen Mehrsprachigkeitsdidaktik. Er ist Autor zahlreicher Publikationen – Monographien, Sammelbände und Aufsätze zur romanischen Interkomprehension. Hauptsächliche Schwerpunkte seiner Forschungen betreffen: Internationalismen, didaktische Lexikographie des Französischen und der romanischen Mehrsprachigkeit, Entwicklung des interkomprehensives Ansatzes, quantitative und qualitative empirische Lern- und Lernerforschung Fremdsprachen, Hörverstehen, Aufgabenkonstruktion, Sprachenpolitik.

F.-J.M. war, z.T. leitend, Mitglied zahlreicher nationaler und internationaler Projekte; darunter der *Referenzrahmen für Plurale Ansätze zu Sprachen und Kulturen* (Graz/Europarat), die wissenschaftliche Beratung für Französisch am Berliner Instituts für Qualitätsentwicklung im Bildungswesen (IQB), *Pour le multilinguisme : Exploiter à l'école la diversité des contextes européens* (mit Unterstützung der Europäischen Kommission), *EuroComDidact: empirische Grundlagen zur Entwicklung einer auf Interkomprehension beruhenden Methode zum Erwerb von Mehrsprachigkeit* und weiterer.

Seit seiner Emeritierung leitet F.-J.M. das Projekt *Kernwortschatz der romanischen Mehrsprachigkeit*. Es umfasst die Entwicklung und den Aufbau einer auf der Computerlexikographie beruhenden Datenbank zu den Zielsprachen Französisch, Italienisch, Portugiesisch und Spanisch sowie Deutsch und Englisch als disambiguierende Sprachen.

Korrespondenzadresse: franz-joseph.meissner@sprachen.uni-giessen.de

Homepage:

<http://www.uni-giessen.de/meissner>

<http://www.eurocomdidact.eu>

Der Band beschreibt die Vermessung des *Kernwortschatzes der romanischen Mehrsprachigkeit* (KRM) mit Blick auf zwischensprachliche Transparenz und Opazität der Lemmata sowie ihrer Frequenzränge. Er schließt an GiF:on 7 an, wo Grundfragen der Erstellung mehrsprachlicher lexikalischer Minimalwortschätze (Frequenzrang <5000) und ihrer Verarbeitung in elektronischen Lernapps behandelt werden.

Da das Inventar des KRM mit einem Copyright belegt ist und sich die zahlreichen Vorteile einer solchen Datenbank nur im Rahmen einer elektronischen Architektur abrufen lassen, behandelt dieses Buch Grundfragen, die sowohl die didaktische Lexikographie der Mehrsprachigkeit als auch die Minimalkorpora betreffen.

Adressaten sind neben Lexikographen und Fremdsprachendidaktiker durchaus auch Lehrerinnen und Lehrer romanischer Sprachen, des Englischen und des Deutschen, die über Fragen des lexikalischen Inputs nachdenken.