Morphologische Produktivität in deutschen literarischen Texten ausländischer Schriftsteller

Das Beispiel der Ableitungen auf *-ung*

Masterarbeit
zur Erlangung des akademischen Grades Master of Arts (M.A.)
im Fach Linguistik

Humboldt-Universität zu Berlin
Philosophische Fakultät II
Institut für deutsche Sprache und Linguistik

eingereicht von: Marta Lupica Spagnolo
geb. am 20.12.1985
in Pavia (IT)

Wissenschaftliche Betreuerin: Prof. Dr. Anke Lüdeling
Zweitgutachter: Prof. Dr. Manfred Krifka

Berlin, den 21. November 2011

Inhaltverzeichnis

1. Einleitung 6

1.1 Interkulturelle Literatur: eine Definition 7

1.2 Die Geschichte der interkulturellen Literatur 8

1.3 Die Sprache der interkulturellen Autoren 12

1.4 Die Hypothese 13

2. Die erste Unterhypothese: Anwesenheit gemeinsamer linguistischen Merkmalen 16

2.1 Die Lernersprache 16

2.2 Die soziale und soziolinguistische Fossilisation 18

2.3 Lernersprachen und Varietäten 22

2.4 Zwischenfazit 25

3. Die zweite Unterhypothese: Merkmal und Möglichkeit einer Asymmetrie im linguistischen Verhalten von interkulturellen und deutschen Schriftstellern 27

3.1 Die Wortbildungsmuster 28

3.2 Anlässe und Beschränkungen der Wortbildung als Bedarfsystem 30

3.3 Die Produktivität als variables Maß 33

3.3.1 Die Baayen'sche Produktivität und die probabilistische Morphologie 35

3.3.2 Die drei Maße der Produktivität 38

3.3.3 Die soziolinguistische Variation der Produktivität 40

3.4 Die Produktivität und die Lernersprache: Erwerb der Wortbildungskompetenz 41

4. Die dritte Unterhypothese: Gebrauchsunterschied und dessen Untersuchungsmethode . 44

4.1 Die Variable und die Korpuslinguistik 44

4.2 Die Lernerkorpora 46

4.3 Korrelativer und funktionaler Ansatz: Begriff von Stil 48

4.3.1 Der Stil in der Stilometrie 50

4.3.2 Der Stil in der Stilistik 51

4.4 Methodische Zusammenfassung 53

5. Empirischer Teil ... 56
5.1 Der Korpusaufbau ... 57
5.1.1 Das Chamisso-Korpus ... 57
5.1.2 Datenerhebungstechnik ... 59
5.1.3 Metadaten ... 61
5.1.4 Vergleichskorpus ... 63
5.2 Die empirische Untersuchung ... 64
5.2.1 Die deskriptive Analyse: Auswahl und Zählung der Belege ... 64
5.2.2 Analytische Analysis I: die Auswertung der Belege nach der Variable *Nativeness* ... 67
5.2.3 Analytische Analysis II: die anderen unabhängigen Variable ... 69
5.2.3.1 Die Variable Auswanderungsgrund ... 69
5.2.3.2 Die Variable Muttersprache ... 71
5.2.3.3 Die Variable Autor ... 72
5.2.3.4 Outsider ... 73
5.2.4 Analytische Analyse III: Gebrauchsvariationen anderer Wortbildungsmuster nach der Variable *Nativeness* ... 73
5.3 Diskussion der Ergebnisse ... 75

6. Fazit und Ausblick ... 78

7. Quellenverzeichnis ... 80
7.1 Literatur ... 80
7.2 Internet-Link ... 88

8. Anhänge ... 89

9. Selbstständigkeitserklärung ... 95

Abbildungsverzeichnis

Abbildung 1: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *Nativeness* und Berechnung der *realized* und *potential productivity*. 68
Abbildung 2: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *Auswanderungsgrund*. 70
Abbildung 3: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *L1*. 71
Abbildung 4: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *Autor*. 72
Abbildung 5: Normalisierte Verteilung der Belege der morphologischen Kategorie *-keit/-igkeit/ -heit*, *-bar*, *ver-* nach der Variable *Nativeness* und Berechnung der *realized* und *potential productivity*. 74

Abkürzungen

ANNIS2: *ANNotation of Information Structure*
C: morphologische Kategorie
CEA: *Computer-aided Error Analysis*
CIA: *Contrastive Interlanguage Analysis*
CL: *Contrastive Linguistics* (kontrastive Linguistik)
COSMASII: *COrpus Search, Management and Analysis System*
DaF: Deutsch als Fremdsprache
DEREKO: DEutsches REferenzKOrpus
EA: *Error Analysis* (Fehleranalyse)
ESF: *The European Science Foundation Project*
H0: Null-Hypothese
H1: Alternativhypothese
IL: *Interlanguage* (Lernersprache)
L1: Erst- bzw. Muttersprache
L2: Zweit- bzw. Fremdsprache
LAD: *Language Acquisition Device*
N: Token
NL: native Sprache bzw. Muttersprache
NNV: nicht-native Varietät
prop.test: *Test of Equal or Given Proportions*
p-Wert: *probability-value* (Signifikanzwert)
SLA: *Second Language Acquisition* (Zweitspracherwerb)
TL: Targetsprache
V: Typ

*: ungrammatisch
?: nicht vollkommen wohlgeformt bzw. nicht vollkommen akzeptabel

Ich musste mich in Geduld fassen, aufs Neue beginnen; die ödesten italienische Texte, die keine Spur von Tragischem zeigten, musste ich lesen und verdauen, um mich mit toskanischen Wendungen einzudecken – wäre der Ausdruck nicht so albern, so würde ich sagen: ich musste mich den ganzen Tag lang *entdenken*, um dann *wieder zu denken*.

Vittorio Alfieri, *Mein Leben*

1. Einleitung

Die Abgrenzung einer „interkulturellen Literatur“ begründen einige Kritiker auf der sprachlichen Ebene, insofern sie das Schreiben in einer Fremdsprache als Anlass zu einem kreativen Umgang mit derselben sehen. In der linguistischen Forschung stellt man darüber hinaus fest, dass produktive Wortbildungsstrategien von Lernern einer Fremdsprache (L2-Lerner) angewendet werden, um Löcher im lernersprachlichen Wortschatz zu füllen. Das pragmatische Bedürfnis, die Abwesenheit bestimmter Beschränkungen, sowie eine unterschiedliche Zusammensetzung des mentalen Lexikons verursachen diesbezüglich bei L2-Lernern die Generalisierung morphologischer Strukturen. Ein Übergebrauch einiger produktiver Wortbildungsmuster könnte daher in Bezug auf die linguistische Kreativität der „interkulturellen Autoren“ einen Aspekt bilden, der unmittelbar mit dem Sprachwechsel zur deutschen Sprache zusammenhängt.

Durch eine korpuslinguistische Analyse untersucht die vorliegende Studie Variationen in der Produktivität der Wortbildungsmuster auf *-ung*, und zwar im Vergleich von Autoren, die in Deutsch als ihrer Muttersprache schreiben, mit solchen, die in Deutsch als einer Fremdsprache schreiben (Variable=*Nativeness*). Die Korpora sind zwei: das Chamisso-Korpus enthält Texte zwanzig interkultureller Schriftsteller mit zehn unterschiedlichen Muttersprachen, während das Vergleichskorpus aus Texten zwölf deutscher Autoren besteht. Die Produktivität der Wortbildungsmuster der ung-Suffigierung quantifiziere ich anhand der Formeln der *realized* und *potential productivity* nach Baayen (2009) und vergleiche sie nach dem „variable-corpus approach“ von Gaeta und Ricca (2006).

Unterschiede in der Produktivität der morphologischen Kategorie *-ung* zwischen fortgeschrittenen DaF-Lernern und Muttersprachlern untersucht auch Berth (2009) mit einer vergleichbaren Methode. Seiner quantitativen Analyse zufolge ist die *potential productivity* der Wortbildungsmuster auf *-ung* in den Texten der L2-Lerner kleiner als die bei den deutschen Muttersprachlern. Trotzdem beweist eine qualitative Analyse der Belege die Anwesenheit zahlreicher Neubildungen, welche die Lerner auf der Basis von Generalisierungen der Wortbildungsmuster auf *-ung* und teilweise gegen die Akzeptabilitätsbeschränkungen der zielsprachlichen Gebrauchsnorm bilden.

Die Vermutung, dass die Ergebnisse meiner quantitativen Untersuchung anders als die von Berth (2009) sein könnten, begründet meine Fragestellung. Die Autoren der literarischen Texte des Chamisso-Korpus’ könnten produktive morphologische Kategorien uneingeschränkter verwenden als die Autoren der argumentativen Texte im Berth’schen Korpus. Soziolinguistische Faktoren könnten darüber hinaus eine Rolle spielen, um Besonderheiten der Sprache der interkulturellen Schriftsteller nach einem ähnlichen Pattern zu stabilisieren. Die linguistischen Merkmale, die sich

als kreativ auswirken, würden in diesem Zusammenhang die interkulturellen Autoren in ihrer literarischen und sozialen Identität kennzeichnen („soziolinguistischen Fossilisation“).
Der Analyse nach der Variable *Nativeness* folgt eine weitere, die den Einfluss anderer Faktoren (des *Auswanderungsgrundes*, der *Muttersprache* und des *Autors*) auf die Produktivität der Wortbildungsmuster auf *-ung* untersucht. Sie zieht daher die Möglichkeit einer Korrelation eines linguistischen Merkmals jeweils mit einem sozialen Zustand, mit Transfer-Phänomenen und mit einer individuellen Neigung in Betracht. Zuletzt analysiere ich den Produktivitätsgrad von drei weiteren morphologischen Kategorien (*-keit/-igkeit/-heit*, *-bar* und *ver-*) im Chamisso- und Vergleichskorpus. Insofern es sich um morphologische Prozesse mit unterschiedlichen strukturellen Eigenschaften und Produktivitätsgraden im gegenwärtigen Deutsch handelt, ist es interessant, die Ergebnisse dieser Untersuchungen mit den Ergebnissen bei den Derivaten auf *-ung* zu vergleichen.
Die allgemeine Zielsetzung meiner Arbeit präzisiere ich im Lauf dieses ersten Kapitels: Die Darstellung der Debatte über die „interkulturelle Literatur“ wirft diesbezüglich einige Fragen auf, die den Rahmen dieser Arbeit bilden. Das zweite, dritte und vierte Kapitel bilden den theoretischen Teil meiner Arbeit, in der die Fragestellung detailliert erklärt wird. Das fünfte Kapitel stellt schließlich die Untersuchungsmethode und die Ergebnisse dar.

1.1 Interkulturelle Literatur: eine Definition

Obwohl das Phänomen einer Literatur von ausländischen nicht-muttersprachlichen Schriftstellern ab den 1980er Jahren das Thema verschiedener Monographien, Sammelwerke und Tagungsbände bildet (vgl. u.a. Weigel 1992, Photong-Wollmann 1996, Chiellino 2000, Rösch 2004, Amodeo 2009) und einen eigenen zusammenfassenden Eintrag in einigen Literatur-Lexika besitzt (vgl. Konzett 2000: 707–709, Burdorf et al. 2007), sind die zu verwendende Denomination wie auch eine einheitliche Definition noch umstritten, wobei in jüngerer Zeit „interkulturelle Literatur“ und „Migrationsliteratur“ die häufigsten Ausdrücke darstellen.
Während sich das Attribut *Migration* hauptsächlich auf den behandelten Themenkreis und auf die aufgeworfenen heiklen sozio-politischen Fragen dieser Literatur bezieht, benennt das Adjektiv *interkulturell* auf eine allgemeinere Weise das Ergebnis des Zusammentreffens von zwei oder mehreren Kultur- und Sprachräumen. Im literarischen Bereich kann *interkulturell* sowohl die Themen als auch die Biographie der Autoren oder deren Einstellung gegenüber den nationalen Literaturtraditionen bezeichnen. Dabei herrscht jedoch Uneinigkeit über die Relevanz der Klassifikationskriterien.

Diesbezüglich versucht Chiellino (2000a: 387–398) eine präzisere Definition zu formulieren, insofern er auf die „drei innovative[n] Komponenten“ hinweist, welche die interkulturelle Literatur im Gegensatz zu einer mono- oder bi-kulturellen kennzeichnen. Die ersten zwei Merkmale identifiziert er in dem besonderen Verhältnis zur Tradition bzw. in dem Adressat des Werkes. Indem er zwei literaturwissenschaftliche Begriffe verwendet, schreibt Chiellino der interkulturellen Literatur nämlich einerseits die Begründung eines „interkulturellen Gedächtnisses“ zu – und zwar eines mit der literarischen Tradition weder des Herkunfts- noch des Auswanderungslandes übereinstimmenden (vgl. auch Biondi & Schami 1981), andererseits die Adressierung an einen „interkulturellen Gesprächspartner“ als idealen Leser. Die dritte und aus meiner Perspektive besonders interessante Eigenschaft bildet die sogenannte „dialogische Zusammensetzung der Sprache“, die in der unauffälligen Einbindung fremder Sprachelemente besteht (Chiellino 2000a: 395).

Im Rahmen meiner Masterarbeit gehe ich von dem Chiellino’schen Begriff der „interkulturellen Literatur“ aus und konzentriere ich mich vor allem auf den dritten Aspekt seiner Definition. Der Ausdruck „interkulturelle Literatur“ scheint mir nämlich infolge seiner Aktualität, aber auch seiner neutraleren Konnotation besser als andere dazu geeignet, die Texte und Autoren, den Gegenstand meiner Untersuchung, als Sammelbegriff zu bezeichnen. Einerseits ergänze ich seine als Ausgangspunkt geltende Definition im Lauf dieses ersten Kapitels mit einigen von anderen Kritikern hervorgehobenen Aspekten, was in der Absicht geschieht, meine Fragestellung genau zu begründen. Andererseits werde ich sie auf bestimmte Kriterien einschränken, wenn ich die Auswahl von Autoren und Texten meines Korpus’ begründen werde (vgl. Kap. 5.1).

Eine erste grundsätzliche Beschränkung kann jedenfalls schon hier ausgesprochen werden und bezieht sich auf die Veröffentlichungssprache der Texte: Meine Arbeit beschäftigt sich nämlich mit „deutschsprachiger interkultureller Literatur“, und zwar nur mit von Nicht-Muttersprachlern auf Deutsch geschriebenen literarischen Texten.

1.2 Die Geschichte der interkulturellen Literatur

Dass die Schwierigkeiten in der Findung einer einheitlichen Terminologie für das Phänomen der deutschen interkulturellen Literatur auch daher rühren, dass es nicht ausschließlich um eine ästhetisch-literarische Bewertung, sondern auch um eine implizite Stellungnahme zu komplexen sozio-politischen Fragen geht, ist den Kritikern bewusst. Daher diskutieren sie rege die Abfolge und die Implikationen der verschiedenen Denominationsvorschlage. Hinzu kommt folgende

Beobachtung von Chiellino (2000a), die darauf verweist, dass die Schaffung einer literarischen Einheit durch die Erfindung eines Begriffs immer auch mit politischen oder ökonomischen Absichten verbunden ist.

> „Hier sei als Zwischenergebnis festgehalten, daß es bislang bei der Suche nach einer griffigen Benennung dieser neuen Literatur keineswegs um eine kritische Würdigung der Werke ging. Denjenigen, die nach einer Bezeichnung suchten, lag sehr daran, entweder die Entwicklung eines vielversprechenden Phänomens zu steuern oder öffentliche Akzeptanz für eine aufkommende deutschsprachige Literatur zu stiften, d. h. ihr eine Leserschicht zu erschließen, um ihr dadurch Kontinuität zu sichern“ (Chiellino 2000a: 391–392).

Die folgende zusammenfassende Darlegung der Definitionen und Denominationen für die deutsche interkulturelle Literatur hat keinen Anspruch auf Vollständigkeit: Ihr Zweck ist einerseits, die Stichhaltigkeit der Kriterien und Gründe für die Ausprägung eines Sammelbegriffs zu diskutieren, andererseits, die Variation in der Rezeption dieser Literatur hervorzuheben, die sich meiner Meinung nach auch aus der Darstellung der Phasen der deutschen Ausländerspolitik der Nachkriegszeit entnehmen lassen. Die grundlegende Hypothese, für deren Erörterung ich auch auf die nachfolgenden Kapitel verweise, ist nämlich die, dass die Bildung eines Sammelbegriffs sowie die von dem sozio-politischen Hintergrund beeinflusste Einstellung des Publikums gegenüber den interkulturellen Autoren soziolinguistische Konsequenzen auf ihre Sprache haben könnte (vgl. vor allem Kap. 2.2, Kap. 2.3, Kap. 3.3.3, Kap. 4.3.1 und Kap. 5.2.3).

Ein wichtiges Datum für die Migrationsgeschichte nach Deutschland ist der Dezember 1955, als die Bundesregierung die erste Anwerbevereinbarung mit der italienischen Regierung schloss, um die Anstellung ausländischer Arbeitskräfte im deutschen Arbeitsmarkt zu kontrollieren. Wenngleich einige deutsche Gewerkschaften sich anfänglich wegen der unangemessenen Lohn- und Arbeitsbedingungen gegen die ersten und die nachfolgenden Anwerbeverträge[1] stellten und die Bundesregierung nach dem „Anwerbestopp“ 1973 besondere Maßnahmen für die Zuwanderungsbegrenzung und teilweise auch die Rückkehrförderung der Gastarbeiter eingeleitet hat, blieb doch eine heftige Auseinandersetzung über das Thema in der deutschen Gesellschaft zu dieser Zeit aus (vgl. Yano 2000: 2–6). Die Gründe dafür erkennt Yano (2000) teilweise in dem von den Anwerbeverhandlungen beförderten Rotationsprinzip der Arbeiter und in der strengen Regulierung der Aufenthaltsgenehmigung nach der Markterfordernis.

[1] Nach der Anwerbevereinbarung mit Italien schloss die Bundesregierung ähnliche Verträge 1960 mit Griechenland und Spanien, 1961 mit der Türkei, 1963 mit Marokko, 1964 mit Portugal, 1965 mit Tunesien und 1968 mit Jugoslawien (vgl. Yano 2000: 2).

Eine Wende in der deutschen Ausländerpolitik findet jedoch zwischen 1979 und 1981 statt: Sie tritt mit dem Beginn der „Integrationspolitik [...] in den Mittelpunkt der Debatte“ und in die „Anerkennung der Bundesrepublik als faktisches Einwanderungsland“ nicht nur seitens der Administration, sondern auch in der Wahrnehmung der Einwohner und Emigranten (Yano 2000: 6). Infolge der zweiten Ölpreisexplosion 1979 und der Entfesselung des sogenannten „Türkenproblems“ entwickelt sich die öffentliche Diskussion über die verschiedenen „Integrationskonzepte“ und Einbürgerungsoptionen. Außerdem ist die Rede von „Überfremdungsängste[n]“, die mit einer Furcht vor einem Qualitätsverlust der deutschen Sprache einhergehen, während 38 % der Ausländer seit mehr als zehn Jahren in Deutschland wohnen und oft ihre Rückkehrpläne fallengelassen haben. Gleichzeitig gewinnt das Phänomen der interkulturellen Literatur im literaturwissenschaftlichen Diskurs Relevanz.

Diesbezüglich stellt Esselborn (1997) eine enge Verbindung zwischen der öffentlichen Migrationsdebatte, der Veränderung der Einstellung der deutschen Gesellschaft gegenüber Ausländern und der Aufmerksamkeit der Kritik gegenüber der interkulturellen Literatur fest (vgl. auch Weigel 1992: 208–211):

> „Die deutschsprachige Literatur nichtdeutscher Autoren trat in der Bundesrepublik zuerst Ende 70er, Anfang der 80er Jahre ins öffentliche Bewußtsein [...]. Nicht zu übersehen ist dabei aber, daß Anfang der 80er Jahre weder die Situation der „Gastarbeiter“ in Deutschland noch die literarische Reaktion der Betroffenen darauf eigentlich neu waren. Als eine wichtige Zäsur könnte man um 1979/1980 allenfalls [...] den Übergang vom Gastarbeiterstatus zur ethnischen Minderheit betrachten, zeitweise von der Diskussion unterschiedlicher Integrationskonzepte begleitet“ (Esselborn 1997: 327).

Eine Kennzeichnung der Gastarbeiter als „ethnische Minderheit“ findet sich auch in einem Aufsatz von Franco Biondi und Rafik Schami (1981: 125), zwei Autoren, die in meinem Korpus vertreten sind: Die Wahl des Begriffs der „ethnischen Minderheit“ begründen sie mit der Feststellung, dass die Reaktion der Gastarbeiterliteraten auf das „Ausländerdasein“ und die Behandlung der „Frage der Identität“ „trotz der Verschiedenheit ihrer Herkunft“ ähnlich seien (1981: 125). Obgleich die Gastarbeiter aus unterschiedlichen Ländern kommen und keine eigentliche ethnische Minderheit bilden, werden sie im Einwanderungsland als eine homogene Gruppe empfunden und stellen sich teilweise auch selbst als eine solche dar.

Dazu kommt der vereinheitlichende Faktor der in der Literaturwissenschaft ausgeprägten gemeinsamen Benennungskategorie der *Gastarbeiterliteratur*, die chronologisch gesehen am Ende der 1970er Jahre die erste Benennung für das Phänomen der interkulturellen Literatur in

Deutschland ist. Sie findet ihr quasi programmatisches Manifest in dem oben zitierten Aufsatz von Biondi und Schami (1981). Die beiden Autoren kritisieren den Begriff „Gast-arbeiter" als wegen des langfristigen Aufenthalts der Auswanderer unzutreffenden Ausdruck (1981: 134 Fußnote 1), trotzdem nehmen sie die von der Kritik aufgedrängte Bezeichnung auf und erklären das Ziel des Schreibens auf Deutsch in der Betonung des Gemeinsamen zwischen den verschiedenen Schriftstellern (1981: 134). Indem die Gastarbeiterautoren Kunstvereine begründen und ihre Werke in gemeinsamen Anthologien und Sammelbänden veröffentlichen (vgl. z.B. der zwischen 1980 und 1987 aktive Verband *PoLiKunst* bzw. die Buchreihe *Südwind Gastarbeiterdeutsch*), verhalten sie sich zudem als eine Art literarische Gruppe.

Die interkulturelle Literatur ist aber nicht nur Gastarbeiterliteratur: Darum werden andere Bezeichnungen vorgeschlagen, welche die Identifizierungskriterien jeweils vermehren und verschiedene Aspekte des Phänomens hervorheben.[2] Biografisch und thematisch begründet sind jeweils die Definition der *Migranten*- und die der schon erwähnten *Migrationsliteratur*, die auf die Lebenssituation der Autoren bzw. den Gegenstand des Werkes abzielen, jedoch im Vergleich zu der *Gastarbeiterliteratur* jede Art der Auswanderung, freiwillig oder unfreiwillig, erfassen (vgl. Photong-Wollmann 1996 und Rösch 2004).

Nicht nur vor dem Hintergrund von Episoden von Fremdenfeindlichkeit in Deutschland im Laufe der 1980er Jahre bis zum Anfang der 1990er Jahre (vgl. Yano 2000: 14–15) ist der Zustand des Emigranten wie auch der des Gastarbeiters tendenziell von sozialer Ausgrenzung geprägt und bildet so zugleich einen wahrscheinlichen Ausgangspunkt für Prozesse der Identitätsaufbau und -zuschreibung. Indem sie die Sprache auch als Mittel der Konstruktion und des Ausdrucks der eigenen Identität betrachtet und auf die linguistischen Konsequenzen der Sprachkontakte und Mehrsprachigkeit Bezug nimmt, stellt Heller (2005: 1584) diesbezüglich fest, dass der Auswanderungszustand „produce empirically observable phenomena such as the use of linguistic resources traditionally understood as belonging to different language and group [...] in ways which make it increaslingly difficult to decide which piece of language belongs to what or to whom".

---

[2] Andere Denominationsvorschläge für die interkulturelle Literatur sind zum Beispiel: *Minority Literatur*, die den Ausgrenzungszustand der Autoren in dem Auswanderungsland als grundsätzlich betrachtet (vgl. Konzett 2000: 707–709), oder *Außen*- bzw. *Rand*- und *Brückenliteratur*, die teilweise auf den Diskurs über die Rolle des Zentrums und der Peripherie in den modernen Gesellschaften Bezug nimmt (vgl. Chiellino 2000a: 391). Weigel (1992) fasst dagegen das Phänomen als *Literatur der und in der Fremde* zusammen und schließt in die Kategorie nicht nur ausländische und zeitgenössische Autoren, sondern auch die Deutschen ein, die in ihrer Muttersprache über Fremdheitserfahrungen schreiben, während Biondi & Schami (1981) ihre Literatur als Gastarbeiterliteratur, aber auch als eine *Literatur der Betroffenheit* definieren, um damit die politischen Themen und Zwecke ihrer Werke hervorzuheben. Die Ausdrücke der *Ausländer*- und *Gast-Literatur* stammen schließlich von den Forschern des DaF-Instituts in München und erkennen das Schreiben in einer Fremdsprache als gemeinsame Eigenschaft (vgl. Kap. 1.3): Der zweite von Weinrich ausgeprägte Begriff bezieht sich auf einen Satz von dem Schriftsteller Elias Canetti, der sich, indem er das Deutsche gerade einmal mit acht Jahren als dritte Sprache gelernt hat, als einen „Gast in der deutschen Sprache" bezeichnet (vgl. Weinrich 1983: 920).

Chiellinos Beschreibung der interkulturellen Autoren soll also in diesem Zusammenhang mit dem daraus resultierenden Bild einer Literatur ergänzen werden, die in ihrer Rezeption und wahrscheinlich auch in ihren verschiedenen Eigenschaften von sozialen Faktoren bedingt ist.

1.3 Die Sprache der interkulturellen Autoren

Unter den verschiedenen Definitionen für die interkulturelle Literatur ist darüber hinaus die des DaF-Instituts in München besonders interessant, insofern sie nicht auf biografischen oder thematischen Elementen beruht, sondern linguistische Kriterien fokussiert. Da ausländische, auf Deutsch schreibende Schriftsteller sich nicht nur unter den Gastarbeitern, sondern auch unter Flüchtlingen, Aussiedlern und Asylanten, sowie unter Studenten und Praktikanten befinden[3], erklärt Weinrich (1983) den Sprachwechsel und das Schreiben in einer Fremdsprache als konstitutive und vereinheitlichende Faktoren für die Gruppierung. Ihnen weist er auch einen ästhetischen Wert zu, insofern eine gewisse Distanz und Fremdheit zu der deutschen Sprache, unabhängig von in der Regel sehr guter Beherrschung, immer bestehen bleibt und Konsequenzen für die Produktion und Rezeption der Texte bedeutet (1983: 918).

Wie sich die ausländischen Schriftsteller linguistisch genau verhalten und welches die Anzeichnen und Wirkungen des Sprachwechsels sind, wird aber nicht eindeutig festgestellt. Einerseits spricht Weinrich (1983) im Allgemein über die Bereitschaft, „sich mehr als andere auf die Sprache einzulassen, sich ihrer Führung williger anzuvertrauen und genauer auf die Wörter und ihre Bedeutung zu achten“ (918), sowie über die konkrete Anwesenheit einiger von der deutschen Sprachnorm abweichender Ausdrücke in den Texten (919–920). Anderseits stimmen zwei Autorinnen meines Korpus’, Léda Forgó und Yoko Tawada, in einem Gespräch mit Amodeo (2009: 143–176) überein, dass die Wahl einer Fremdsprache besondere kreative Möglichkeiten bietet, warnen aber davor, alle abweichenden sprachlichen Konstruktionen auf den Einfluss der Muttersprache zurückzuführen.[4]

Ackermann (1997) fasst diesbezüglich die Position gegenüber dem Phänomen der interkulturellen Literatur des DaF-Instituts in München wie folgt zusammen. Von „eine(r) größere(n)

[3] Die drei von Weinrich hier erwähnten Kategorien geben mir für die Analyse meines Korpus’ Anlass, jeweils auch nach dem Auswanderungsgrund der interkulturellen Schriftsteller zu fragen. Von Weinrich (1983) abgeleitet lauten nämlich die drei festgestellten Ausprägungen der Variable *Auswanderungsgrund* in meiner Untersuchung: „Arbeit (Gast-)“, „Exil“ und „Privat“ (vgl. Kap. 4.2 und Kap. 5.2.3.1).

[4] Sowohl thematisch als auch theoretisch angegriffen ist die Frage über die Konsequenzen der Mehrsprachigkeit auch in Texten anderer in meinem Korpus repräsentierten Autoren zu finden: Die türkische Schriftstellerin Emine Özdamar veröffentlichte zum Beispiel 1990 den Erzählungsband *Mutterzunge*, während Tzveta Sofronieva in den 1990ern das Projekt *Verbotene Worte* initiiert, das Texte über belastete Wörter, ihre Geschichte und Veränderungen auch durch die Begegnung verschiedener Sprachen und Kulturen enthält.

Sensibilisierung für sprachliche Besonderheiten“ der mehrsprachigen Autoren im Vergleich zu den einsprachigen ausgehend (1997: 21), erkennt Ackermann gemeinsame Modelle für die von dieser Literatur verlangten „sprachlichen Neusituierung“. Auf der einen Seite thematisiert der gezielte Gebrauch von inkorrekten und deformierten Ausdrücken des Deutschen oder von muttersprachlichen Elementen die „sprachliche Heimatlosigkeit“ der Autoren und zielt auf die Verfremdung der Lesegewohnheiten ab (1997: 22). Auf der anderen Seite führen der kreative Umgang mit der deutschen Sprache und die Bemühung, ihr Angebot weitest möglich auszuschöpfen, zu „eigenwillige[n] Wortschöpfungen, grammatische[n] Konstruktionen ʼgegen den Strichʼ, ungewohnte[n] Wortverbindungen und Sprachspiele[n]“ (1997: 26).
Aus der Aufmerksamkeit des DaF-Instituts in München für diese Literatur folgt darüber hinaus deren Förderung durch Preisvergabe und Anthologieveröffentlichungen (vgl. Esselborn 1995). Darunter findet sich der auf Veranlassung von Harald Weinrich und der Robert Bosch Stiftung eingerichtete *Adelbert von Chamisso-Preis für Beiträge zur deutschen Literatur von Autoren nichtdeutscher Muttersprache*, dessen Name auf den deutschen Dichter französischer Herkunft – Adelbert von Chamisso (1781–1838) – zurückgeht. Von diesem Preis, der seit 1985 jedes Jahr an einen oder mehrere auf Deutsch schreibende Nicht-Muttersprachler-Autoren verliehen wird, werde ich noch sprechen, insofern diese Autoren die erste Grundlage für den Aufbau meines Korpus’ bilden (vgl. Kap. 5.1.1).

1.4 Die Hypothese

Die Überlegungen der Forscher aus München, insbesondere in der von Ackermann (1997) dargestellten Form, sind der Ausgangspunkt für die Frage, die meine Masterarbeit aufwirft. Es handelt sich um die Frage, ob konkrete linguistische Merkmale in der Sprache ausländischer Schriftsteller zu finden sind, die den zweiten Teil von Ackermanns Aussage rechtfertigen, und eine „zusätzliche Qualifikation“ der interkulturellen Schriftsteller im Vergleich zu den Deutschen im Hinblick auf das „Experimentieren mit der Sprache“ belegen (1997: 26).
Aus der bisherigen Forschung ergibt sich Annahme, dass die untersuchten Eigenschaften die Texte der interkulturellen Autoren unabhängig von der jeweiligen Muttersprache kennzeichnen und daher zumindest auf der theoretischen Ebene nicht als Transferphänomene beschreibbar sind. Davon ausgehend, dass ausländische Schriftsteller ihre Texte vor der Veröffentlichung mehrmals korrigieren, erwarte ich schließlich, keine unbeabsichtigten Fehler oder rein ungrammatische

Konstruktion zu finden, sondern einen Über- oder Mindergebrauch einiger Strukturen des Deutschen.

Die hier vorgestellte Hypothese fasst daher drei unterschiedliche, aber miteinander verbundene Unterhypothesen zusammen, die ich anhand von Theorien und Begriffen aus verschiedenen Bereichen der Linguistik in den drei nachfolgenden Kapiteln erläutern werde.

Nach der ersten Unterhypothese weist die Sprache ausländischer Schriftsteller trotz der individuellen Unterschiede gemeinsame Eigenschaften auf. Als Anlässe zu einer solchen Möglichkeit gelten zuerst der Zustand der ausländischen Autoren als Fremdsprachen-Lerner (L2-Lerner), ferner ihr besonderer Status in der Auswanderungsgesellschaft als Emigrant, sowie in dem literaturwissenschaftlichen Diskurs, der sie als literarische Gruppe betrachtet. Um diesen Punkt zu erklären, möchte ich mich deshalb auf einige Begriffe aus der Zweitspracherwerbsforschung und der Soziolinguistik beziehen: das Konzept der *Interlanguage* sowie die Begriffe *Fossilisation* bzw. *Varietät* (vgl. Kap. 2).

Die zweite Unterhypothese nimmt einen Unterschied zwischen dem linguistischen Verhalten der interkulturellen und der deutschen Schriftsteller bezüglich einer für die erste Gruppe gemeinsamen sprachlichen Eigenschaft an, die mit einem Aspekt der „sprachlichen Kreativität" zusammenhängt.[5] Im dritten Kapitel geht es deshalb um die Feststellung des zu untersuchenden linguistischen Merkmals und die Begründung der Annahme seiner eventuellen Variation. Da mein Merkmal aus dem derivativen und produktiven Wortbildungssystem stammt, werde ich einerseits den von mir angenommenen Begriff der „morphologischen Produktivität" und die damit zusammenhängende morphologische Theorie erörtern, anderseits die Gründe für die Wortschöpfung und den Ablauf des Erwerbs der morphologischen Kompetenz seitens der L2-Lerner (auch im Vergleich zu L1-Lernern) besprechen.

Die dritte Unterhypothese betrifft die Art und Weise einer eventuellen Asymmetrie zwischen den Texten beider Gruppe. Die Vermutung, dass der Unterschied in der Frequenz der analysierten Variable liegt und deswegen in einem Über- oder Mindergebrauch einiger produktiver Wortbildungsmuster des Deutschen besteht, ist mit der Beschaffenheit der Texte und dem „fortgeschrittenen Niveau" der interkulturellen Autoren als L2-Lerner eng verbunden. Um überprüft werden zu können, setzt sie darüber hinaus eine bestimmte und zu rechtfertigende

[5] Mit „sprachlicher Kreativität" bezeichnet man konventionell in Linguistik die Fähigkeit eines Sprechers, (potenziell) unendlich viele Wörter und Satze zu produzieren und zu verstehen, und darunter auch solche, die neu sind und noch nicht ausgesprochen wurden (vgl. Swann et Al. 2004: 60). Der Begriff, der schon von der Humboldt'schen Sprachtheorie thematisiert wurde, nimmt im Rahmen der Generativen Grammatik eine zentrale Stelle ein. In diesem Zusammenhang wird die linguistische Kreativität zumindest in einigen seiner Formen als „regelgeleitet" definiert. Ein gutes Beispiel dafür stellen die mittels produktiver morphologischer Kategorie zusammengesetzten Neubildungen dar: Neue Wörter werden nämlich durch regelgesteuerte und regelbegrenzte morphologische und semantische Muster gebildet (vgl. Kap. 3).

Untersuchungsmethode voraus (vgl. Kap. 4). Da ich mich mit literarischen Texten beschäftige, werde ich ferner im dritten Teil des vierten Kapitels auf einige Ergebnisse der stilometrischen und korpusstilistischen Forschung Bezug nehmen. In diesem Zusammenhang ist die gegenseitige Wirkungsbeziehung zwischen Stil und Morphologie Grundlage für die Hypothese, dass der von mir untersuchte Texttyp[6] für andere Ergebnisse im Vergleich zu der Berth'schen Analyse verantwortlich sein könnte.

Schließlich stelle ich im fünften Kapitel den Verlauf des Aufbaus meines Korpus' und die Ergebnisse meiner Untersuchung vor. Insofern ich mich nur auf die vollständige Analyse eines linguistischen Merkmals und eine teilweise Untersuchung dreier anderer Merkmale beschränke (und zwar der Produktivitätsgrad der Wortbildungsmuster auf *-ung* bzw. auf *-keit/-igkeit/-heit*, auf *-bar* und auf *ver-*), ist aber schon hier zu erwähnen, dass meine Arbeit bezüglich der Frage der angeblichen sprachlichen Kreativität seitens der interkulturellen Schriftsteller keinesfalls alle Fragen beantworten kann. Dies gilt auch für die Betrachtung der Korrelation zwischen morphologischer Produktivität und soziolinguistischen Faktoren. Diesbezüglich verstehe ich sie in einer ersten Analyse als weitere Bausteine, welche die gemeinsame Gruppierung der „interkulturellen Autoren" begründen und die Fossilisation ihrer *Interlanguage* sozial oder soziolinguistisch verursachen könnten (vgl. Kap. 2.2). In einer zweiten Analyse betrachte ich einen einzigen Faktor etwas genauer, um seinen Einfluss auf den Gebrauch der morphologischen Kategorie *-ung* zu überprüfen. Es sei hier aber hervorgehoben, dass andere soziale Eigenschaften und weitere linguistische Variablen meine Untersuchung ergänzen müssten, um eine genauere Analyse der soziolinguistisch bedingten sprachlichen Variationen bei interkulturellen Schriftstellern durchzuführen.

[6] Unter „Texttypen" versteht man die verschiedenen Arten von Texten, die in einer Sprachgemeinschaft vorkommen. Texte werden sowohl anhand ihrer Funktion als auch anhand ihrer linguistischen Merkmale klassifiziert (vgl. Swann et Al. 2004: 313). Man unterscheidet zum Beispiel zwischen *argumentativen*, *narrativen*, *explikativen* und *deskriptiven* Texten gemäß ihrer Funktion (überreden, erzählen, erklären bzw. beschreiben) sowie gemäß den i.d.R. verwendeten verbalen Tempora oder Personalpronomina. In der neueren Forschung werden häufig Bündel von positiv oder negativ ko-okkurierenden linguistischen Variablen durch *principal component analyses* oder *factor analyses* ermittelt und denen eine Funktion zugeschrieben (vgl. unter den in dieser Arbeit zitierten Aufsätzen, Baayen 1994a). Eine Klassifikation nach Texttyp wird dann auf der Grundlage der dadurch festgestellten „Dimensionen" (Bündel von Merkmalen + Funktion) durchgeführt.

2. Die erste Unterhypothese: Anwesenheit gemeinsamer linguistischen Merkmalen

2.1 Die Lernersprache

Mit dem Begriff der *Interlanguage* (IL, auch Lernersprache) bezeichnet Selinker (1972) das von der Muttersprache (NL) und der Targetsprache (TL) separate sprachliche System, dem die bedeutungsvollen Reproduktionsversuche einer Fremdsprache seitens ihrer Lerner unterliegen und dessen Zusammensetzung sich aus der Analyse der inkorrekten wie auch korrekten Äußerungen der L2-Sprecher, sowie aus ihrem Vergleich mit den Äußerungen der Muttersprachler bestimmen lässt (vgl. auch Tarone: 1994). Gewisse Regelmäßigkeiten im sprachlichen Verhalten der Lerner zusammen mit den Besonderheiten der Mechanismen, die den Erwerbsablauf steuern und oftmals einschränken, lassen sich nämlich laut Selinker (1972) mit der Aktivierung einer latenten psychologischen Struktur erklären, die aber nicht mit dem für den Erstspracherwerb zuständigen *Language Activation Device* (LAD) übereinstimmt.[7]

Die Theorie der *Interlanguage* findet ihre Grundlage in zwei früheren Ansätzen der Zweitspracherwerbsforschung (SLA): der Kontrastiven Linguistik (CL) und der Fehleranalyse (EA) (vgl. Tarone 1994 und Nickel 1998). Diese Methoden nehmen sich das Vorhersagen der von den Lernern getroffenen Erwerbsschwierigkeiten vor, indem die eine sie in den Unterschieden zwischen TL und NL feststellt, die andere sie aus den Fehlern und einer Untersuchung ihrer möglichen Gründe heraus ableitet. Wenngleich die *Interlanguage*-Forschung die Ergebnisse der CL und EA teilweise in Betracht nimmt, fasst sie sie doch in einem neuen theoretischen Rahmen zusammen und weist deutliche methodische Unterschiede auf. Deren Erörterung zielt hier auf die Präzisierung des Begriffs der „Lernsprache“ und des damit verbundenen Konzepts der „Fossilisation“ ab.

Im Gegensatz zur Kontrastiven Linguistik spielt nach der *Interlanguage*-Hypothese weder nur die Muttersprache im Erwerbsablauf eine Rolle, noch ist allein der sprachlichen Transfer der einzige für diesen Zweck aktivierte Mechanismus. Selinker (1972) erwähnt nämlich fünf dafür wesentliche linguistische bzw. psychologische Prozesse: Neben dem Transfer aus der Muttersprache ist auf der

[7] Da die Erwachsenen im Gegensatz zu den Kindern nur selten eine vollständige Beherrschung der TL erreichen und ihre Äußerungen von eigenartigen Phänomenen, wie dem Transfer aus dem NL, gekennzeichnet sind, nimmt Selinker (1972) einen wesentlichen Unterschied zwischen dem Zweitsprach- und Erstspracherwerb an. Eine solche Hypothese ist aber noch umstrittenen. Für ihre Bestimmung spielt u.a. die Frage eine wichtige Rolle, ob die *Interlanguage* eine natürliche Sprache bildet und den Prinzipien der Universalen Grammatik unterliegt. Im Gegensatz zu Selinker (1972) sind nämlich z.B. für Adjemian (1976) auch die Lernersprachen natürliche Sprachen und daher ist ihr Erwerb von demselben *Acquisition Device* wie bei Kindern gesteuert. Die oben erwähnten Besonderheiten erklären in diesem Zusammenhang die *Permeability* der IL-Grammatiken auf die Anwendung bzw. Generalisierung von fremden und eigenen Regeln und die Annahme, dass das Funktionieren des LAD sich nach einer ersten Parametrisierung leicht verändert.

einen Seite die Generalisierung von Strukturen der Targetsprache zu beachten, auf der anderen Seite kommen die allgemeinen Lern- und Kommunikationsstrategien sowie das kulturspezifische *transfer-of-training* hinzu.[8] Auf diese Prozesse lassen sich ferner Fossilisationsphänomene zurückführen, welche in sprachlichen Elementen („items, rules, and subsystems") bestehen, die die Lerneräußerungen unabhängig von dem Alter oder der Studiumsdauer kennzeichnen und die, auch wenn sie ausgemerzt scheinen, stetig wiederkehren (vgl. Selinker 1972: 215). Insofern fossilierte Elemente nicht nur von Transfers, sondern auch von TL-spezifischen Strukturen bzw. generellen Lern- und Kommunikationsprinzipien bedingt sind, kann die *Interlanguage* eine eigene Norm entwickeln, die NL-unabhängige Eigenschaften aufweist und teilwiese systematischen und allgemeinen Tendenzen folgt (vgl. Vogel 1990: 47–50, Klein et al. 1992 und Kap. 2.3).[9]

> „Mit Hilfe testmethodischer-statistischer Reduktionsverfahren (wie der Lernergruppenbildung nach Lerneralter, Wissensstand, Kenntnis weiterer Fremdsprachen, Begabung, sozialer Hintergrund etc.) und der Standardisierung von Lerneräußerungen ist es trotz der Instabilität individueller Lerneräußerungen möglich, für einen ‚idealisierten gruppentypischen Normal-Französischlerner', [...] das Sprachwissen, also die Lernergrammatik als relativ stabile Größe zu rekonstruieren [...] [und] Fremdspracherwerbsuniversalien an[zu]nehmen" (Vogel 1990: 49).

Die Erklärungskraft einer nur auf Transfer beruhenden Theorie für SLA zweifelt auch ein auf Fehler gerichteter Ansatz an: Obwohl man inkorrekte Ausdrücke anfänglich mit dem Zweck neuer empirischer Beweise für die Hypothesen der Kontrastiven Linguistik untersucht hat, verursachte die Entdeckung der Verschiedenheit ihrer Ursachen die Entwicklung alternativer Theorien.
Die Betrachtung der Fehler als Zeichnen des Erwerbsverfahrens setzt jedoch eine mangelhafte Beschreibung der linguistischen Fähigkeiten eines Lerners voraus, indem sie zum Beispiel die

---

[8] Der Transfer ist ein Phänomen von „interlingual identification" (Selinker 1972: 211), infolgedessen ein Lerner eine linguistische Einheit als identisch in drei verschiedenen sprachlichen Systemen (TL, ML, IL) wahrnimmt. Das Transfer-of-Training folgt dagegen aus der Anwendung von in Handbüchern oder im Unterricht gelernten Regeln. Die Kommunikations- und Lernstrategie benutzt der L2-Sprecher oft bewusst, um konkrete Probleme im Gespräch zu lösen bzw. die Targetsprache zu meistern: hierunter zählt man jeweils Paraphrasen von einem unbekannten Wort oder das sich ins Gedächtnis Einprägen von Verbdeklinationen. Vereinfachungstendenz können in diesem Zusammenhang sowohl als kommunikativ- als auch als Lernen-bedingt betrachtet werden (vgl. Tarone 1994: 1716–1717 und Selinker 1972: 217–221). Schließlich verursacht z.B. die Generalisierung von TL-Strukturen nach Selinker die kritiklose Anwendung einer Regel ohne die Betrachtung ihrer Ausnahme: Der Prozess kennzeichnet auch den Erstspracherwerb und bildet den allgemeinen Hintergrund, auf dem Variationen meiner Variable entstehen können (vgl. Kap. 3).

[9] Die Frage nach den ersten Gründen der Fossilisation, danach, ob es sich um eine spezifische und daher unvermeidliche Konsequenz des Funktionierens der latenten psychologischen Struktur oder um eine Wirkung der schon durchgeführten Parametrisierung des LAD handelt (vgl. auch Fußnote 7), kann hier nicht diskutiert werden. Es ist aber in diesem Zusammenhang interessant anzumerken, dass Fossilisation ein verbreitetes Phänomen darstellt und für die Mehrheit der Lerner (nach Selinker 1972 und Nickel 1998: 95 % der Fälle) das Erreichen einer vollständigen Äquivalenz mit der Target-Sprache verhindert.

Anwendung von Vermeidungsstrategien oder Frequenzvariationen im Gebrauch einiger sprachlichen Strukturen zwischen Lernern und Muttersprachlern nicht erfasst.

Im Gegensatz zu der Fehleranalyse versteht die *Interlanguage*-Theorie die linguistische Kompetenz eines Lerners als eine veränderliche, aber vollständige Hypothese des Funktionierens der Target-Sprache, die sich auf alle linguistischen Ebenen – Phonologie, Morphologie, Syntax und Pragmatik – bezieht (vgl. Tarone 1994: 1716). *Interlanguage*-spezifische bzw. fossilierte Eigenschaften bestehen darüber hinaus nicht nur aus Lernerfehlern, sondern „‚correct things' can also reemerge when thought to be eradicate, especially if they are caused by processes other than language transfer" (Selinker 1972: 216 Fußnote 14). Die IL-Norm kann daher Variationen auch im Gebrauch korrekter Ausdrücke bestimmen, deren Erkennung eine andere Untersuchungsmethode bzw. umfangreichere Lerneräußerungsaufnahmen im Vergleich zu den in CL oder EA verwendeten benötigt (vgl. Kap. 4.1 und 4.2).

Infolge der Selinker'schen (1972: 209–210) Unterscheidung zwischen Lerner- und Lehrerstandpunkt tritt schließlich in Bezug auf die CL- und EA-Perspektive ein Fokuswechsel auf. Man betrachtet nämlich die *Interlanguage* nicht nur als imperfektes Abbild der Target-Sprache, sondern untersucht auf der Grundlage der Legitimität des Lernerstandpunkts ihre Eigenschaften auch als Ergebnisse der produktiven Leistung von Lern- und Kommunikationsstrategien und hebt ihre kreativen Aspekte hervor (vgl. Nickel 1998: 69). Dieser Ansatz begründet teilweise meine Herangehensweise an die Analyse der Sprache der interkulturellen Autoren: Diese wird hier nämlich aufgrund der Erwerbssituation als eine Art *Interlanguage* betrachtet und ihre eventuellen linguistischen Besonderheiten werden dabei nicht nur als Erwerbsschwierigkeiten verstanden, sondern auch als die Ausdrucksmöglichkeiten, welche die Lernersprache als System für den einzelnen Lerner bietet (vgl. Kap. 3.4). Die Theorie der *Interlanguage* ist für mich zugleich Grundlage für die These, dass sich die beobachtbaren Differenzen eventuell als eine für die gesamte Gruppe der interkulturellen Autoren geltende sprachliche *Varietät* beschreiben lassen (vgl. Kap. 2.3).

2.2 Die soziale und soziolinguistische Fossilisation

Wenn die Annahme der Systematizität der Lernersprache ihre wissenschaftliche Untersuchung ermöglicht, insofern sie die Struktur der Lerneräußerungen und die Erwerbsabfolge teilweise vorhersagbar macht, leugnet man dadurch nicht, dass die zeitliche und individuelle Variabilität, sowie die Empfindlichkeit gegenüber außerlinguistischen Faktoren wie etwa dem psychophysischen

Sprecherzustand oder der Situation weitere wesentliche Eigenschaften der *Interlanguage* sind (vgl. Vogel 1990: 46–57). Die lernersprachliche Instabilität, verursacht durch den stetigen Kontakt mit der TL und der NL bzw. durch den Mangel von verschriftlichten Übereinkünften, ist einigen Forschern zufolge ein zusätzlicher Punkt für die Unterscheidung der *Interlanguages* von den „natürlichen Sprachen". Diese Instabilität bildet auf den ersten Blick einen möglichen Einwand gegen die hier vertretene Hypothese der Entwicklung gemeinsamer Eigenschaften in der Sprache der interkulturellen Autoren:

> „Eine Lernersprache ist nicht wie eine natürliche Sprache die Sprache einer Sprachgemeinschaft, sie ist streng genommen ein individuelles und nicht ein soziales Phänomen in dem Sinne, dass sie nicht eine *langue* ist, [...] somit auch nicht ein Produkt, das zwar virtuell im einzelnen Sprecher, vollkommen jedoch nur in der Gesamtheit der Individuen einer Sprachgemeinschaft existiert" (vgl. Vogel 1990: 14).

Trotz der Variabilität und der fehlenden Institutionalisierung der *Interlanguages* (aber vgl. Kap. 2.3) beeinflussen den *individuellen* Erwerb einer Fremdsprache nicht nur strukturelle bzw. psychologische Faktoren, sondern vor allem bei Erwachsenen auch soziale und interrelationale Variablen (vgl. u.a. Perdue 1984: 68–97, Züngler 1989, Preston 1993, Dittmar 1997). Relevantes Beispiel dieser Interdependenz ist die Einwirkung der „Identität" eines L2-Sprechers auf das Verhalten der Muttersprachler ihm gegenüber und der beiden auf einige lernersprachliche Eigenschaften.

Unter dem schon in Kapitel 1.2 erwähnten Begriff der „Identität" versteht man „a set of practices and representation regarding social categories which are produced and reproduced in social interaction in everyday life" (Heller 2005: 1584). Die Identitätszuschreibung und ihre sozialen Ausprägungen entstehen aus dem Sozialisierungsprozess und von daher zum Teil aus der sprachlichen Interaktion eines Individuums mit anderen Gesellschaftsmitgliedern, wobei die Erkennung der Zugehörigkeit einer Gruppe eine wesentliche Rolle spielt. Obwohl die Identitätseigenschaften in der neueren Forschung als je nach Situation und Zeit dynamisches Bündel von Merkmalen eines Individuum erfasst werden und von daher an Bezugsfestigkeit verloren haben (vgl. u.a. Heller 2005, Berruto 1995: 29–32 und Kap. 4.3), können sie trotzdem von den Gesprächspartnern bewertet werden und dadurch eine gewisse Konnotation innerhalb einer Sprachgemeinschaft gewinnen, die auch die sie charakterisierenden linguistischen Eigenschaften einbezieht.

Unter den Identitätseigenschaften sind *Nativeness* und *Ethnicity* für die SLA-Forschung und diese Arbeit besonders relevant. Erstere umfasst alle L2-Lerner als „nicht-native Sprecher", obwohl man

Nativeness-Stufen unterschieden kann, die mit dem Sprachbeherrschungsgrad (*Fluency*) des Einzelnen zusammenhängen, während Letztere die Frage der mehr oder weniger starken Zugehörigkeit zu einer ethnischen Minderheit einbezieht. Trotz der Differenzen sind jedoch die Auswirkungen beider Eigenschaften auf den Spracherwerbsprozess nach Preston (1993) vergleichbar. Sowohl aus dem *Nativeness* als auch aus der Ethnizität der L2-Sprecher können nämlich Fossilisationsphänomene folgen, die daher als sozial- und soziolinguistisch-bedingt beschrieben werden (Preston 1993: 78–91 und 252–269).

Die soziale Fossilisation ist die unmittelbare Konsequenz der sozialen Attribute der Lerneridentität und der Einstellung der neuen Sprachgemeinschaft ihnen gegenüber, da „[j]ust as an empathic interlocutor can have an effect on the IL speaker, so can an interlocutor who threatens the IL speaker's ethnicity identity" (Züngler 1989: 85). Davon unterscheidet sich die soziolinguistische Fossilisation dahingehend, dass die fossilierten linguistischen Eigenschaften einen metaphorischen Wert gewinnen. Indem sie die Identität des Lerners als Nicht-Muttersprachler symbolisieren, werden daher diese Eigenschaften oft in einer Sprachgemeinschaft mit *Stigma* oder *Prestige* besetzt, wobei deren Anwesenheit und Frequenz mit dem Willen des Sprechers zum Ausdruck seiner Identität variieren. Die soziolinguistische Fossilisation kennzeichnet zum Beispiel Fälle additiver Zweisprachigkeit (oder *Accommodation*), wobei der Sprecher einer linguistischen Minderheit in der Kommunikation außerhalb seiner Gruppe den „dominierenden Code" bzw. die Sprache der Mehrheit benutzt, aber fossilierte sprachliche Merkmale beibehält, die Aspekte seiner Identität darstellen:[10]

> „In accommodation situation, a dominated language group uses its own language in interethnic communication, and it may persist in using a *distinctive ethnic accent* in the dominant language" (Preston 1993: 90. Von mir hervorgehoben).

Von Prestons Zitat ausgehend sind zwei weitere Punkte zu erörtern. Obgleich Preston (1993) die soziale und soziolinguistische Fossilisation hauptsächlich mit Beispielen aus ethnischen Minderheiten erklärt, beschränkt er sie nicht nur auf diese Fälle. Diesbezüglich bildet in der Literatur ein oft erwähntes Beispiel genau das der „low-status immigrant workers in Europe" (vgl. Preston 1993: 117 und u.a. Perdue 1984: 69, Vogel 1990: 48). Als diese Perspektive

[10] Am anderen Ende des Kontinuums finden sich Fälle von „subtractive bilingualism" oder „assimilation", in denen eine Minderheit bzw. ein Individuum durch den L2-Erwerb ihre bzw. seine ursprüngliche Sprache verliert. Von einer engen Verbindung zwischen Sprache und Identität ausgehend sind nach Preston (1993: 78–91) die Ergebnisse sowohl auf individueller als auch auf sozialer Ebene eher negativ: Assimilation folgt nämlich aus einem Überdruck der neuen Sprachgemeinschaft und verursacht oft Anomie im L2-Lerner bzw. Misskredit seitens der nativen Sprecher.

zusammenfassend gilt der folgende Passus von Züngler (1989), der das oben erwähnte Zitat über das Phänomen der sozialen Fossilisation direkt fortsetzt:

> „The most notable example of this [dem Effekt eines bedrohenden Ansprechspartners] on a large scale concerns the *Gastarbeiter* research in Europe [...]. It is suggested that because the adult immigrant workers occupy clearly disadvantage roles in their host communities, any pre-existing tendencies toward racism and discrimination are heightened. Interacting with TL speakers in such a threatening social context impedes the workers' second language acquisition" (Züngler 1989: 85).

Die wechselseitigen Identitätszuschreibungsprozesse und das daraus resultierende linguistische Verhalten sind also das Ergebnis der Übereinstimmung verschiedener Faktoren, unter denen neben der ethnischen Gemeinsamkeit auch das Geschlecht, die berufliche Tätigkeit oder die schulische Erziehung eine Rolle spielen (vgl. Züngler 1985: 84). Angesichts dieser Tatsache untersuche ich die Sprache der interkulturellen Autoren nicht nur bezüglich der Variablen *Nativeness* und *L1*[11], indem ich sie einerseits mit der Sprache deutscher Schriftsteller, andererseits untereinander vergleiche und so auf den spezifischen Einfluss ihrer Muttersprache hin analysiere. Sondern gehe ich auch der Variable *Auswanderungsgrund* nach, die den sozio-ökonomischen Status[12] der Autoren in der deutschen Gesellschaft zumindest teilweise widerspiegelt (vgl. Kap. 5.2.3.1). Eine weitere Variable meiner Untersuchung ist schließlich die der *Individuality*: Sie bildet in der Soziolinguistik traditionell keine besonders relevante Bezugskategorie, vor allem in der SLA-Forschung kann sie aber laut Preston (1993: 119–121) nicht außer Acht gelassen werden (vgl. Kap. 5.2.3.3).
Zweitens sind, wie das oben erwähnte Preston´schen Zitat andeutet, die fossilierten Eigenschaften, die am wahrscheinlichsten Identitätszeichen werden, diejenigen, die den Sprechern relativ bewusst sind und daher häufig schon linguistische „Marker" bzw. „Stereotype" bilden.[13] Es handelt sich

---

[11] Mit „nativ" oder „nicht-nativ" unterscheide ich zwischen Schriftstellern, die ihre Texte in der Muttersprache, und solchen, die in einer Fremdsprache schreiben, ohne aber im Gegensatz zu Preston (1993) auf den Sprachbeherrschungsgrad Bezug zu nehmen. Der Variable *L1* teilt dagegen die Gruppe der interkulturellen Autoren je nach ihrer Muttersprache auf. Da in meinem Fall die Muttersprache mit der Nationalität bzw. *Ethnizität* der Autoren übereinstimmt, könnten Transfer-phänomene auch soziolinguistisch bedingt sein.

[12] Ursprünglich erfasste die Variable *Status* kulturelle Aspekte der sozialen Schichtung und stimmte nicht völlig mit der auf ökonomischen Kriterien festgestellten Variable *Klasse* überein (vgl. Swann et al. 2004: 297, der darauf hinweist: „For example, the social status of immigrant workers is usually lower that the status of indigenous workers, although both may share the same class position"). Im variationistischen Ansatz entwickelt sich jedoch die zusammenfassende Kategorie des *sozio-ökonomischen Status*', der Informationen über die berufliche Tätigkeit, den Lohn, die Wohnsituation und die schulische Erziehung zusammenstellt, um die korrelierende linguistische Variation innerhalb einer Sprachgemeinschaft zu beschreiben. Da die soziale Herkunft der interkulturellen Autoren in meiner Arbeit nur durch die Variable *Auswanderungsgrund* erfasst wird, ist meine Betrachtungsweise diesbezüglich etwas vereinfacht.

[13] Nach der Labov'schen dreiteiligen Unterscheidung spricht man von *Indikatoren* für linguistische Merkmale, welche die Sprache eines Sprechers mit sozialer Zugehörigkeit prägen, aber keine stilistischen Verteilung aufweisen: Sie sind linguistische Variable, die dem Sprecher unbewusst sind. *Marker* variieren nach sozio-ökonomischem Status, aber auch

demnach vor allem bei L2-Lernern um phonetische Merkmale oder L1-bedingte Strukturen, sodass man den Transfer als nicht nur psycholinguistisch-, sondern auch sozial-bedingtes Phänomen erfasst (vgl. Züngler 1989: 87). Nicht ausgeschlossen ist aber, dass ein Element, das aufgrund eines anderen Selinker'schen Lernprozesses fossiliert, einen symbolischen Wert gewinnt:

> „Of course, the aetiology for some bilingual 'mannerism' [...] may be originally learning problem of transfer, overgeneralisation, and so on. Such differences would, however, quickly acquire the kinds of symbolic values suggested here" (Preston 1986: 20–21).

Ein von der Generalisierung einer morphologischen TL-Struktur bestimmtes lernersprachliches Merkmal könnte deshalb einer auch soziolinguistischen Fossilisation unterliegen und einen Identitätsmarker bezüglich der *Nativität* bzw. der *Ethnizität*, des *Status'* und der *Individualität* für einen oder eine Gruppe von L2-Sprechern bilden.

2.3 Lernersprachen und Varietäten

Die Feststellung von allgemeinen Fossilisationstendenzen wie auch die Beispiele der soziolinguistischen Fossilisation deuten die Möglichkeit an, dass Eigenschaften in der Sprache mehrerer L2-Lerner nach ähnlichen *Patterns* fossilieren. Auch wenn fossilierte Merkmale zu Beginn individuell vorkommen, können sie nämlich unabhängig voneinander bei unterschiedlichen Lernern entstehen und sich im Nachhinein als Kennzeichnung der Sprache einer Lernergruppe stabilisieren, erst recht dann, wenn andere auch außerlinguistische Faktoren die Wahrscheinlichkeit der Kontakte, sowie die Identifikation zwischen den betroffenen L2-Sprechern erhöhen. Schon Selinker (1972) weist diesbezüglich auf eine gruppespezifische Dimension, in der eine Lernersprache erscheinen kann, und daher auf ihre mögliche Entwicklung als eine Art *Varietät* hin:

> „not only can entire IL competences be fossilized in individual learners performing in their own interlingual situation, but also in whole groups of individuals, resulting in the emergence of a new dialect" (Selinker 1972: 217).

Der Begriff „Dialekt" wird hier in einer allgemeineren Bedeutung als der heutigen verwendet (vgl. auch das nachfolgende Zitat von Labov 1971). Verstanden wird darunter gemeinhin eine

---

Stil: Sie markieren die soziale Identität eines Sprechers und werden bewusst verwendet. Schließlich nennt man die Merkmale, die als typisch für eine Varietät gelten und die man z.B. für die Nachahmung dieser Varietät auch außerhalb einer Sprachgemeinschaft benutzt, *Stereotype* (vgl. Dittmar 1997: 69–70 und Swann et Al. 2004).

Sprachvarietät (oder -lekt), und zwar eine mit außersprachlichen Variablen ko-okkurierende Menge von kongruenten linguistischen Merkmalen (vgl. Berruto 1995: 74–79), die sich von standardsprachlichen und anderen Varietäten regelhaft unterscheidet.[14] Von der Bewertung als „restringierte Kodes“ verzögert, beginnt eine systematische Auseinandersetzung mit den Non-Standard-Varietäten wenige Jahre vor der Formulierung der *Interlanguage*-Hypothese. Dabei ist eine gewisse Ähnlichkeit der Definition der „Non-Standard-Varietät“ in Labov (1971) mit der Definition der „Lernersprache“ bei Selinker (1972) zu bemerken:

> „Alle Linguisten stimmen darin überein, daß Nonstandard-Dialekte *hoch strukturierte Systeme* sind; sie sehen diese Dialekte *nicht als Anhäufungen von Fehlern* an, die entstanden sind, weil es ihnen Sprechern mißlingt, Standardenglisch zu beherrschen [...], [sie] erkennen an, daß [sie] *ein separates System* darstell[en], das eng mit dem Standardenglisch verbunden ist, aber [...] *durch eine Anzahl beständiger und systematischer Unterschiede abgesondert* ist“ (Labov 1971: 92–93. Hervorhebungen von mir).

Während die Annahme einer Differenz- statt einer Defizithypothese das allgemeine Interesse für die „defekte Sprache“ rechtfertigt (vgl. auch Kap. 4.3.2 und teilweise die im Kap. 2.1 dargestellte Selinker’sche Unterscheidung zwischen Lerner- und Lehrerstandpunkt), legitimiert ihre Auffassung als „separates System“, sie zum Gegenstand der Forschung zu machen. Eine solche Auffassung setzt ferner die Beobachtung ihrer Regelmäßigkeit in Abhängigkeit allgemeiner Bezugsdimensionen (wie Raum, Situation oder Gruppe) sowohl innerhalb eines Sprechers, als auch zwischen verschiedenen Sprechern (Intra-Sprecher und Inter-Sprecher) voraus.

Abgesehen vom obigen Zitat ist Selinker (1972) bezüglich einer gruppenspezifischen Dimension der *Interlanguages* wenig deutlich. Davon ausgehend, dass jede konkrete *Interlanguage* eine (Kontakt-)Varietät der Target-Sprache (ein „Iterimlekt“ oder eine „Lernervarietät“) bildet, stellen jedoch neuere Forschungen Zwischenstadien zwischen Lernersprache und Varietäten fest. Von diesem Standpunkt aus untersucht man sie als Idiolekte oder Soziolekte, indem man jeweils entweder ihre psychologische oder ihre gruppenspezifische Dimension hervorhebt (vgl. Dittmar 1997: 240–244 und Fußnote 14).

---

[14] Nach der heutigen Definition sind Dialekte diatopisch (in Bezug auf den Ort) definierte sprachliche Varietäten, wobei andere mögliche Kriterien diastratische oder diaphasische Variablen (wie den sozialen Status bzw. die Funktion und den Kontext) betrachten, um damit eine Varietätstypologie zu erstellen (für einen Ausblick über den Begriff „Varietät“ und die unterschiedlichen Untersuchungstraditionen vgl. Berruto 1995: 74–79, 174–152 und Dittmar 2008b; für ihre Abgrenzung vom Stil Kapitel 4.3). Interessant ist ferner die weitere Unterscheidung bei Dittmar (1997: 173 ff.) nach sechs Ordnungsdimensionen, und zwar nach Person, Raum, Gruppe, Gebrauchsfunktion, Situation und Kontakt. Zu den Kontakt-Varietäten gehören u.a. die sogenannten „Lernervarietäten“, welche „die *Interlanguages*, die bei der Erlernung der Zielsprache jeweils zur Verfügung stehen“ (240), bilden.

Prototypisches Beispiel solcher Zwischenstadien sind die sogenannten „Non-Native-Varieties“ einer Sprache (NNV): Sie entstehen zum Beispiel, wenn eine Sprache sich in einem nicht-nativen Kontext als Lingua Franca oder als Verkehrssprache durchsetzt (NNV sind z.B. die Varietäten des Englischen, die in ehemaligen Kolonien wie Indien und Nigeria gesprochen werden). Obwohl die Gruppendimension bei IL-Sprechern einen geringeren Institutionalisierungsgrad und eine geringere Stabilität als bei den NNV-Lernern aufweist, kennzeichnet Züngler (1989) den Prozess der Aneignung der Target-Sprache (oder *Nativization*) in beiden Situationen als einen individuellen sowie kollektiven:

> „Learners of English in Nigeria, sharing norms and often native languages, would, therefore, transfer what Nigerians as a group transfer. At the same time, though, one could argue that nativization is also an individual phenomenon, since the learner's degree of nativization would presumably be a function of his strength of cultural identity and the type of NNV styles he is acquiring. In IL settings, we can speak, as well, of nativization as both a group and an individual process [...] for the same reasons as in NNV“ (Züngler 1989: 92–93).

Sowohl die individuelle, als auch die kollektive Dimension werden zudem durch die Begriffe der „basic variety“ und der „fortgeschrittenen Lernervarietäten“ berücksichtigt, welche weitere Zwischenstadien zwischen Lernersprache und Varietäten darstellen. Diese Begriffe sind von der Forschung des ungesteuerten L2-Erwerbes in natürlichem Kontext bei erwachsenen ausgewanderten Arbeitern geprägt worden (vgl. die Studien von Perdue 1984 und Klein et al. 1992 im Rahmen von „The European Science Foundation Project“ (ESF)).

Die Basisvarietät ist eine erste einfache Form von *Interlanguage*, die sich für alle Lerner unabhängig von der Mutter- und der Targetsprache, dem Alter oder der Unterrichtsart nach allgemeinen kognitiven Prinzipien entwickelt und eine mögliche erste Stufe darstellt, auf der sich Sprachkompetenzen fossilieren können (vgl. Klein et al. 1992: 301–316, Vogel 1990: 47–50 und Walter & Grommes 2008: 10–14). Während die Eigenschaften der Basisvarietäten hauptsächlich vom Einfluss von *Spracherwerbsuniversalien* abhängen, weist die Entwicklung einer „fortgeschrittenen Lernervarietät“ das regelmäßige Auftreten von Strukturen auf, die sich auf die Ziel- oder die Ausgangssprache beziehen. Der „fortgeschrittene Lerner“ wird in diesem Zusammenhang als ein L2-Sprecher definiert, der die Basisvarietät hinter sich gelassen hat, aber dennoch nicht notwendigerweise eine dem Muttersprachler äquivalente Kompetenz erreicht hat (vgl. Walter & Grommes 2008: 12). Die in dieser Phase stattfindenden Fossilisationsprozesse sind i.d.R. Generalisierungen oder Transferphänomene, die sich teilweise nach ähnlichen Mustern

stabilisieren.[15] Die Hypothese der Entwicklung gemeinsamer Eigenschaften in der Sprache der interkulturellen Autoren ist auf dieser Basis und in Anlehnung an diesen Beispielen berechtigt.

2.4 Zwischenfazit

Die Begründung der Hypothese der Anwesenheit gemeinsamer Eigenschaften in der Sprache der interkulturellen Autoren war das Ziel dieses zweiten Kapitels. Dafür sprechen meiner Meinung nach zwei Arten von Gründen, welche eventuelle Ähnlichkeiten katalysieren könnten.
Auf der einen Seite steht der Zustand der interkulturellen Schriftsteller als L2-Lerner und daher die mögliche Betrachtung ihrer Sprache als eine Art *Interlanguage*. Da der Prozess des Erwerbs einer Fremdsprache tendenziell nach einer invarianten Abfolge geschieht bzw. die lernersprachliche Kompetenz nach bestimmten Tendenzen fossiliert, kann man ein ähnliches sprachliches Verhalten zwischen den Lernern vermuten. Die Tatsache, dass es unter den Erwerbsmechanismen auch Übergeneralisierungen von TL-Strukturen gibt, rechtfertigt ferner eine solche Annahme nicht nur bezüglich des Lernprozesses, sondern auch hinsichtlich des Vorkommens konkreter linguistischer Merkmale, insofern die Target-Sprache für alle interkulturellen Autoren das Deutsche ist und die eventuellen Abweichungen nicht nur vom Transfer abhängen. Schließlich betrachtet man Übergeneralisierungen aus der TL oftmals als kreative Vorgänge, sodass die davon bedingten Strukturen zu einer Verwirklichung des zweiten Modells der „sprachlichen Neusituierung“ zählen könnten (vgl. Ackermann 1997 und Kap. 1.3):

> „Übergeneralisierungen sind kreative Abweichungen von der Zielsprachennorm, die als Anzeichen für regelgesteuertes, exploratives und generatives Lernerverhalten gesehen werden“ (Vogel 1990: 49).

Auf der anderen Seite ist der Zustand der interkulturellen Autoren als „Minderheit“ relevant. Von verschiedenen Standpunkten aus kann man sie nämlich im Allgemeinen als nicht-native bzw. ethnische oder sozial stigmatisierte Gruppe erfassen (vgl. Esselborn 1997, aber auch die Untersuchungen im Rahmen des ESF-Projekts: Perdue 1984 und Klein et al. 1992) bzw. sie in Untergruppen je nach Nationalität oder Status aufteilen. Sozial und soziolinguistisch bedingt

[15]Obwohl ein Vergleich mit den fortgeschrittenen Lernervarietäten nur bedingt durchgeführt werden kann, stellen auch einige „urbane Multiethnolekte“ Beispiele einer nach gemeinsamen Pattern sich entwickelten Kontakt-Varietät. Diesbezüglich nennt das sogenannte „Kiezdeutsch“ eine Sprachvarietät, die vornehmlich Jungen und Jugendliche mit unterschiedlicher ethnischer Herkunft in urbanen multi-ethnischen Stadtteilen sprechen und die auf der Basis von Deutsch und fremdsprachlichen Einflüssen entsteht (vgl. Jannedy 2011).

könnten sich daher fossilierte Merkmale nach einem ähnlichen *Pattern* stabilisieren und eventuell je nach einer der verschiedenen Identitätseigenschaften variieren.

Eine genauere Erörterung des von mir untersuchten Merkmales führe ich im nächsten Kapitel durch. Wie schon vorweggenommen wurde, bezieht sich der vermutete sprachliche Unterschied auf den Produktivitätsgrad einiger Wortbildungsmuster des Deutschen und insbesondere auf deren Übergebrauch seitens der interkulturellen Autoren.

Daher sind zwei weitere Frage zu erwähnen, die später besprochen werden. Zuerst kann man sich fragen, ob der Übergebrauch eines Wortbildungsmusters ein lernsprachliches fossiliertes Phänomen darstellen könnte. Fossilierte Phänomene können diesbezüglich auf allen linguistischen Ebenen erscheinen (vgl. Tarone 1994: 1716) und insbesondere gehört die kreative Wortbildung zu den Elaborationsstrategien, die Anfänger und Fortgeschrittene beim Lernen der L2 verwenden, um kommunikative Probleme zu lösen (vgl. González Àlvarez 2004 und Kap. 3.4). Hinsichtlich der Frage nach dem Einfluss soziolinguistischer Faktoren stellt die Forschung darüber hinaus Variationen im Produktivitätsgrad von Wortbildungsmustern je nach Register, sozialer Schicht und Region fest (vgl. Baayen 1994a, Keune et al. 2006, Säily 2011 und Kap. 3.3.3).

Ebenso kann man die Betrachtung der interkulturellen Autoren als (fortgeschrittene) Lerner als gezwungen empfinden: Einerseits insofern die meisten seit vielen Jahren in Deutschland leben und ihre Sprachbeherrschung sehr gut ist, andererseits insofern ihre für die Veröffentlichung erstellten Texten hoch elaboriert sind. Von der großen Verbreitung der Fossilisationsphänomene bei L2-Lernern ausgehend datieren aber einige Linguisten die kritische Periode für den Spracherwerb auf die ersten zwölf Lebensjahre, nach denen man nicht oder nur mit großen Schwierigkeiten eine TL-äquivalente Kompetenz in einer L2 erreichen kann (vgl. Selinker 1972: 210 Fußnote 4 und Kap. 5.1.1). Darum habe ich meinen Korpus auf Autoren beschränkt, die nach dem 12. Lebensjahr nach Deutschland ausgewandert sind.

Die sorgfältige Bearbeitung der Texte und ihr mögliches Lektorat durch Muttersprachler schließt Besonderheiten im Gebrauch einer lernersprachlichen Struktur nicht aus, vor allem wenn sie nicht als Fehler zu erkennen sind. Diesbezüglich kann man eine Parallele zu der neueren Übersetzungsforschung ziehen, die sich ebenfalls mit Sprachkontaktphänomenen im weiteren Sinne beschäftigt. Ihnen zufolge bildet die Übersetzungssprache eine Art „dritten Kode“, der sich nicht nur durch Transfer, sondern auch durch spezifische Übersetzungseffekte von der Ausgangs- und der Targetsprache unterscheidet, z.B. durch Über- oder Unterrepräsentation von linguistischen Merkmalen oder durch allgemeine Simplifizierungsprozesse. Baroni & Bernardini (2006) etwa stellen gemeinsame *translationese*-Merkmale in geschriebenen, hoch elaborierten und veröffentlichen Texten fest. Auch ihre Ergebnisse rechtfertigen meine Fragestellung.

3. Die zweite Unterhypothese: Merkmal und Möglichkeit einer Asymmetrie im linguistischen Verhalten von interkulturellen und deutschen Schriftstellern

Auf Basis einer quantitativen Analyse der Derivate auf *-ung* in einem Lernerkorpus (*Falko*) und eines Vergleichs mit deren Gebrauch in Texten von Muttersprachlern[16] zieht Berth (2009) in seiner Magisterarbeit einige Schlussfolgerungen über den Erwerb der Wortbildungsregularitäten und die morphologische Kompetenz von fortgeschrittenen Lernern des Deutschen als Fremdsprache. Obwohl sich die *potential productivity* der Wortbildungsmuster auf *-ung* (vgl. Kap. 3.3.2) größer bei den Muttersprachlern als bei den L2-Lernern zeigt, weist Berth wie folgt darauf hin:

> „dass Lerner komplexe Neubildungen produzieren, die ungrammatisch sind, und die also von Muttersprachlern aller Wahrscheinlichkeit nach nicht produziert würden, was bedeutet, dass ein Lerner theoretisch mehr -ung-Suffigierungen bilden könnte“ (Berth 2009: 55).

Angesicht dessen könnte der Gebrauch von sehr produktiven Wortbildungsmustern des Deutschen durch interkulturelle Schriftsteller Besonderheiten aufweisen, die ihre Texte im Vergleich zu denen der deutschen Autoren kennzeichnen und so einen Aspekt des von Ackermann (1997: 26) behaupteten „kreativen Umgangs“ mit der deutschen Sprache bilden. Die Anwesenheit signifikanter Gebrauchsunterschiede wiederum würde auf eine Asymmetrie in der morphologischen Kompetenz bzw. in der Haltung bezüglich der pragmatischen und strukturellen Bedingungen der Wortbildung zwischen den interkulturellen Autoren als L2-Lernern und den deutschen als Muttersprachlern verweisen.

Indem ich die Produktivität eines Affixes qualitativ als seine Fähigkeit neue Wörter zu schaffen und quantitativ als die Wahrscheinlichkeit des Gebrauchs in der Sprachproduktion definiere (vgl. aber Kap. 3.3), stelle ich in den nachfolgenden Kapiteln den Begriff „Wortbildungsmuster“ (Kap. 3.1) und die allgemeinen linguistischen Faktoren dar, welche für eine größere Bereitschaft zur Anwendung der Wortbildungsstrategien seitens der L2-Lerner sprechen (Kap. 3.2). Anschließend erläutere ich mit Bezug auf einige empirische Studien den in dieser Arbeit angenommenen Begriff der morphologischen Produktivität und die davon vorausgesetzte Struktur des mentalen Lexikons genauer (Kap. 3.3). Diese Darlegung, die teilweise von der Erörterung meiner Fragestellung wegführt, begründet die für meine Hypothese wesentliche Annahme einer Variation des Gebrauchs

[16] *Falko* ist ein im Internet frei zugängliches fehlerannotiertes Lernerkorpus des Deutschen als Fremdsprache. Es besteht aus drei Subkorpora: ein „Zusammenfassungskorpus“ und ein „Essaykorpus“ mit Textzusammenfassungen bzw. Aufsätze über vorbestimmte Themen von fortgeschrittenen DaF-Lernern und ein „Longitudinalkorpus“ mit Daten von DaF-Studenten der Georgetown University auf unterschiedlichen linguistischen Niveaus. Vergleichskorpora mit kontrollierten Texten von Muttersprachlern stehen dazu jeweils zur Verfügung (vgl. http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/falko/standardseite. Zuletzt geprüft am 10.11.2011).

der produktiven Wortbildungsmuster zwischen Sprechern, indem sie auf psycho- und soziolinguistische Variablen zurückführt.
Schließlich betrachte ich das Problem des Erwerbs und der Beschaffenheit der morphologischen Kompetenz von L2-Sprechern. Dabei werden auch einige mögliche Parallelen mit dem Verhalten von L1-Lernern gezogen (Kap. 3.4).

3.1 Die Wortbildungsmuster

Den Begriff „Wortbildungsmuster" übernehme ich von Motsch (2004). Dieser stellt sie „als Paare von semantisch-syntaktischen und phonologisch-morphologischen Beschreibungen" (1) dar, die als „Muster für die Analyse komplexer Wortzeichen" (1) fungieren und die „mögliche[n] Wortbildungen einer Sprache vorher[sagen]" (18). Insofern sie mit weiteren systematischen sowie idiosynkratrischen Informationen in den lexikalischen Einträgen komplexer Wörter gespeichert werden, sind die Wortbildungsmuster laut Motsch „Bestandteil[e] der Lexikon-Komponente der Grammatik" (2004: 1).
Im Gegensatz zu anderen Auffassungen im Rahmen der Generativen Grammatik, welche die „Wortbildungsregeln" als in die Grammatik eingegliedert betrachten, weist eine solche „lexikongebundene" Annahme den Vorteil auf, die „Besonderheiten" der zusammengesetzten Wörter im Vergleich zu den durch syntaktische Regeln generierten Strukturen zu erfassen (Motsch 2011: 53–59). Unter diesen Besonderheiten befinden sich die Akzeptabilitätsbeschränkungen regelmäßig gebildeter Ausdrücke bzw. die unscharfe semantische Kategorisierung der Basis oder die Geltung einiger Wortbildungsregularitäten als Tendenzaussagen, die auch in den nachfolgenden Kapiteln besprochen werden (vgl. Kap. 3.2 und Kap. 3.3.3).
Der Begriff „Wortbildungsmuster" gehört jedenfalls nach Motsch zu der Beschreibungsebene der theoretischen Linguistik. Er erhebt daher weder den Anspruch auf psychologische Realität noch kann er die erwähnten Besonderheiten genau erklären (2004: 25). Indem sich „Wortbildungsfakten der angeführten Art [Akzeptabilitätsbeschränkungen, unscharfe semantische Kategorisierung der Basis usw.] [...] auf die Anwendung von Regeln, nicht auf die Regeln selbst" beziehen, benötigt nämlich deren ausführlichere Erörterung sprachpsychologische und kommunikationspragmatische Betrachtungen, während die Wortbildungsmuster nur die möglichen morphologischen Strukturen einer Sprache in ihren allgemeinen Zügen beschreiben (Motsch 2011: 59).[17]

[17] Anders gesagt: die Auflistung von Wortbildungsmustern betrifft die Unterscheidung von *möglichen* und *unmöglichen* Bildungen einer Sprache, während die Bestimmung der Gründe und Bedingungen ihres *Vorkommens* (oder *Nicht-Vorkommens*) innerhalb der möglichen Wörter sprachpsychologische und kommunikationspragmatische Betrachtungen

Die Unterscheidung zwischen diesen zwei Standpunkten führt Motsch zu einer Definition der Wortbildungsstrategien als psychologischen Realitäten, die mit der in den nachfolgenden Kapiteln dargestellten Perspektive von Baayen und seinen Kollegen Gemeinsamkeiten aufweist (Kap. 3.3 und ff.). Wenn nämlich das Konzept der Wortbildungsmuster in einer grammatikalisch-deskriptiven Hinsicht die Anwesenheit von vorbestimmten bzw. *deterministischen* semantischen und phonologischen Modellen für die Zusammenstellung *diskreter* morphologischer Einheiten annimmt, sind dieselben dabei „aus psychologischer Sicht nur potentielle Regeln […], die im Bedarfsfall aus Lexikoneinheiten mit gemeinsamen Eigenschaften abstrahiert werden müssen" (Motsch 2004: 24). Unabhängig von ihrem Produktivitätsgrad (in der Motsch'schen Terminologie „Aktivitätsgrad") und ihrer Systematizität unterliegen nämlich alle Wortbildungsmuster demselben psychologischen Mechanismus der Generalisierung aufgrund des gespeicherten mentalen Lexikons, während die als graduell erfassten Aktivitätsunterschiede einerseits von der semantischen Transparenz und der pragmatischen Nützlichkeit des Musters, andererseits von quantifizierbaren Variablen wie der Größe der morphologischen Familien oder der Häufigkeit ihres Auftretens abhängen (vgl. Motsch 2004: 21–23 und Kap. 3.3). Daher ist der Motsch'sche Begriff von „Wortbildungsmuster" mit der in den nachfolgenden Kapitel dargestellten Theorie der morphologischen Produktivität bei Baayen weitestgehend vereinbar. Ich verwende ihn hier durchaus als Synonym des Baayen'schen Ausdrucks „morphological category", der als „sets of words sharing aspects of form structure and aspects of meaning" definiert wird (Baayen 2009: 900), um die gemeinsame Struktur bestimmter Wortbildungen zu erfassen.

Außerdem ist die Motsch'sche Unterscheidung zwischen einer grammatikalischen und psycholinguistischen Betrachtungsweise der Wortbildungsregularitäten deshalb interessant, weil sie sich als besonders anwendbar auf den Zweitspracherwerbsprozess erweist. Im Gegensatz zu einem L1-Lerner erwirbt nämlich ein L2-Sprecher als Erwachsener die Target-Sprache sowohl unbewusst durch die Interaktion in einem natürlichen Kontext mit den Muttersprachlern (*acquisition*), als auch bewusst im Sprachunterricht durch Grammatiken (*learning*), die oft auf linguistische „Regeln" und ihre Ausnahmen hinweisen (für einen Ausblick über verschiedene Ansätze diesbezüglich vgl. Vogel 1990: 75–98). Daher hat ein L2-Lerner Zugang zum Erwerb der produktiven Wortbildungsstrategie einer Sprache, sowohl durch die Beobachtung des linguistischen Verhaltens der Muttersprachler und die Verarbeitung des umliegenden sprachlichen Inputs, als auch durch deren grammatikalische Definition bzw. ein systematisches Studium der Fremdsprache. Eine Asymmetrie in dem Gewicht

---

benötigt. Die These, dass diese zwei Ebenen der linguistischen Beschreibung des Lexikons auch unterschiedliche morphologische Kompetenzen eines Sprechers widerspiegeln, vertritt z.B. Coseriu 1975, der das Wissen über die funktionellen und konventionellen morphologischen Oppositionen unterscheidet. Der Mangel der vollständigen Kenntnis der konventionellen morphologischen Oppositionen kann einen L2-Lerner auf der Basis seiner Kenntnis der funktionellen morphologischen Oppositionen zu Generalisierungen verleiten (vgl. Kap. 3.4).

dieser Strategie beim Spracherwerb könnte die Generalisierung einiger Wortbildungsmuster seitens der L2-Lerner begründen.

3.2 Anlässe und Beschränkungen der Wortbildung als Bedarfsystem

Aufgrund der Annahme, dass „words are only formed as and when there is a need for them“ (Bauer 2001: 143), bespricht die Literatur die genauen Anlässe und Beschränkungen für die Verwendung des Wortbildungssystems einer Sprache (vgl. u.a. Motsch 2004: 17–26, Plag 1999: 37–61, Rainer 2004, Fleischer et al. 2007: 59–60). Obwohl sie im Allgemeinen gelten, erweisen sich einige von ihnen als im Zweitspracherwerb besonders aktiv.

Unter den Neubildungsanlässen finden sich auf die lexikalische Erweiterung gerichtete pragmatische Gründe, die Motsch von denen der syntaktischen Umkategorisierung bzw. der stilistischen Variation differenziert (2004: 20–22). Insofern die Wortbildungsstrategien den Aufbau neuer Wörter aus schon vorhandenem sprachlichem Material ermöglichen, verwendet man sie nämlich u.a. als eine Art „Bedarfssystem“, um absolute oder relative Löcher im Wortschatz zu füllen, und zwar wenn ein Wort tatsächlich in einer Sprache oder in einer *Interlanguage* fehlt bzw. als dem Kommunikationsbedarf des Sprechers nicht-genügend wahrgenommen wird. Da die L2-Lerner durchschnittlich ein ärmeres Lexikon als die Muttersprachler besitzen, aber dieselben Ausdrucksbedürfnisse aufweisen, kann man erwarten, dass für sie die Gelegenheiten häufiger und die Notwendigkeit zur der Bildung neuer Wörter größer sind:

> „This ‚list of lexemes‘ – which includes the whole lexicon of established language – will be considerably shorter for L2 learners due to their limited IL lexicon. Therefore, we should expect them to feel the need to coin new lexeme more often than native speaker of a language“ (González Álvarez 2004: 35).

Bei den Wortbildungseinschränkungen befinden sich außerdem auf Regeln nicht reduzierbare *Akzeptabilitätsbeschränkungen*, infolge deren der Sprecher regelmäßig zusammengesetzte Bildungen als nicht-akzeptabel bewertet. Sie unterscheiden sich nach Plag (1999) in *pragmatic* und *general structural restrictions* (vgl. auch Rainer 2004).[18] Während Erstere z.B. die Bildung

[18] Indem Plag (1999) und Rainer (2004) eigentlich von „productivity restrictions“ bzw. „Produktivitätsbeschränkungen“ sprechen, ist ihr Ausdruck „begrifflich an Theorien geknüpft, die eigene Wortbildungsregeln enthalten, an denen Beschränkungen ‚festgemacht’ werden können“ (Rainer 2004: 880). Daher und aufgrund meines nachfolgenden Bezugs auf die Baayen’sche Theorie, die die morphologische Produktivität ohne Annahme von Regeln und *Constraints* erklärt, verwende ich hier den allgemeineren Ausdruck der Akzeptabilitätsbeschränkungen. Die Frage, ob sie die Bildung oder eher die Verwendung eines neuen Wortes einschränken, bleibt deswegen offen. Rainer (2004) nimmt darüber hinaus

unsinniger Wörter (**köpfig*) verhindern und vermutlich gleich auf L2-Lerner und Muttersprachler wirken, könnte eine allgemeine Beschränkung, wie das *Blocking*, gemäß der Variable *Nativeness* ein unterschiedliches Gewicht aufweisen. Von Blockierung spricht man, wenn die Anwesenheit eines schon lexikalisierten Wortes die Bildung oder den Gebrauch eines regelmäßig zusammengesetzten Synonymes bzw. teilweise Homonymes sperrt (**Stealer* ~ *Thief*). Die Blockierung, die man in dieser Form auch als *Tokens-Blocking* im Gegensatz zu der hier nicht betrachteten *Typen-Blocking* definiert,[19] ist ein psycholinguistisches Phänomen, „the effect of word-storage and word-processing mechanism", das also mit der Wortfrequenz und der Beschaffenheit des mentalen Lexikons zusammenhängt (Plag 1999: 52).

Aufgrund des *Tokens-Blocking* gründen Fleischer et al. (2007: 59) die Akzeptabilität als Bereitschaft des Sprechers, neue Lexeme anzunehmen, zum Teil „auf den bekannten/geläufigen Wortschatz und seinen Strukturen", während Plag (1999) die Korrelation der Blockierung mit der Frequenz der betroffenen Wörter und der Stärke ihrer mentalen Repräsentationen beweist. Die daraus abzuleitenden Konsequenzen in der SLA bespricht Berth (2009). Davon ausgehend, dass L2-Lerner weder einen Target-gleichen Wortschatz, noch semantische Relationen zwischen den gespeicherten Wörtern vollständig hergestellt haben, nimmt man bei ihnen eine Senkung der Blockierungskraft im Vergleich zu den Muttersprachlern an:

> „Diese Art von Lernerfehlern [Schwangerschaftsabbrechung statt Schwangerschaftsabbruch] ist in gewissem Maße erwartbar, da bei fortgeschrittenen Lernern angenommen werden kann, dass sie die regelhafte Bildung von -ung Derivaten verinnerlicht haben, aber noch nicht denselben Wortschatz wie Muttersprachler erworben haben, sodass die Blockierung in vielen Fällen nicht wirksam werden kann" (Berth 2009: 23).

Um ein unterschiedliches Verhalten zwischen Lernern und Muttersprachlern bezüglich der Verwendung des Wortbildungssystems zu vermuten, gelten schließlich die nicht-regelspezifischen Beschränkungen als dritter Grund, welche die syntaktische und semantische Kategorie der Basis eines Wortbildungsmusters bestimmen (vgl. Plag 1999: 45–46). Insofern ihre semantische Subkategorisierung oft unscharf ist (Motsch 2004: 5–6), könnte nämlich die *Interlanguage* eines

---

eine von Plag (1999) unterschiedliche Terminologie an: Er spricht nämlich von einzelsprachigen und übereinzelsprachigen Aspekten, die sich von den regel-spezifischen Beschränkungen (nach Plag 1999: *structural rule-specific restrictions*) unterscheiden.

[19] Mit *Typen-Blocking* bezeichnet man die Blockierung eines regelmäßigen morphologischen Prozesses durch einen anderen funktionsgleichen, aber mit bestimmterem Anwendungsbereich: ?*decentness* ~ *decency* (vgl. Plag 1999: 50–54 und Rainer 2004). Da tatsächliche semantische Rivalen unter den Affixen fehlen bzw. Ausnahmen häufiger sind, ist das Vorkommen des *Typen-Blocking* umstritten und wird hier nicht in Betracht gezogen (vgl. Plag 1999). Plag 1999 bestreitet darüber hinaus den Bestand eines *Tokens-Blocking* wegen Homonymie, insofern eindeutige Kontexte seine Wirkung aufheben würden.

Lerners auch infolge der daraus resultierenden Erwerbsschwierigkeiten *nicht-Target-gleiche Basis-Beschränkungen* für einige Affixe aufweisen.

Zu bemerken ist jedoch, dass dieser letzte Aspekt mit Vorsicht zu betrachten ist. Im Gegensatz zu den zwei oben erwähnten (dem *nicht-Target-gleichen Tokens-Blockierung* und dem *pragmatischen Bedürfnis*) begründen *nicht-Target-gleiche Basis-Beschränkungen* nämlich eine abweichende Verwendung der Wortbildungsstrategien seitens eines Lerners, aber keine Tendenz zu ihrer Generalisierung, insofern sie sowohl einen Über- als auch einen Mindergebrauch des Wortbildungssystems verursachen könnten. Darüber hinaus treten sie seltener und unsystematischer auf: L1- und L2-Lerner halten nämlich nach González Álvarez (2004: 43) die grammatikalischen Kriterien der Affigierung normalerweise ein und nur gelegentlich selegieren sie eine für einige Affixe grammatikalisch unverträgliche Basis.

Ich gehe davon aus, dass aus einem oder aus dem Zusammenwirken dieser drei Faktoren eine Asymmetrie, wenn nicht gar ein Übergebrauch des deutschen Wortbildungssystems durch die interkulturellen Autoren folgen könnte. Trotzdem nehme ich eine genauere Bestimmung des Grundes Fall für Fall in meiner Masterarbeit nicht vor. Die dafür verlangte qualitative Analyse der vorkommenden Derivate wäre übrigens aufgrund der Beschaffenheit meiner Daten vermutlich nicht immer schlüssig (vgl. Kap. 4 und Kap. 5). Wenn man z.B. in einer psycholinguistischen experimentellen Untersuchung zusätzliche Informationen mit Hilfe bestimmter *Tasks* über die Prozesse der lexikalischen Verarbeitung und den Reichtum des Wortschatzes eines Probanden entnehmen kann, hat eine Korpus-basierte Analyse, die jedoch mit „authentischen" Daten arbeitet (Kap. 4.1), auf solche Auskünfte keinen Zugriff.

Daher ist die Feststellung der verwendeten Strategie bzw. der von einem Autor tatsächlich neu gebildeten Wörter schwieriger. Als Teillösung setzt man in der SLA-Forschung oft die morphologisch relevanten Strukturen den Fehlern gleich, insofern sie am wahrscheinlichsten Neubildungen[20] darstellen, von denen man dann die psychologischen bzw. pragmatischen Bedingungen und Gründen untersuchen kann (vgl. z.B. Berth 2009).

Da mein Korpus aus hoch elaborierten und fast fehlerfreien Texten besteht, erweist sich diese Methode als schwer anwendbar. Die Feststellung der tatsächlichen Neubildungen und ihrer Gründe ist darüber hinaus – wenn auch ideal wünschenswert – für meine Fragestellung nicht zwingend

[20] Unter dem Ausdruck „Neubildung" fasst man zwei aus einem theoretischen Standpunkt unterschiedliche Begriffe zusammen. Auf der einen Seite bilden nämlich die *Neologismen* die neuen Lexeme, die bereits in das Lexikon einer Sprachgemeinschaft aufgenommen wurden. Auf der anderen Seite stellen die *Textwörter* oder *Occasionalismen* in einem Text ad hoc zusammengesetzte Neubildungen dar, die nur einmalig auftreten (vgl. Motsch 2004: 19 und Berth 2009: 24–26). Insofern sie qualitative Ermittlungen und lexikografische bzw. sprachhistorische Untersuchungen benötigt, ist nicht nur die empirische Auseinanderhaltung dieser zwei Unterkategorien, sondern auch die allgemeine Feststellung der tatsächlichen Neubildungen eines Textes in einer quantitativen Analyse problematisch. Daraus folgt die Entwicklung von indirekten Methoden für die Auswertung der Anzahl von Neubildungen eines Textes, die man nämlich anhand der Hapax Legomena einschätzt (vgl. Baayen 1993, 1994b, 2009).

notwendig. Einerseits ist nämlich der Produktivitätsgrad eines Wortbildungsmusters nicht nur durch die absolute Anzahl seiner Neubildungen, sondern z.B. auch durch die seiner Typen und Tokens bzw. ihres Verhältnisses untereinander zu erfassen. Anderseits stellt Baayen (1993, 1994b, 2009) die Anzahl der *Hapax Legomena* – der Wörter, die nur einmal in einem Korpus vorkommen – als indirekte, aber praktikable Schätzung der Neubildungen dar, um die Maße der *expanding* und *potential productivity* einer morphologischen Kategorie ohne eine qualitative Analyse bzw. auch in kleinen Korpora zu berechnen (vgl. Kap. 3.3). Aus der unterschiedlichen Zusammensetzung meines und des Berth'schen Korpus' folgt schließlich, dass ein Vergleich der Ergebnisse unserer quantitativen Untersuchungen nur teilweise und mit Blick auf die verschiedenen Produktionssituationen der betrachteten Texte möglich ist.

3.3 Die Produktivität als variables Maß

Wenngleich der Mangel an bestimmten Beschränkungen und das pragmatische Bedürfnis die Hypothese eines Unterschieds zwischen L2-Lernern und Muttersprachlern in dem Gebrauch des Wortbildungssystems rechtfertigen, reichen sie weder aus, um die Art, noch um die genaue Bestimmung der dazu verwendeten Wortbildungsstrategien zu erklären. Dem Bedarf der Wortbildung könnte ein L2-Lerner nämlich mit fortwährend wechselnden Bildungsmethoden nachkommen, sodass die sprachliche Asymmetrie in einer Vermehrung idiosynkratrischer analogischer Neubildungen bestehen würde. Diesbezüglich unterscheidet man zwischen „konzeptuellen" und „linguistischen" Wortbildungsstrategien. Unabhängig davon, ob sie zwei verschiedene Mechanismen oder die gegenteiligen Spitzen einer Skala (Motsch 2004) benennen, bezeichnen Erstere die durch analogische und idiosynkratrische Prozesse realisierten Wortbildungen, während Letztere, die man in der Literatur entweder als Ergebnisse morphologischer Regeln oder als Generalisierungen auf Basis des schon gespeicherten mentalen Lexikons beschreibt, sich auf die sogenannten produktiven Wortbildungsmuster einer Sprache zurückführen lassen (vgl. auch Kap. 3.1).

Nicht nur wegen der erwähnten Schwierigkeiten in der Feststellung der Neubildungen durch quantitative Methoden, sondern auch aus positiven Gründen nimmt meine Hypothese einen Unterschied im Gebrauch gerade der sogenannten produktiven Wortbildungsmuster zwischen interkulturellen und deutschen Schriftstellern an. Die erste Voraussetzung dafür bildet die Baayen'sche Theorie der morphologischen Produktivität (vgl. u.a. Baayen 1993, 1994b, 2009, Hay

& Baayen 2005, Keune et al. 2006), indem sie die Möglichkeit einer nicht nur diachronen, sondern auch synchronen Variation des Produktivitätsgrades einer morphologischen Kategorie begründet. Bezüglich einer zeitlichen Variation der Produktivität ist sich die Literatur durchaus einig und baut darauf z.B. die Unterscheidung zwischen den noch produktiven und den zu einem Zeitpunkt einfach generalisierten (*generalized*) morphologischen Prozessen auf (vgl. u.a. Motsch 2004, González Álvarez 2004: 61). Ihre synchrone Variation erklärt insbesondere Baayen ausführlich, indem er sie auf psycholinguistische bzw. soziolinguistische und stilistische Faktoren zurückführt. Infolgedessen erfasst Baayen den Produktivitätsgrad einer morphologischen Kategorie nicht als Produkt kategorialer Regeln, sondern als variables Ergebnis der Wechselwirkung struktureller und vor allem kognitiver und interaktioneller Faktoren:

> „If structural constraints as such do not directly drive morphological productivity […], research on morphological productivity should be directed toward other factors. There are two clusters of such factors that play a demonstrable role, one cluster pertaining to societal factors, the other to the role of processing constraints in the mental lexicons of individual speakers“ (Baayen 2009: 907).

In Anlehnung an die Baayen'sche Theorie verhandle ich in den nachfolgenden Kapiteln einerseits die für die Kodierung und den Erwerb der morphologischen Produktivität zuständigen prozessualen Prinzipien und *Constraints* (vgl. Kap. 3.3.1 und 3.3.2), anderseits stelle ich einige Studien vor, die Variation der Produktivität je nach Texttyp, Region, Alter, Geschlecht und Schulbildung beweisen (Kap. 3.3.3). Dass einige solcher Faktoren einen Einfluss auf die *Interlanguage* eines Lerners bzw. auf die Sprache der interkulturellen Autoren haben könnten, lässt das nachfolgende Zitat von Baayen (2009) vermuten:

> „morphological productivity can be understood as resulting from a great many factors such as the individual language user's experience with the words of her language, her phenomenal memory capacities, her conversational skills, her command of the stylistic registers available in her language community, *her knowledge of other language, her communicative needs, her personal language habits and those of the people with which she interact*“ (Baayen 2009: 901. Von mir hervorgehoben).

3.3.1 Die Baayen'sche Produktivität und die probabilistische Morphologie

Indem er die oben erwähnte Unterscheidung zwischen den qualitativen und quantitativen Aspekten des Produktivitätsbegriffs voraussetzt, spricht Bauer (2001: 205–213) von zwei Attributen für die Beschreibung eines produktiven Wortbildungsprozesses: seine „Verfügbarkeit" (*availability*) und seine „Ergiebigkeit" (*profitability*). Während ein „morphological process is available if it can be used to produce new words as they become necessary" (205), spiegelt ihre Ergiebigkeit nämlich „the extent to which its availability is exploited in language use" (211) wider, denn „possible words are not all equally probable" (207). Angesichts dessen erfasst Bauer die erste Eigenschaft als für jedes Wortbildungsmusters bzw. jede Wortbildungsregel anwendbare binäre Bestimmungskategorie, wohingegen die zweite ein stufenartiges Merkmal bildet, das eher von dem üblichen Sprachgebrauch, der Anwesenheit rivalisierender Strategien sowie dem pragmatischen Bedürfnis, und weniger von dem grammatikalischen Sprachsystem bedingt ist.

Den Begriff „Verfügbarkeit" zusammen mit einer absoluten Unterscheidung zwischen produktiven und unproduktiven Wortbildungsmustern bezweifelt jedoch Baayen (2009), ausgehend davon, dass ein einziges kognitives Prinzip der „exemplar-driven analogical generalisation" (912) allen morphologischen Bildungsprozessen zu Grunde liegt (vgl. auch Motsch 2004 und Kap. 3.1). Die Baayen'sche Annahme hat ihre Grundlage in einer probabilistischen Theorie der morphologischen Strukturen, nach der diese nicht aus der Anwendung kompositioneller bzw. deterministischer Regeln folgen, sondern „emerge[s] from the statistical regularities that characterize the forms and the meanings of word" (Hay & Baayen 2005: 342). Die Produktivität einer morphologischen Kategorie hängt daher von den Regelmäßigkeiten ihres Auftretens bzw. der Wahrscheinlichkeit der Erkennung ihres Musters seitens des Sprechers ab, die nämlich die Stärke und die Art der mentalen Repräsentationen ihrer Vollformen, Stämme und Affixe beeinflussen.

Notwendig sind in diesem Zusammenhang weitere Erklärungen, welche die Bedingungen und die Folgen dieser Annahme darstellen. Erstere beziehen die im Baayen'schen Modell vorausgesetzte Beschaffenheit der mentalen Zugriffsrepräsentationen der lexikalischen Einträge ein. Obwohl die experimentellen Beweise[21] für die graduelle Natur der morphologischen Strukturen im Rahmen einer Theorie wie etwa der *Word and Paradigm Morphology* die Begriffe „Morphem" und

[21] Belege für die graduelle Natur der morphologischen Strukturen bietet u.a. die Anwesenheit von *affix-ordering constraints*, die man nämlich am besten durch die unterschiedlichen Komplexitätsgrade der betroffenen Affixe vorhersagen kann. Diesbezüglich folgen in Fällen von mehrfacher Affigierung innerhalb eines Wortes die am schwierigsten zu verarbeitenden Affixe, die auch die produktivsten bilden, auf die einfacheren und weniger produktiven. Damit beweist man, dass „the gradient behaviour in experimental tasks reflects gradience [...] in underlying structure" (Hay & Baayen 2005: 343).

„morphologische Regel“ in Zweifel ziehen bzw. stattdessen die Vollform selbst als basale Verarbeitungseinheit bestimmen, nehmen Hay & Baayen (2005) eine Zwischenposition ein.

Ihres Erachtens besitzen nämlich nicht nur Vollformen, sondern auch Stämme und Affixe ihre eigene Zugriffsrepräsentation im mentalen Lexikon, wobei ihre Existenz sich als graduell bzw. probabilistisch erweist, indem sie von paradigmatischen Relationen und lexikalischen Verarbeitungsprozessen bedingt ist. Auf die Begriffe „Morphem“ und „morphologische Struktur“ kann man nämlich nicht ohne Weiteres verzichten, wenn man etwa betrachtet, dass Morpheme bei Sprachproduktion und -rezeption *Priming*-Effekten unterliegen bzw. ihre Salienz innerhalb des zusammengesetzten Wortes das Komplexitätsurteil des Sprechers und die Stärke der paradigmatischen Relationen beeinflusst.

Die zweite Frage betrifft die genauen Faktoren, die nach Baayen und seinen Kollegen die Stärke der mentalen Repräsentationen und daher die Produktivität einer morphologischen Kategorie bestimmen. Diesbezüglich spielt traditionell die Wortfrequenz eine wesentliche Rolle. Ausgangspunkt ist die Annahme, dass die linguistische Erfahrung andauernde Spuren im Gedächtnis hinterlässt: je häufiger eine Form im Input vorkommt, desto präsenter wird daher ihre mentale Repräsentation. Wenn man sie als reine Wortwiederholung versteht (Oberflächenhäufigkeit), stellt jedoch die Vorkommenshäufigkeit nach Baayen eine, aber nicht die einzige Variable dar, die den Verarbeitungs- und Erwerbsprozess beeinflusst. Hingegen muss man zwischen verschiedenen Frequenzmaßen unterscheiden, deren Bezugsbereich und Wirkung nur im Rahmen eines Modells des mentalen Lexikons zu verstehen sind.

Diesbezüglich nimmt Baayen (1993, 1994b) anfänglich ein *dual-Route*-Modell mit lexikalischem Eintrag in Form von eines Logogen sowie mit einer dynamischen Komponente an (das *Morphological Race Model* von Frauenfelder & Schreuder, vgl. Schriefers 1999: 124–127), um die Korrelation von Frequenz und Produktivität zu begründen. Aufgrund dessen geschieht der Zugriff eines komplexen Wortes sowohl auf einem direkten, als auch auf einem indirekten Weg, indem der Sprecher den Eintrag der Vollform ohne Weiteres wiederfindet bzw. die Bildung zerlegt und morphologisch analysiert. Beide Wege werden bei der Sprachproduktion und -rezeption gleichzeitig aktiviert; den Erfolg der einen oder der anderen Strategie bestimmt aber der Ruheaktivationsgrad der Vollform einerseits, des Stammes und der Affixe andererseits.

Wesentlich sind daher für die Verarbeitungsprozesse nicht nur das Maß der Oberflächen-, sondern auch der kombinierten Wortstamm- und der Affixhäufigkeit,[22] die nämlich jeweils den

[22] Die *Oberfächenhäufigkeit* definiert man genauer als die Vorkommenshäufigkeit einer bestimmten flektierten Wortform, während die *kombinierte Wortstammhäufigkeit* und die *Affixhäufigkeit* jeweils die Vorkommenshäufigkeit eines Stamms über alle flektierten, abgeleiteten und zusammengesetzten Formen bzw. die eines bestimmten Affixes benennen (Schriefers 1999: 136).

Aktivationsrest der entsprechenden mentalen Repräsentationen beeinflussen. Die Konsequenzen dieser Perspektive auf die morphologische Produktivität kann man mit einem Beispiel verdeutlichen: Wenn eine Bildung relativ selten im Vergleich zu ihrer Basis auftritt, verarbeitet der Sprecher sie wahrscheinlicher auf dem indirekten Weg der morphologischen Analyse. Die von diesem Zugriff verursachte Rückkopplungsaktivierung der Repräsentation des Stammes und des Affixes vergrößert wiederum die Zugänglichkeit bei der Bildung eines neuen Lexems und daher die Produktivität des Musters.

Außerdem müssen „mikro-linguistische" Faktoren benannt werden, welche Baayen zufolge zuallererst für die Zusammensetzung der mentalen Repräsentationen, sowie für den graduellen Erwerb der morphologischen Kompetenz zuständig sind (vgl. Hay & Baayen 2005, Baayen 2009, und vor allem Baayen 2010).[23] Die Frequenz in all ihren Formen bildet nach der neueren Baayen´schen Theorie nämlich eher ein „Epiphänomen": Ihre Voraussagekraft bezüglich der Prozesse lexikalischer Verarbeitung ergibt sich aus der Widerspiegelung weiterer distributioneller und kontextueller lexikalischer Eigenschaften, „that are all co-determining learning", wobei „the learning experience is what drives speed of lexical processing" (Baayen 2010: 2–3).

Unter diesen Eigenschaften stellen die morphologische und syntaktische Entropie wichtige Faktoren dar, denn desto unwahrscheinlicher der Gebrauchskontext eines Wortes ist, desto mehr wird es als eigenständige Einheit wahrgenommen. Auf der Ebene der morphologischen Produktivität bedeutet das u.a., dass die Wahrscheinlichkeit der phonotaktischen Fugen eines Morphems, die seine Salienz im sprachlichen Input bestimmt, die Stärke seiner mentalen Repräsentation und seiner Produktivität beeinflusst (Hay & Baayen 2005, Baayen 2009 und 2010).

> „Affixes represented by more words which are infrequent relative to their bases, and which contain low probability phonotactics, are not only the most likely to be more highly segmentable and to develop stronger independent representations, they are also more readily available for use in new word; that is, they tend to be more PRODUCTIVE" (Hay & Baayen 2005: 345).

Als dritter die Produktivität beeinflussender Faktor gilt schließlich die Größe der morphologischen Familien eines Wortbildungsmusters, insofern semantische oder phonologische Überlappungen die Aktivierung seitlich bzw. durch paradigmatische Beziehungen im mentalen Lexikon verbreiten.

[23] Die Rolle dieser „mikro-linguistischen" Faktoren bei den Erwerbs- und Verarbeitungsprozessen führt bei Baayen (2010) u.a. zu der Entwicklung eines neuen Modells für das mentale Lexikon (das *Naive Discriminative Reader Modell*), nach dem die Verbindungen zwischen lediglich zwei Repräsentationsebenen (die eine für Unigramm/Bigramm und die andere für die lexikalische und grammatikalische Bedeutung) auf Basis des Auftrittskontexts und der Häufigkeit sprachlicher Elemente und Strukturen *on going* und inkrementell hergestellt werden.

Die Relevanz dieser theoretischen Überlegungen für meine Hypothese ist eine doppelte. Dass der Produktivitätsgrad einer morphologischen Kategorie eine gegenüber der Struktur und Zusammensetzung des mentalen Lexikons empfindliche Variable bildet, stellt erstens einen weiteren Grund neben den in Kap. 3.2 erwähnten dar, um einen unterschiedlichen Gebrauch der in diesem Fall gerade *produktiven* Wortbildungsstrategien durch die interkulturellen Autoren im Vergleich zu den deutschen zu vermuten. Ich gehe davon aus, dass sie als L2-Lerner eine *nicht-Target-gleiche Beschaffenheit* des mentalen Lexikons aufweisen könnten. Zweitens begründen die Art und Weise des Erwerbs der Wortbildungskompetenz die Möglichkeit einer *Generalisierung* eben der produktiven Wortbildungsmuster. Ein L2-Lerner erwirbt nämlich zuerst die im sprachlichen Input häufigeren und auffälligeren morphologischen Strukturen: Aufgrund des Baayen'schen Modells bilden diese aber auch die produktivsten Wortbildungsmuster, die dann als „Achievement-Strategie" zur Lösung kommunikativer Probleme dienen und in dieser Funktion fossilieren könnte (vgl. Kap. 3.4).

3.3.2 Die drei Maße der Produktivität

Die oben erörterte Korrelation zwischen morphologischer Produktivität und prozessualen *Constraints* beschreibt Baayen (2009: 914–916) als „Conspiracy", wodurch man u.a. die Bedeutung der drei Maße für die Quantifizierung des Produktivitätsgrades eines Wortbildungsmusters deutlicher ausformulieren kann. Insofern sie ein korpuslinguistisches Herangehen an das Problem der Produktivität rechtfertigen und besonders wesentlich für die Auswertung meiner Daten sind (vgl. Kap. 5.2.2 und ff.), erörtere ich im Folgenden die drei Baayen'schen Produktivitätsmaße etwas genauer.

Das Maß der *expanding productivity* bildet das Ergebnis der Division der Anzahl der Hapax Legomena einer morphologischen Kategorie in einem Korpus von N Tokens durch die aller Hapax Legomena desselben Korpus': die Formel lautet daher V(1,C,N)/V(1,N)[24]. Davon ausgehend, dass die Zahl der Hapax Legomena als Schätzung für die der Neubildungen eines Textes gilt, stellt diese Größe den Beitrag eines bestimmten Wortbildungsmusters zur Neubildung im Vergleich zu jenem der anderen morphologischen Kategorien in einem Korpus dar. Da ich mich alleine auf die Analyse von vier Wortbildungsmustern des Deutschen beschränke und nicht die geeigneten

[24] V steht für Typen, C für morphologische Kategorie und N für Tokens, wobei N in diesem Fall die Tokens des Korpus', während N(C) die Tokens einer morphologischen Kategorie darstellt. Die Zahl 1 bezieht sich auf die Vorkommenshäufikeit, sodass V(1) die Hapax Legomena bezeichnet (vgl. Baayen 2009).

Computerprogramme für die automatische Auffindung der Gesamtzahl der Hapax Legomena meines Korpus' zur Verfügung habe, berechne ich diese in meiner Untersuchung nicht.

Gemäß der Formel der *realized productivity* teilt man dagegen die Zahl der Typen einer morphologischen Kategorie V(C,N) durch die ihrer Tokens N(C). Einerseits sind nämlich produktive Wortbildungsmuster auf vielen unterschiedlichen Basen anwendbar, sodass die Verwendungswahrscheinlichkeit verschiedener Typen zunimmt. Anderseits ist normalerweise die Token-Frequenz einer regelmäßig zusammengesetzten Form niedriger als die von unregelmäßigen bzw. durch unproduktive Muster hergestellten Bildungen, deren Existenz im mentalen Lexikon nur von einer hohen Vorkommenshäufigkeit gewährleistet wird. Daraus folgt, dass die totale Zahl der Tokens im Verhältnis zu jener der Typen bei den produktiveren morphologischen Kategorien abnimmt.

Indem die Formel der *realized productivity* die Menge von Neubildungen bzw. Hapax Legomena nicht als selbständigen Indikator der Produktivität betrachtet, quantifiziert diese die Ergiebigkeit einer morphologischen Kategorie, ohne einen Unterschied zwischen ihrer diachronen und synchronen Leistung festzustellen. In meiner Masterarbeit betrachte ich sie als erstes Maß, um einen Vergleich zwischen interkulturellen und deutschen Schriftstellern durchzuführen. Obwohl das Reichtum des Wortschatzes positiv die Zahl der bekannten und verwendeten verschiedenen Typen beeinflusst, ist das Ergebnis einer Gegenüberstellung fragwürdig, insofern die *Interlanguage* der interkulturellen Schriftsteller nicht Target-gleiche Akzeptabilitätsbeschränkungen aufweisen könnte. Insbesondere bei generalisierten, aber nicht stark produktiven Wortbildungsmustern sind darüber hinaus Generalisierungen zu erwarten, weil die Frequenz im sprachlichen Input nicht die tatsächliche Produktivität der morphologischen Kategorie widerspiegelt.

Die *potential productivity*, die man auch als „category-conditioned degree of productivity" definiert, „estimates" schließlich nach Baayen (2009: 902) „the growth rate of the vocabulary of the morphological category itself". Die Formel (V(1,C,N)/N(C)) teilt die Zahl der Hapax Legomena einer morphologischen Kategorie in einem Korpus durch die aller Tokens dieser Kategorie und stellt so die Wahrscheinlichkeit dar, mit welcher ein mit einem bestimmten Wortbildungsmuster zusammengesetztes Wort ein Hapax Legomena (bzw. eine Neubildung) bildet. Daher hängt ihr Wert eng mit der Art der geförderten Verarbeitungsprozesse und der oben erwähnten psycholinguistischen Faktoren zusammen, welche die Stärke der Repräsentationen im mentalen Lexikon bestimmen (vgl. Kap. 3.3.1).

> „The statistic *P* [potential productivity] can be interpreted as the probability that at this point in the production process, lexical procedural knowledge will have to be invoked […] is determined by the semantic transparency of the affix, the frequencies of the already existing items, the

> presence of rival affixes, the frequency with which the procedural knowledge specific to some affix has had to be invoked in the past, etc..“ (Baayen 1994b: 454).

Davon ausgehend könnte sich die *potential productivity* einiger produktiver Wortbildungsmuster in den Texten der interkulturellen Autoren größer als in denen der deutschen Schriftsteller erweisen. Sie würde nämlich einen Gebrauchsunterschied quantitativ erfassen, die von prozessualen *Constraints* bzw. von einer Asymmetrie im pragmatischen Bedürfnis wie auch in der Wirkung von Akzeptabilitätsbeschränkungen bedingt wäre.

3.3.3 Die soziolinguistische Variation der Produktivität

Zum Schluss ihres theoretischen Aufsatzes über die probabilistische Morphologie stellen Hay und Baayen (2005) Fragen für zukünftige Untersuchung, von denen eine wie folgt lautet:

> „If structure emerges gradiently from paradigms of encountered words, then an individual's social networks and personal linguistic experience could considerably affect the degree to which they represent and process specific words as morphologically complex. To what extent is this the case? “ (Hay & Baayen 2005: 346).

Vor dem Hintergrund, dass „given the many linguistic variables that have been shown to be involved in such variation […], morphological productivity is likewise expected to be a determinant and exponent of variation“, beklagen Baayen und seine Kollegen noch 2006 den Mangel an Studien, die Variationen in der Kodierung, im Verarbeitungsprozess und im Gebrauch der morphologischen Strukturen in Beziehung zu soziolinguistischen Variablen betrachten (vgl. Keune et al. 2006: 572). Es gibt aber Ausnahmen, die methodologische Hinweise für die Durchführung meiner Untersuchung bieten, obwohl die Beschaffenheit ihrer Daten und ihre Fragestellung nur teilweise mit meinen übereinstimmen.

Indem Baayen (1994a) die zentrale Annahme vieler qualitativer Studien ablehnt, welche die morphologische Produktivität als ein der Anzahl von strukturellen Beschränkungen indirekt proportionales Maß darstellen (vgl. für das Deutsche u.a. Fleischer et al. 2007), nennt er zum Beispiel anhand eines in stilometrischer Analyse entwickelten Verfahrens eine Methode, um die Korrelation des Produktivitätsgrades mit stilistischen und soziolinguistischen Variablen zu erfassen (vgl. auch Kap. 4). Durch eine *principal component analysis* der Verteilung von 27 Affixen des Englischen in einem literarischen Korpus beweist er die Mitwirkung einerseits des Texttyps,

andererseits der individuellen Abneigungen sowie der diachronischen Faktoren bei der Verwendung von Wortbildungsmustern. Ferner stellen Keune et al. (2006) Variationen des Produktivitätsgrades einiger Affixe in gesprochenem Niederländisch bezüglich des geografischen Ortes, des Alters und der Schulbildung fest, während Säily (2011) sich auf die Abhängigkeit der Produktivität von der Variable Geschlecht anhand einer Untersuchung des *Britisch National Corpus* konzentriert.

Die letzte Studie ist für meine Analyse nicht nur methodologisch, sondern auch theoretisch nützlich. Einerseits arbeitet Säily nämlich mit der Baayen'schen Formel der *realized* und *potential productivity*. Jedoch betrachtet sie die beiden Maße aufgrund ihrer Empfindlichkeit auf die Subkorpusgröße als wenig zuverlässig. Deshalb kontrolliere ich meinen Korpus sowie die Subkorpora für die Tokenanzahl und ergänze den Baayen'schen Ansatz mit dem „variable-corpus approach" von Gaeta und Ricca (2006), der auf die Überwindung der damit verbundenen Auswertungsprobleme gerichtet ist (vgl. Kap. 5.2.2). Andererseits hebt Säily die probabilistische mehr als die kategoriale Natur der Korrelation zwischen dem Produktivitätsgrad und dem betrachteten soziolinguistischen Faktor hervor. Wenn die Männer im Vergleich zu den Frauen ein bestimmtes Wortbildungsmuster häufiger verwenden, schließt dies nicht die Möglichkeit aus, dass der Sprachgebrauch einiger Individuen gegen die festgestellte Norm verstößt. Trotzdem ist das Kriterium der statistischen Signifikanz gültig, um gewisse Tendenzen zu erfassen.

Von diesen Bemerkungen ausgehend untersuche ich anhand meiner Daten, ob Variationen in dem Produktivitätsgrad einiger Wortbildungsmuster auch in *geschriebenen literarischen deutschen* Texten je nach dem Auswanderungsgrund der Autoren zu finden sind.

3.4 Die Produktivität und die Lernersprache: Erwerb der Wortbildungskompetenz

Der Erwerb der „lexical knowledge" hängt einerseits von der Fähigkeit ab, die mentalen Repräsentationen bestimmter Lexeme für die Zwecke der Sprachproduktion und -rezeption zu aktivieren und sie syntaktisch, semantisch und pragmatisch zutreffend zu verwenden, andererseits von der Bereitschaft, bekannte Elemente für die Bildung neuer Ausdrücke zu kombinieren (González Álvarez 2004: 11–12). In diesem Zusammenhang ist das Erreichen der Wortbildungskompetenz eines L2-Lerners ein wesentlicher Prozess, der graduell geschieht und Ähnlichkeit mit dem bei L1-Lernern aufweist. Trotz der möglichen Wechselwirkung mit Strukturen der Muttersprache steuern das Erlernen der Wortbildungsstrategien bzw. das innovative Verhalten der L2-Sprecher nämlich dieselben Prinzipien und Anlässe, die auch bei Kindern ursächlich sind:

„Despite the difference between first and second language acquisition, the assumption can be made that L2 learners' innovative behaviour in this area is driven by basically the same reasons that lie behind children's coinages. The need to express at a particular moment a concept for which they have not yet learnt the exact form makes children as well as L2-learners coin new terms, whose formation is based on observed adult or native speaker linguistic behaviour" (González Àlvarez 2004: 35).

Die hier an zweiter Stelle genannte Ähnlichkeit zwischen dem Verhalten von L1- und L2-Lernern hebt die zentrale Rolle der Beobachtung des linguistischen Verhaltens von Erwachsenen bzw. Muttersprachlern beim Erwerbsprozess hervor. Was die produktiven morphologischen Kategorien angeht, weist diese Annahme einerseits auf die Rolle der Frequenz eines Wortbildungsmusters für die Reihenfolge seines Erwerbs hin. Anderseits erklärt sie Abweichungen und Fehler in beiden Lernsituationen, die nämlich häufig aus der Nicht-Übereinstimmung zwischen generalisierten und tatsächlich produktiven Wortbildungsprozessen entstehen (González Àlvarez 2004: 42–43).
Diesbezüglich spricht auch Coseriu (1975) über unterschiedliche morphologische Kompetenzen, welche das System oder die Norm einer Sprache betreffen.[25] Die hier angenommene Hypothese, dass deren Erwerb getrennt verläuft und daher Fehler entstehen können, wird indirekt von Coseriu bekräftigt:

„[D]ie Aufschlüsse über den Unterschied zwischen Norm und System in der Morphologie […] liefern uns wohl am ehesten die Flexionsfehler der Kinder oder ganz allgemein die nicht hinreichend mit der Norm vertraute Person. Denn Irrtümer entstehen hier fast immer aus der einfachen Anwendung der funktionellen Oppositionen des Systems, die der in der betreffenden Sprachgemeinschaft als normal sanktionierten Anwendung zuwiderläuft" (Coseriu 1975: 68).

Aufgrund ihrer geringen Vertrautheit mit der sprachlichen Norm sind oft die konventionellen Oppositionen für L2-Lerner weniger verbindlich als für Muttersprachler. Indem man diese Perspektive mit jener der vorangehenden Kapiteln ergänzt, könnten produktive Wortbildungsmuster, welche die L2-Sprecher anhand ihrer Frequenz und vielleicht über das Studium im Sprachunterricht zuerst erwerben, in Ermangelung anderer Beschränkungen generalisiert werden und fossilierte Merkmale der *Interlanguage* darstellen.

[25] Nach Coseriu (1975: 11–101) bildet das *System* einer Sprache die Menge der funktionellen Oppositionen bzw. distinktiven Merkmale, die für ihr Funktionieren unerheblich sind. Die *Norm* stellt dagegen die Gesamtheit der gewöhnlichen und konventionellen Realisierungen eines Sprachsystems dar, die sozial und kulturell zu einem bestimmten Zeitpunkt bzw. auf eine gegebene Sprachgemeinschaft festgelegt sind.

Die zweite Ähnlichkeit zwischen L1- und L2-Lernern bezieht die Anlässe zur Wortprägung mitein, die bei den Lernern im Vergleich zu den Erwachsene, bzw. Muttersprachlern zunehmen (vgl. Kap. 3.2). Nach González Àlvarez (2004: 13–15 und 20–28) führt diesbezüglich der von dem unvollständigen Lernprozess bedingte lexikalische Mangel zur häufigeren Anwendung allgemeiner „kommunikativer Strategien", unter denen sich auch die Wortbildungsstrategie befindet.

Unter dem Oberbegriff der „kommunikativen Strategien" versteht man linguistische oder außerlinguistische Mechanismen, die ein Sprecher für die Vermittlung einer *Message* verwendet, falls die dafür benötigte grammatikalische Struktur ihm selbst oder dem Gesprächspartner nicht bekannt ist. In der Literatur unterscheidet man zwischen *reduction*-, *interactive*- und *achievement*-Strategien, wobei erste und zweite Art jeweils zur Vermeidung unbekannter Elemente oder Strukturen bzw. zur Verwendung von Gesten oder Mimik führen, während die dritte Sorte von Strategien das kommunikative Problem durch einen kreativen Umgang mit schon vorhandenen sprachlichen Elementen löst (vgl. u.a. González Àlvarez 2004: 26–27). Die Anwendung von produktiven Wortbildungsmustern bildet einen Fall der letzten Kategorie: sie reagiert zusammen mit anderen *achievement*-Strategien (Transfer, Periphrase, Bedeutungserweiterung) auf lexikalische Mängel in der *Interlanguage* eines Lerners.

Indem sie sich auf weitere Studien bezieht, erwähnt González Àlvarez 2004 darüber hinaus, dass Lerner auf unterschiedlichen sprachlichen Niveaus verschiedene *achievement*-Strategien bevorzugen. Während Anfänger kommunikative Probleme eher durch Bedeutungserweiterungen und Transfers lösen, stellt man bei fortgeschrittenen L2-Lernern oft spontane Neubildungen (*word-coinages*) anhand von Strukturen der TL fest (2004: 42). Obwohl das pragmatische Bedürfnis bei Anfängern wahrscheinlich am stärksten wirkt, kann man daher einen Übergebrauch von produktiven Wortbildungsstrategien auch bei fortgeschrittenen Lernern nicht ausschließen.

4. Die dritte Unterhypothese: Gebrauchsunterschied und dessen Untersuchungsmethode

4.1 Die Variable und die Korpuslinguistik

Anhand dieser Prämissen ist es möglich, dass die ausländischen Schriftsteller unabhängig von ihrer Muttersprache sich in der Anwendung produktiver Wortbildungsmuster gegenüber dem Deutschen signifikant unterscheiden. Indem ich mich im Rahmen dieser Arbeit fast ausschließlich auf die Analyse einer morphologischen Kategorie beschränke, führe ich für die Auswertung dieser Hypothese eine quantitative Untersuchung des Produktivitätsgrads der Wortbildungsmuster auf *-ung* in einem Korpus von Texten interkultureller Autoren durch und vergleiche ihn mit demselben in einem Korpus deutscher Schriftsteller.[26]

Die Anwendung eines quantitativen Verfahrens für die Bearbeitung meiner Fragestellung hängt sowohl mit der Beschaffenheit meiner Daten (vgl. Kap. 4.2), als auch mit der Variable zusammen. Zunächst bezüglich Letzterer ist ein quantitativer Ansatz notwendig, insofern die Produktivität eines Wortbildungsmusters kein binäres Maß bildet, sondern nur graduell erfassbar ist (vgl. Kap. 3.3.2). Deshalb beschreibt man auch ihre Variationen nicht (allein) durch kategoriale Behauptungen der Art: Ein Affix ist oder ist nicht produktiv in einem Text bzw. eine Wortbildungsregel ist oder ist nicht einem Lernern bekannt, sondern eher durch eine Aussage hinsichtlich der Gebrauchswahrscheinlichkeit, wie etwa: Ein Affix ist mehr oder weniger produktiv in einem Text im Vergleich zu einem anderen (Texten oder Affixen) bzw. ein morphologisches Muster wird mehr oder weniger häufig von einem Sprecher verwendet. Der Vergleich der Produktivitätsgrade einer morphologischen Kategorie in zwei Korpora beruht nämlich auf der Gegenüberstellung der Anzahl der damit zusammengesetzten *Tokens*, unterschiedlichen *Typen* und Neubildungen (bzw. *Hapax Legomena*) sowie auf dem Verhältnis ihrer Frequenz: Die Untersuchung der Vorkommenshäufigkeit verlangt eine quantitative Analyse der betrachteten Daten.

Diesbezüglich stellt die Korpuslinguistik eine Methodologie für die Datenerhebung und die Datenauswertung dar, die sich als besonders geeignet für die Durchführung solcher quantitativer Untersuchungen auch in der SLA-Forschung erweist (vgl. Granger 2007, Granger 2009, Lüdeling & Walter 2010):

[26] Außer einer quantitativen Analyse des Produktivitätsgrads der Wortbildungsmuster auf *-ung* je nach der Variable *Nativeness*, *Auswanderungsgrund*, *L1* und *Autor* führe ich auch die der morphologischen Kategorie *-keit/-igkeit/-heit* bzw. *-bar* und *ver-* durch, jedoch nur bezüglich der Variable *Nativeness*. Ich spreche in all diesen Fällen jeweils von Wortbildungsmustern im Plural in Anlehnung an Motsch 2004, der nämlich unterschiedliche semantische Muster für die Derivate auf *-ung*, auf *-bar* und so weiter feststellt (vgl. Kap 4.3 und Kap. 5.2.4).

„Although corpora are but one source of evidence among many […] there is general agreement today that they are ‚the only reliable source of evidence for such features as frequency' […] an aspect of language of which we have very little intuitive awareness but one that plays a major part in many linguistic applications which requires a knowledge not only of what is possible in a language but what is likely to occur" (Granger 2007: 4).

Aus einem Korpus, der eine „für einen bestimmten Zweck zusammengestellte (elektronische) Textsammlung[...]" ist (Lüdeling & Walter 2010: 315), kann man nämlich durch Korpusabfrage „authentische" Belege einer bestimmten linguistischen Struktur generieren (316) und durch Korpuswerkzeuge „Wortlisten mit Korpusfrequenzen erstell[en]" (317), die systematische und reproduzierbare Analysen bestimmter Aspekte des Sprachgebrauchs ermöglichen.
Die „Authentizität" bildet dabei eine der wichtigsten Eigenschaften der Daten eines Korpus'. Obwohl man sie beim Korpusaufbau nicht immer in einem absoluten Sinn versteht und z.B. verschiedene Natürlichkeitsgrade bei der Textproduktion feststellt, sind nämlich die Daten einer korpuslinguistischen Untersuchung als „continuous text" zu verstehen (im Gegenteil z.B. zu den mittels Experimenten bzw. Introspektion und teilweise auch durch Elizitierung (*elicitation*) gesammelten Daten), wobei der sprachliche Kontext einer bestimmten Variablen immer zur Verfügung steht und die Sprecher nicht zur Produktion einer besonderen Struktur animiert werden, sondern es ihnen freisteht, „to choose their own wording" (vgl. Granger 2009: 261).
Ein weiteres grundsätzliches Kriterium der Datensammlung in der Korpuslinguistik besteht in der „Repräsentativität". Repräsentativ sind Korpora, die man nach einem vorbestimmten *Design* herstellt, um adäquat die linguistische Variation innerhalb einer Population gemäß vorgegebener Parameter darzustellen (vgl. auch Kap. 5.1). Auf den Kriterien der Authentizität und der Repräsentativität baut man die Auffassung eines Korpus' als „model for linguistic experience" auf, die nach Lüdeling und Evert (2003: 1) dem Baayen'schen Herangehen an das Problem der Produktivität zu Grunde liegt und auch hier vorausgesetzt wird:

„The approach adopted by Baayen need not to be seen as a purely statistical approach, though. It can also be understood as a cognitive model which assumes that a speaker somehow ‚knows' that she can use process X to produce new types precisely *because* she has seen many new types produced by X (i.e., she knows that other speakers have also produced new words by X)" (Lüdeling & Evert 2003: 1. Von den Autoren hervorgehoben).

Die Verteilung und die Frequenz eines bestimmten Wortbildungsmusters in einem Korpus spiegelt von diesem Standpunkt aus die von einem Sprecher aufgrund des Inputs gesammelte linguistische

Erfahrung sowie seine Kompetenz bezüglich der Anwendungsmöglichkeiten einer morphologischen Kategorie wider. Davon ausgehend gilt der Wert des Produktivitätsgrads der Wortbildungsmuster auf *-ung* im Vergleichskorpus als *Baseline*, auf der die Abweichungen in der morphologischen Kompetenz der interkulturellen Autoren zu schätzen ist (vgl. Kap. 4.2 und Kap. 5.1.4).

4.2 Die Lernerkorpora

Die Beschaffenheit meiner Daten bildet den zweiten Grund für die Durchführung einer quantitativen Analyse. Da die interkulturellen Schriftsteller nicht von einem bestimmten *Task* gezwungen schreiben bzw. ihre Texten vor der Veröffentlichung lektorieren lassen, sind mögliche sprachliche Unterschiede mit den deutschen Autoren, wie schon erwähnt, eher in der Verwendungshäufigkeit einiger Strukturen des Deutschen als in tatsächlichen Fehlern zu suchen.
Die hier vorgenommene Untersuchung folgt dementsprechend dem Modell der sogenannten *Contrastive Interlanguage Analysis* (CIA), die zusammen mit der *Computer-aided Error Analysis* (CEA) eine der verbreitetsten Methoden für die korpuslinguistische Auswertung von Lernerdata darstellt (vgl. Granger 2007, 2009). Beide Verfahren unterscheiden sich jeweils von denen der Kontrastiven Linguistik und der Fehleranalyse (vgl. Kap. 2.1), gerade für die Wahl eines elektronischen Korpus' als Datenquelle, welcher die Miteinbeziehung des sprachlichen Kontextes und eine systematischere Suchabfrage ermöglicht. Da man in der CIA Äußerungen von L2-Lernern mit denen von *nativen* Sprechern oder anderen Lernern vergleicht, erweist sich diese Untersuchungsmethode und ihre Leistungen für meine Fragestellung als besonders passend:

> „NS/NNS [native speech/non native speech] comparisons can highlight a range of features of non-nativeness in learner writing and speech, i.e. not only errors, but also instances of under- and overrepresentation of words, phrases and structures" (Granger 2007: 12).[27]

Die Durchführung einer *contrastive interlanguage analysis* benötigt sodann den Aufbau zweier Korpora. Das eine enthält Texte der zu analysierenden Population von L2-Lernern, während das andere aus kontrollierten Texten von Muttersprachlern (bzw. anderen Lernern) besteht, welche die Rolle eines Prüfsteins einnehmen. Falls es repräsentativ ist, stellt nämlich das sogenannte

[27] Diesbezüglich sind gerade Phänomene von Über- bzw. Untergebrauch von Strukturen der Targetsprache nach Granger (2009: 267) für „the foreign-soundinness of perhaps otherwise error-free advanced Interlanguage" verantwortlich, während Walter und Grommes (2008) von einem Zustand der „Near-nativeness" bezüglich fortgeschrittener Lernervarietäten sprechen, die „trotz keiner erkennbaren Fehler von Sprecher der Zielsprache als abweichend angesehen" werden (13) und daher am besten durch quantitativen Methoden zu untersuchen sind (17).

„Vergleichskorpus“ die *native* Norm dar, aufgrund der man den Abweichungsgrad der durch das Target-Korpus festgestellten lernersprachlichen Norm bewertet. Davon ausgehend betrachte ich in meiner Arbeit den Produktivitätsgrad der Wortbildungsmuster auf *-ung* in den Texten der deutschen Schriftsteller als den „normalen“ bzw. *standard* Produktivitätswert. Durch einen statistischen Test überprüfe ich im Nachhinein, ob der anhand des Korpus' der interkulturellen Autoren berechnete Produktivitätsgrad des Affixes *-ung* von dem des Vergleichskorpus' signifikant abweicht oder die erfassten Variationen sich als unerheblich erweisen, indem der eine Wert zu derselben (idealen) Population des anderen gehört (vgl. Kap. 5.2.2).

Im zweiten Teil meiner kontrastiven lernersprachlichen Analyse stelle ich dagegen *non-native* Varietäten gegenüber. Dafür lasse ich das Vergleichskorpus außer Acht und überprüfe die Anwesenheit von Unterschieden im Produktivitätsgrad der Wortbildungsmuster auf *-ung* nach drei anderen unabhängigen Variablen, welche die Target-Population jeweils in drei, zehn bzw. fünfzehn Subkorpora aufteilen (vgl. Kap. 4.4). Obwohl das Vorgehen dem der ersten Fragestellung durchaus ähnlich ist, ist es von einem theoretischen Standpunkt aus wichtig anzumerken, dass hier die Bestimmung einer Prestigehierarchie zwischen den Gruppen bzw. ein *Standard* weniger unmittelbar ist. Daher stellt man die Wahrscheinlichkeit fest, mit der die unterschiedlichen Werte in dem Produktivitätsgrad des Affixes *-ung* aus einer selben Population stammen, ohne den normalen Wert jedoch zu erkennen.

Ausgangspunkt des hier vorgestellten Herangehens bildet ein Aufsatz von Labov (2004), der die Bedingungen und Ziele der quantitativen Beweisführung in der Linguistik ausführlich erläutert. Indem man z.B. Abhängigkeitsbeziehungen unter zwei Variablen bzw. Interaktions- oder Wettbewerbsrelationen zwischen mehreren festsetzt, ist nach Labov „the major thrust of quantitative reasoning […] to characterize surface variation in terms of the largest (and most abstract) set of objects that vary in the same way“ (2004: 10).

Um die Dimensionen der sprachlichen Variation zu untersuchen, ist der erste Schritt die binäre Aufteilung der Population in zwei wesentliche Gruppen gemäß der Feststellung einer unabhängigen Variablen, die in meinen Fall die der *Nativeness* darstellt. Da aber die eventuelle Beobachtung eines oder keines sprachlichen Unterschieds auch das Produkt einer durch die gewählte binäre Aufteilung uneinheitlich verteilten dritten Variable sein kann, kann das Ergebnis überprüft werden, indem man andere hinsichtlich des Merkmals und der Population möglicherweise relevante Faktoren in Betracht zieht. Diese bestehen in meinem Fall im *Auswanderungsgrund*, in der *Muttersprache* und im *Autor*.

Die Lernersprache „is influenced by a wide variety of linguistic, situational and psycholinguistic factors“. Darum fordert der Aufbau eines Lernerkorpus' z.B. die Kontrolle über weitere „L2-

specific variables“ außer jenen der üblichen Korpora (Granger 2009: 263). Deswegen sind die von mir betrachteten Variablen auf keinen Fall die einzigen möglichen Faktoren, auf die man eventuelle Variationen in der Sprache der interkulturellen Schriftsteller zurückführen kann (vgl. Kap. 5.1.3).

4.3 Korrelativer und funktionaler Ansatz: Begriff von Stil

Insofern ich die Abhängigkeit eines quantifizierbaren linguistischen Merkmals (des Produktivitätsgrads der Wortbildungsmuster auf *-ung*) von außerlinguistischen vorbestimmten Variablen untersuche (der *Nativeness*, dem *Auswanderungsgrund*, der *L1* und dem *Autor*), ist mein Ansatz hinsichtlich einer eventuellen sprachlichen Variation in meinem Korpus in seinen allgemeinen Zügen korrelativ. Diesbezüglich steht in der Literatur ein „korrelativer“ Standpunkt einem „funktionalen“ oder „interpretativen“ gegenüber, wobei die soziolinguistischen Arbeiten von Labov bzw. die der Ethnographie der Kommunikation die beiden Standpunkte jeweils emblematisch darstellen (vgl. Berruto 1995: 29–32).

Eine der bedeutendsten Differenzen ergibt sich gerade in der Betrachtungsweise der außerlinguistischen Variablen. In einem korrelativen Ansatz bilden diese, die man übrigens als unabhängig definiert, nämlich keinen selbständigen Untersuchungsgegenstand, sondern festere „makro-strukturelle Konstrukte“, welche die „Referenzidentität“ eines Sprechers und daher Variationen in seinem sprachlichen Verhalten einseitig verursachen. Ein interpretativer Ansatz dagegen betrachtet außerlinguistische und linguistische Variablen als sich gegenseitig bestimmend, indem man das Verhältnis zwischen Sprache und Kontext als wechselseitig, sowie die Sprecheridentität als eine veränderliche und pluralistische, sich in der konkreten kommunikativen Handlung organisierende Struktur versteht (vgl. u.a. Berruto 1995: 29–32, Androutsopoulos & Spreckels 2010 und Kap. 2.2).

Ein weiterer Unterschied zwischen beiden Ansätzen tritt darüber hinaus in der Auslegung des „Stil“-Begriffes und der Bewertung seines Verhältnisses zu dem der „Varietät“ auf. Varietät und Stil definiert Dittmar (2008b: 1670) als zwei „Hyponyme des Hyperonyms ‚sprachliche Variation‘“, die man für die Beschreibung des „zwischen ‚langue‘ und ‚parole‘ anzusetzende[n] weite[n] Feld[es] des (mündlichen) Sprachgebrauchs“ verwendet. Die varietätenlinguistische Forschung ist im Gegensatz zu der auf den Stil fokussierten hauptsächlich korrelativ, insofern man eine Menge von ko-okkurierenden sprachlichen Eigenschaften auf der grammatikalischen Ebene in Korrelation mit allgemeingültigen und außersprachlichen Bezugsdimensionen setzt, um sie als ein vollständiges Subsystem zu beschreiben. Innerhalb ihres Rahmen entwickeln sich zwei Hauptauffassungen von

„Stil“ als diaphasische Variable, die ihr doch eine zweitrangige Rolle im Vergleich zu den diatopischen und diastratischen Faktoren in der Bestimmung der sprachlichen Variationsdimensionen zuschreiben (vgl. Berruto 1995: 147–152, Androutsopoulos & Spreckels 2010 und Fußnote 14).

Indem einerseits Labov den Stil als den Aufmerksamkeitsgrad auf die Sprache betrachtet, andererseits u.a. Preston (1993) ihn als die vom „audience-design“ bedingte linguistische Variation erfasst, erweist sich die Schichtung der durch diese Definitionen festgestellten stilistischen Eigenschaften innerhalb einer Sprachgemeinschaft als von den sozialen Klassen unabhängig bzw. von den diastratischen Variablen abgeleitet, da „each social class follows a parallel path of style schifting“ (vgl. Labov 2004: 9, der sich auch auf die Studien von Preston bezieht).

Solche Schlussfolgerungen scheinen mit dem Stilkonzept eines interpretativen Ansatzes unverträglich, der nämlich den Stil als von dem Adressat *interpretierbare* Menge von sprachlichen und nicht-sprachlichen Merkmalen bzw. als Kennzeichnen der sozialen Identität und Gruppenzugehörigkeit eines Sprechers darstellt:

> „Stile sprachlichen Verhaltens sind ein wesentliches soziales Unterscheidungsmerkmal, und ihre Ausprägung ist mit der Ausbildung von sozialen Welten und der sozialen Identität von Gruppen und größeren Gemeinschaften verbunden“ (Kallmeyer 1994: 30–31).[28]

Zentral ist für diese Auffassung die Ansicht, dass stilistische Unterschiede nicht als Ergebnis einer Wahl zwischen formal verschiedenen, aber semantisch identischen Realisierungsmöglichkeiten verstanden werden, sondern als Funktion-gerichtete und Bedeutung-tragende Merkmale zu betrachten sind (vgl. Kallmeyer 1995: 7). Obwohl ich in meiner Arbeit eine korrelativ-variationische Methode bzw. quantitative-statistische Kriterien für die Bestimmung der Variation und der Signifikanz meines linguistischen Merkmales anwende, betrachte ich in Anlehnung an einen interpretativen Ansatz die Beziehung zwischen linguistischen und außerlinguistischen

[28] Davon ausgehend, dass der Stil nicht nur als Kennzeichnung eines Textes oder eines Individuum, sondern auch einer Gruppe bzw. eines Aktivitätstyps und einer Institution fungiert, entwickelte sich der Begriff des „sozialen kommunikativen Stils“ (vgl. u.a. Kallmeyer 1994, 1995, Dittmar 2008b, 2008b). Er wird als Menge von linguistischen und außer-linguistischen Merkmalen definiert, die für den Ausdruck und die Zuschreibung einer sozialen Identität relevant sind, wobei man unter „sozialer Identität“ die Selbst- und Fremdzuordnung eines Individuums zu allgemeinen sozialen Kategorie versteht. Im Rahmen eines interpretativen Ansatzes besteht jedoch diese Zugehörigkeitszuschreibung in einem sich in der kommunikativen Handlung bestimmenden Prozess, in dem u.a. das Wissen der Teilnehmer über die Unterschiede zwischen *standard* und *non-standard* Varietäten bewusste Symbolisierungsvorgänge ermöglicht. Ferner besitzt ein Individuum von diesem Standpunkt aus die Kompetenz über unterschiedliche soziale kommunikative Stile, indem es zu verschiedenen Netzwerken, Aggregaten und Gruppen für die Bearbeitung sozialer Bedürfnisse (sozialer Welten) gehört. Die Feststellung der Stilunterschiede benötigt schließlich einem interpretativen Ansatz zufolge eine quantitative sowie qualitative Analyse, insofern einige Merkmale Schlüsselphänomene in der Stilcharakterisierung unabhängig von ihrer Vorkommenshäufigkeit darstellen.

Variablen nicht immer als unwillkürlich und automatisch geschehend, sondern teilweise als bewusst und durch Stereotype bzw. Stilisierungsprozesse vermittelt.
Insofern nach Baayen (1994a) auch *stilistische* Faktoren neben den soziolinguistischen die Ausbeutung der Wortbildungsstrategien einer Sprache beeinflussen bzw. meine Analyse sich mit literarischen Texten beschäftigt, die traditionell den Gegenstand *stilistischer* Untersuchungen bilden, erläutere ich in den nachfolgenden zwei Kapiteln etwas genauer die von der Stilometrie und der Stilistik vorausgesetzten Auffassungen vom Stil mit besonderer Aufmerksamkeit auf die gegenseitige Wirkungsbeziehung zwischen Stil und morphologischen Eigenschaften eines Textes (vgl. Kap. 4.3.1 bzw. Kap. 4.3.2).

4.3.1 Der Stil in der Stilometrie

Obwohl es sich auf den Begriff des sozialen Stils im Besonderen bezieht, gilt das nachfolgende Zitat von Kallmeyer (1995) als Ausgangspunkt, um die zweifache und teilweise widersprüchliche Natur des Stils als unwillkürliches Verhalten eines Sprechers (Habitus) bzw. als Bedeutung tragende und Gruppeszugehörigkeit ausdrückende Wahl zwischen verschiedenen Verhaltensmöglichkeiten zu erörtern:

> „Sozialer Stil ist einerseits geprägt durch die tiefsitzenden, fest einsozialisierten Prägungen, die für die meisten Gesellschaftsmitglieder fraglos und vielfach auch unbemerkt den Status von selbstverständlichen und stillschweigenden Voraussetzungen haben. In dieser Hinsicht kommt Stil geradezu einer natürlichen Einstellung gleich. Zugleich wird sozialer Stil aber auch entwickelt, konturiert und verändert in Auseinandersetzung mit der umgebenden Gesellschaft und dem sprachlichen Verhalten der anderen. Dieser Prozeß der Stilentwicklung und -abgrenzung ist reflexiv und mit (naiver) Theoriebildung verbunden“ (Kallmeyer 1995: 8).

Vom ersten Aspekt des Stilbegriffes gehen stilometrische Analysen aus, die hier relevant sind, um den möglichen Einfluss des Stils auf die linguistischen und insbesondere morphologischen Eigenschaften der Sprache der interkulturellen Schriftsteller zu betrachten.
Die Stilometrie definiert Stil als die für eine Person, aber auch für eine Periode oder eine Gattung typische und distinktive Verteilung sprachlicher Merkmale (vgl. u.a. Holmes 1994, Stamatatos 2009). Mit dem Hauptzweck der automatischen *Authorship-attribution* analysiert man auf dieser Grundlage quantitativ linguistische Eigenschaften eines Textes, welche in der Regel einer bewussten Manipulation des Autors unzugänglich sind. Obwohl sie sich als wenig zuverlässig

erweisen (im Vergleich zu anderen Ähnlichkeitsgrößen wie z.B. der auf der Verteilung von Funktionswörtern bzw. n-Grammen beruhenden), befinden sich unter den relevanten Stilsmaßen auch solche, die Reichtum und Beschaffenheit des Wortschatzes eines Sprechers erfassen und die man jeweils durch die *typen-tokens-ratio* bzw. die Frequenz der Hapax Legomena und Dislegomena (die Wörter, die in einem Text nur zwei Mal vorkommen) quantifiziert. Die Ähnlichkeiten dieser zwei Maßgaben mit den Maßen der *realized* und *potential productivity* einer morphologischen Kategorie (vgl. Kap. 3.3.2) werden noch offensichtlicher, wenn man Baayens erfolgsreichen Versuch betrachtet, sie als Grundlage für Textszuschreibungsaufgabe nach Autor und Texttyp zu verwenden (1994a).

Interessant für meine Fragestellung ist darüber hinaus die Ergänzung der stilometrischen Perspektive durch den Begriff des „sozialen kommunikativen Stils“ (vgl. Fußnote 28) und mit der in den Kapiteln 2.2 und 2.3 erwähnten Möglichkeit einer soziolinguistischen Fossilisation auf Gruppenebene. Insofern sind einerseits „Stile ein Medium, mit dem Akteure im sozialen Alltag ihre Identität in semiotischen Gestaltformaten Ausdruck verleihen können“ (Dittmar 2008a: 1267). Anderseits fungieren fossilierte Eigenschaften einer *Interlanguage* manchmal als „Marker“ der Identität als Nicht-Muttersprachler eines oder einer Gruppe von L2-Lernern, weshalb man fragen kann, ob ein eventueller Übergebrauch der Wortbildungsmuster auf *-ung* bei den interkulturellen Schriftstellern auch diesen Funktionen nachkommt.

4.3.2 Der Stil in der Stilistik

Der zweite Aspekt des in dem Zitat von Kallmeyer erwähnten Stilkonzepts bezieht sich auf die Entwicklung des „natürlichen“ Stils durch die Umgebung und auf seine mögliche Veränderung infolge des Willens eines Sprechers, um z.B. Abgrenzungs- bzw. Zugehörigkeit auszudrücken. Mit diesen Prozessen verbunden ist die Frage nach der Interpretation bestimmter stilistischer Eigenschaften bzw. deren bewusster Verwendung als bedeutungstragende Einheiten.

Dass auch die morphologischen Eigenschaften bzw. die Neubildungen eines Textes stilistische Effekte aufweisen können, nimmt man in der Literatur an. Laut Harnisch (2004: 522) lassen sich z.B. Variationen im morphologischen Bereich, insofern Morpheme eine relativ geschlossene Klasse bilden, prinzipiell gut vom Hörer/Leser erkennen und haben daher oft stilistische Auswirkungen. Nach Handler besteht diesbezüglich „eine stilistische Entscheidung [...] schon darin, ein WBP [Wortbildungsprodukt] einzusetzen, wo auch eine andere Formulierungsmöglichkeit bestünde“ (2008: 1567). Deren Bewertung findet dann oftmals auf Basis einer Differenz- oder einer

Defizithypothese statt, indem man morphologische Abweichungen jeweils entweder als eine andere, aber zur eigenen ko-existierende Ausdrucksart, oder als Fehler betrachtet.

Da viele Wortbildungen unauffällig momentane Bezeichnungsbedürfnisse erfüllen, einige der Textgestaltung und -kohärenz genügen, andere für die Kennzeichnung der Figuren oder der Gattung eines Textes verwendet werden, weisen nicht alle dieselbe stilistische Funktion auf (vgl. Motsch 2004 und Handler 2008). Davon ausgehend unterscheidet Dressler (1981: 427–428) in der Bildung neuer Wörter durch derivative morphologische Strategien zwischen verschiedenen Graden von „Audacity", wobei die Neubildungen durch produktive Affixe als am wenigsten auffällig sind bzw. „kühn" wirken im Vergleich zu den Neubildungen durch weniger produktive Wortbildungsmuster oder infolge der Verletzung von Basisbeschränkungen. Erstere verstoßen nämlich nur gegen die lexikalische *Norm* einer Sprache, während Letztere aus Abweichungen von den lexikalischen sowie morphologischen linguistischen Konventionen bzw. aus der Übertretung der funktionellen Oppositionen des *Sprachsystems* folgen (vgl. die in Kap. 3.4 erwähnte Unterscheidung Coserius zwischen Norm und System einer Sprache).

Aufgrund der schon erwähnten Schwierigkeiten bei der Erhebung der Neubildungen durch quantitative Methoden und in der Auseinanderhaltung der zwei Unterkategorien der Neologismen und Textwörter gehe ich der Frage nach den stilistischen Konsequenzen eines eventuellen Übergebrauchs der Wortbildungsmuster seitens der interkulturellen Autoren nicht nach. Voraussetzung dafür bietet mir das Forschungsverfahren der Korpusstilistik, welche die quantitative Methode der Korpusslinguistik auf die Fragestellungen der Stilistik anwendet (Semino & Short 2004).

Die traditionelle Stilistik entwickelt sich als Abweichungsstilistik und geht davon aus, dass ein Autor durch die Verwendung besonderer linguistischer Strukturen vorsätzlich eine weitere Bedeutung als die buchstäbliche vermitteln kann, wobei man ein wichtiges Kriterium für die Identifikation solcher Strukturen in ihrer Auffälligkeit und in ihrer Abweichung von der sprachlichen Norm erkennt. Den absoluten Auffälligkeitsbegriff kritisieren aber spätere Ansätze der Stilistik, die z.B. die Rolle des Lesers sowie den außerlinguistischen Kontext hervorheben. Obwohl sich die Korpusstilistik nur auf die linguistischen Eigenschaften eines Textes konzentriert, nimmt sie auch einen alternativen Begriff der Auffälligkeit (Salienz) an, der nämlich an der Frequenz und der Instabilität einer Variable als von ihrer absoluten Distinktivität orientiert ist. Die Erkennung der relevanten Merkmale ist hierdurch quantitativ festzustellen, während die stilistischen Konsequenzen dann qualitativ auszuwerten sind. Auf dieser Grundlage setze ich auch in dieser Arbeit eine Korrelation zwischen Variabilität und Salienz eines Merkmals voraus, ohne aber eine direkte

Beziehung zwischen den identifizierten sprachlichen Besonderheiten und der stilistischen Intention des Sprechers bzw. dem Effekt auf den Hörer als notwendig zu implizieren.

4.4 Methodische Zusammenfassung

Als Schluss dieses theoretischen Teiles meiner Arbeit und als Einführung in die Darlegung meiner empirischen Untersuchung fasse ich die entscheidenden Bestandteile meiner Hypothese zusammen. Am Anfang steht die Feststellung eines Phänomens, das die Texte der interkulturellen Autoren im Vergleich zu denen der deutschen kennzeichnet. Es besteht in einem besonders kreativen Umgang mit der deutschen Sprache, der nach einigen Literaturkritikern ihre literarischen Werke charakterisiert und mit dem Sprachwechsel verbunden ist. Unter den verschiedenen möglichen linguistischen Realisierungen dieser Kreativität, wie u.a. ein besonderer Gebrauch von Metaphern bzw. von syntaktischen Strukturen, beschränke ich mich im Rahmen dieser Studie auf die Analyse der Verwendung von produktiven Wortbildungsmustern.

Auf der einen Seite folgt die Auswahl dieses Merkmals aus dem Willen, eine allen interkulturellen Schriftstellern gemeinsame sprachliche Eigenschaft zu erfassen, der seinerseits teilweise von methodologischer Anforderung geleitet ist. Die Feststellung einer Variation durch eine quantitative Methode in der Korpuslinguistik benötigt nämlich eine formalisierbare Suchabfrage, sodass z.B. die Frage nach der Bildung neuer Wörter anhand analogischer Muster bzw. nach der Anwesenheit von Bedeutungsabweichungen zwar untersuchbar ist, aber eine zusätzliche qualitative Analyse bzw. semantisch annotierte Korpora erfordert. Indem die Wortbildungsmuster Strukturen der Targetsprache bilden, könnten Variationen in ihrem Gebrauch die Sprache aller interkulturellen Autoren als fossilierte Merkmale kennzeichnen, die dagegen einem gemeinsamen Muster folgen.

Auf der anderen Seite begründet meine Wahl die Feststellung, dass kognitive und soziolinguistische Faktoren den Produktivitätsgrad einer morphologischen Kategorie beeinflussen, sodass der Zustand der interkulturellen Schriftsteller als L2-Lerner und als Emigranten eine Rolle in der Kodierung spielen könnte.

Der Produktivitätsgrad der Wortbildungsmuster auf *-ung*, den ich nach den Baayen'schen Formeln der *realized* und *potential productivity* operationalisiere, stellt meine abhängige Variable dar. Eine Variable ist diesbezüglich ein Symbol für eine Menge von Zuständen, und zwar ein Merkmal, das vorbestimmte Ausprägungen (nominale, ordinale oder *ratio*)[29] aufweist (vgl. Gries 2009). Im Fall

[29] Ausprägungen in der Nominalskala sind ungeordnete nominale Kategorien eines Merkmales (wie z.B. Akkusativ oder Dativ für die Variable „Kasus"); jene in der Ordinalskala sind dagegen geordnet (wie einfach < komplex mit Präpositionalphrase < komplex mit Nebensatz für die „Nominalphrase-Komplexität"), während jene in der *ratio*- bzw.

der *realized* und *potential productivity* handelt es sich um eine *ratio*-Variable, da ihre Ausprägungen numerisch sind.

Die erste unabhängige Variable, anhand der ich Variationen im Gebrauch meines linguistischen Merkmals untersuche, ist die *Nativeness*. Sie ist eine nominale Variable mit zwei Ausprägungen, „nativ“ und „nicht-nativ“, die jeweils den Autoren mit Deutsch als Muttersprache bzw. den interkulturellen Schriftstellern zugeschrieben werden. Meine Hypothese nimmt diesbezüglich eine Korrelation zwischen den zwei Produktivitätsmaßen und der *Nativeness* eines Schriftstellers an. Da die im dritten Kapitel erwähnte Literatur tendenziell für einen Übergebrauch der Wortbildungsmuster auf *-ung* durch die L2-Lerner spricht, ist meine Hypothese „gerichtet“. Die Null-Hypothese (H0) behauptet nämlich, dass der Produktivitätsgrad der morphologischen Kategorie *-ung* in den Texten der interkulturellen Autoren gleich oder kleiner als der bei den deutschen Schriftstellern sei, während die Alternativhypothese (H1) das Gegenteil postuliert. Beide Hypothesen lauten also wie folgt:

- **H0:** die *realized* und/oder *potential productivity* bei den interkulturellen Autoren = oder < die *realized* und/oder *potential productivity* bei den deutschen Autoren.
- **H1:** die *realized* und/oder *potential productivity* bei den interkulturellen Autoren > die *realized* und/oder *potential productivity* bei den deutschen Autoren.[30]

Ähnliches gilt teilweise für die nachfolgenden drei Hypothesen, die eine Korrelation zwischen dem Produktivitätsgrad des Affixes *-ung* und den anderen drei unabhängigen nominalen Variablen (*Auswanderungsgrund*, *L1* und *Autor*)[31] überprüfen. Es handelt in diesen Fällen jedoch um eine „ungerichtete“ Hypothese, deren H0 dementsprechend nur eine nicht-signifikante Variation der Produktivitätsgrade des Affixes *-ung* zwischen den Untergruppen annimmt. Die Gründe für diese untergeordnete Untersuchung erörtere ich genauer ab Kapitel 5.2.3. Hier genügt es anzumerken, dass diese unabhängigen Variablen mit den Identitätseigenschaften des *Status'*, der *Ethnizität* und der *Individualität* Gemeinsamkeiten aufweisen (vgl. Preston 1993 und Kap. 2.2). Von daher haben sie nicht nur psycholinguistische, sondern auch soziolinguistische und stilometrische Relevanz.

Verhältnisskala geordnet und proportional sind (wie die Länge in der Silbe eines Wortes). Die Art der Skala einer Variable ist dann für die Feststellung der möglichen anwendbaren statistischen Tests wesentlich (vgl. Gries 2009).

[30] Die Formel der *realized* und *potential productivity* der morphologischen Kategorie *-ung* sind: V(-ung, N)/N(-ung) bzw. V(1,-ung,N)/N(-ung). Dadurch können die Null- und Alternativhypothese in mathematisch-statistischen Formeln übersetzt werden (*operationalisiert*).

[31] Die Variable *Auswanderungsgrund* besteht aus drei Ausprägungen: „Exil“, „Arbeit (Gast-)“ und „Privat“ (d.h. private Gründe); *L1* aus zehn: „Italienisch“, „Polnisch“, „Bosnisch“, „Japanisch“, „Bulgarisch“, „Tschechisch“, „Rumänisch“, „Ungarisch“, „Türkisch“, „Arabisch“; *Autor* aus 15, die mit den von mir betrachteten Schriftstellern übereinstimmen: „Franco Biondi“, „Radek Knapp“, „Saša Stanišic“, „Yoko Tawada“, „Dimitré Dinev“, „Ota Filip“, „Catalin Florescu“, „Libuse Moníková“, „Terézia Mora“, „Emine Sevgi Özdamar“, „Yüksel Pazarkaya“, „Tzveta Sofronieva“, „Aglaja Veteranyi“, „Rafik Schami“ und „Léda Forgó“.

Bei den verschiedenen morphologischen Kategorien beschränke ich mich zunächst auf die Analyse der Wortbildungsmuster auf *-ung*, und zwar aus zwei Gründen: Einerseits bezeichnet die Literatur das Suffix *-ung* in der deutschen Gegenwartsprache als sehr produktiv (vgl. u.a. Motsch 2004, Fleischer et al. 2007), anderseits kann ich mich so auf die Magisterarbeit von Berth (2009) beziehen, die die *potential productivity* von *-ung* in einem Lernerkorpus berechnet (vgl. Kap. 3.1). Abgesehen von dem schon erwähnten niedrigen „Authentizitätsgrad" der lektorierten Texte meines Korpus' gegenüber dem von Berth besteht eine weitere wichtige Differenz zwischen unseren Arbeiten in der folgenden. Nach Motsch (2004) teilen die Derivate auf *-ung* eine gemeinsame Form, indem sie deverbale feminine Nomina sind. Sie lassen sich jedoch in vier semantische Muster mit unterschiedlicher Produktivität aufteilen, je nachdem, ob das Derivat ein Nomen Actionis, ein Nomen für den resultierenden Zustand, ein Nomina Acti oder ein Nomen Loci ist.

Nicht ohne Bestimmungsschwierigkeiten teilt Berth in seiner Magisterarbeit die Derivate auf *-ung* nach den vier semantischen Mustern qualitativ ein und berechnet die entsprechenden vier Produktivitätsmaße. Die Abbildung der tatsächlichen Wortschatzswachstum-Kurve (*vocabulary growth curve*) einer morphologischen Kategorie auf die Baayen'schen statistischen Modelle kann nämlich gerade wegen der fehlenden Homogenität eines Affigierungsprozesses bzw. der Anwesenheit unterschiedlicher syntaktischer und semantischer „Subpattern" missglücken (vgl. Lüdeling & Evert 2003). Trotzdem kann meine Arbeit eine qualitative Differenzierung und Berücksichtigung der unterschiedlichen semantischen Muster nicht leisten.

Schließlich untersuche ich die Korrelation des Produktivitätsgrades der drei morphologischen Kategorien *-keit/-igkeit/-heit*, *-bar* und *ver-* mit der Variable *Nativeness*: Ihre Eigenschaften bzw. die Gründe meiner Auswahl lege ich ausführlich Kapitel 5.2.4 dar.

5. Empirischer Teil

5.1 Der Korpusaufbau

5.1.1 Das Chamisso-Korpus

Für den Aufbau meines Korpus' versuche ich die methodologischen Hinweise in Biber (1993) so genau wie möglich zu befolgen, die das Vorgehen erörtern, um ein repräsentatives Korpus aufzubauen. Nach Biber (1993) hängt die Repräsentativität eines Korpus' von einer anfänglichen ausführlichen Definition der Target-Population, sowie von der Wahl der geeigneten Datenerhebungstechnik ab. Zum ersten Ziel schreibt Biber:

> „Definition of the target population has at least two aspect: (1) the boundary of the population – what texts are included and excluded from the population; (2) hierarchical organisation within the population – what text categories are included in the population, and what are their definitions" (Biber 1993: 243).

In Anlehnung an den ersten Punkt Bibers definiere ich zuerst die situationellen (nicht-linguistischen) Eigenschaften der zu untersuchenden Autoren und Texte. Da die Anzahl der interkulturellen Schriftsteller in Deutschland sehr hoch ist, stelle ich diesbezüglich eine erste Populationsgrenze fest, indem ich nur die 60 Autoren in Betracht nehme, die den literarischen *Adelbert-von-Chamisso Preis* gewonnen haben. Wie schon in Kapitel 1.3 erwähnt, wird der Chamisso-Preis seit 1985 jedes Jahr mehreren auf Deutsch schreibenden Autoren nicht deutscher Muttersprache verliehen.[32] Die Kriterien für die Vergabe des Preises sowie die Definition der *Chamisso-Literatur* stimmen ferner im Allgemein mit denen der interkulturellen Literatur seitens der Forscher des DaF-Institutes in München überein, insofern diese im Sprachwechsel einen grundsätzlichen Faktor für die Gruppierung sehen:

> „Die ausgezeichneten Autoren haben ganz unterschiedliche kulturelle Hintergründe und sind durch Arbeitsmigration, Asyl, Exil oder Studium nach Deutschland gekommen. Eines aber verbindet sie: Die deutsche Sprache, in die sie eingewandert sind und die sie zu ihrer eigenen und wichtigsten Ausdrucksform gemacht haben" (http://www.bosch-stiftung.de/content/language1/html/14169.asp. Zuletzt geprüft am 10.11.2011).

[32] Neben dem eigentlichen *Adelbert-von-Chamisso-Preis* werden in einigen Jahrgängen bis zu zwei *Förderpreise* vergeben. Für den Aufbau meines Korpus' betrachte ich alle Chamisso-Autoren unabhängig von der Art des gewonnenen Preises. Da die Textsammlung Ende 2010 begonnen hat und die Preisverleihung im Februar stattfindet, berücksichtige ich in meiner Untersuchung die Chamisso-PreisträgerInnen ab 2011 nicht mehr.

Unter den Chamisso-Autoren schließe ich zweitens von meiner Analyse diejenigen aus, die ihre Werke nicht auf Deutsch geschrieben haben, sondern sie haben übersetzen lassen, und denen man den Preis (fast ausnahmsweise) aufgrund ihrer Vorbildrolle für die nachfolgende interkulturelle Literatur verliehen hat. Dies sind die türkischen Schriftsteller Güney Dal und Aras Ören sowie der tschechische Dichter Jiří Gruša und der ungarische Autor Imre Kertész, wobei die letzten zwei zusammen mit Harald Weinrich die sogenannte Ehrengabe zum Chamisso-Preis gewonnen haben.

Eine weitere Beschränkung bezieht den Zeitpunkt des Erwerbs der deutschen Sprache ein. Ich will nämlich nicht die Sprache von völlig bilingualen Autoren untersuchen, sondern von denen, die Deutsch als Fremdsprache erst später gelernt und literale Kompetenzen auch in ihrer Muttersprachen erworben haben. So lasse ich in meiner Untersuchung diejenigen Autoren außer Acht, deren Eltern aus Deutschland stammen oder die mit weniger als zwölf Jahren nach Deutschland ausgewandert sind (jeweils zwei bzw. 21 Schriftsteller). Im Alter von zwölf Jahren hat man schon einen wichtigen Teil seiner Schulausbildung im Herkunftsland absolviert. Außerdem stellt das zwölfte Lebensjahr laut einigen Forschern einen Wendepunkt für den Spracherwerb dar, nach dem „a ‚critical' period in brain maturation has been passed and language development tends to freeze'" (vgl. Selinker 1972: 210 Fußnote 4 und Tarone 1994: 1718). Obwohl die deutsche Herkunft eines der Elternteile das Lernen der deutschen Sprache bei den Kindern nicht bewiest, ist schließlich die Wahrscheinlichkeit dieser Eventualität ziemlich hoch: Daher schließe ich die Texte von Sudabeh Mohafez und Artur Becker aus meiner Population aus.

Auf dieser Grundlage erwiesen sich vorbereitende biographische Recherchen über die Chamisso-Autoren schon im Laufe der ersten Phase meiner Arbeit als notwendig, außerdem haben sich einige Schwierigkeiten in der Auffindung und vor allem in der Kategorialisierung der relevanten biographischen Informationen ergeben, die hauptsächlich die Phase der Erstellung der Metadaten kennzeichneten (vgl. Kap. 5.1.3). Als Hauptquelle für diese gelten das Chiellino'sche Handbuch über die interkulturelle Literatur in Deutschland (2000) bzw. die von verschiedenen Schriftstellern herausgegebene Internetsseite und das im Netz verfügbare *Kritische Lexikon zur deutschsprachigen Gegenwartsliteratur* (KLG).

Die qualitative Begrenzung der Texte nehme ich zuletzt sowohl anhand situationeller, als auch anhand linguistischer Eigenschaften systematisch vor (vgl. den zweiten Punkt im Biber'schen Zitat). Um die Möglichkeit starker diachroner Effekte zu vermeiden, analysiere ich nur zwischen 1990 und 2010 veröffentlichte Texten. Die einzigen Gattungen, die ich in Betracht ziehe, sind die „literarischen", und unter diesen nur Romane, Erzählungen und Theaterstücke. Empirische Untersuchungen haben nämlich die Wirkung des Texttyps auf den Produktivitätsgrad einiger morphologischer Kategorien belegt (vgl. Baayen 1994a und Kap. 3.3.3). Eine Kontrolle des Themas

ist mir dagegen nicht möglich, aber wahrscheinlich auch nicht notwendig. Obwohl das Thema die lexikalische Auswahl eines Textes stark beeinflusst, ist es hinsichtlich seiner grammatikalischen bzw. morphologischen Eigenschaften nicht besonders signifikant.

5.1.2 Datenerhebungstechnik

Die quantitative Begrenzung der zu betrachtenden Texte betrifft dagegen die Frage der Strategie der Datensammlung, die ich wiederum von Biber (1993) aufgreife (vgl. aber auch Gries 2009: 24–29). Die von mir angewandte Datenerhebungsmethode ist diesbezüglich die der *stratified sampling*. Dafür identifiziere ich Untergruppen bzw. „Strata" innerhalb meiner Population (und zwar „Autor" und „Nationalität") und sammle Texte für eine gleiche Wörteranzahl je nach Untergruppen willkürlich (*random*).[33]

Die Wahl dieser Technik sowie der zwei Parameter für die Definition der Unterkategorien hängt von meiner Fragestellung und von einigen Eigenschaften meiner Daten ab. Insofern ich Unterschiede im Sprachgebrauch im Verhältnis zu vorbestimmten Variablen bzw. Populationsschichten untersuche und die Baayen'schen Produktivitätsmaße sich als empfindlich gegenüber der Subkorpusgröße erweisen, betrachte ich dieselbe Tokenanzahl pro Subkorpora, um den Erhalt vergleichbarer Ergebnisse schon in dieser ersten Phase meiner Arbeit zu sichern (vgl. Säily 2010, Kap. 3.3.3, Fußnote 33 und Kap. 5.2.2 für die abschließende Erklärung meiner Herangehensweise an dieses Problem). Da der Transfer in vielerlei Hinsicht die *Interlanguage* eines L2-Lerners beeinflusst bzw. sich Variationen in der Ausnutzung der Wortbildungsstrategien teilweise auf individuelle Neigungen zurückführen lassen, kann man darüber hinaus nicht ausschließen, dass diese zwei Faktoren auf die Produktivität einer morphologischen Kategorie einwirken. Obwohl meine Analyse zuerst auf die Bestimmung gemeinsamer Eigenschaften in der Sprache der interkulturellen Schriftsteller abzielt (vgl. Kap. 1.4 und 4.4), ermöglicht mir die

[33] Eine andere Methode für die Datensammlung bildet die des *simple random sampling*. Da in diesem Fall alle Texte in einer Population dieselbe Chance haben, für das Korpus ausgewählt zu werden, gewährleistet aber ihre Anwendung im Gegensatz zu jener des *stratified sampling* nicht, dass jedes für eine Untersuchung relevante Populationsstratum adäquat in dem Korpus repräsentiert wird (vgl. Biber 1993: 244). Von *proportional stratified sampling* spricht man, wenn die Anzahl der Texte für jeden betrachteten Stratus das tatsächliche Verhältnis innerhalb der Target-Population widerspiegelt, sodass z.B. in einem nach Register kontrollierten Korpus des gegenwärtigen Deutsch 90 % der Texte zur Konversation und 10 % zum Zeitungsartikel gehören sollen. Diesbezüglich unterscheidet Gries (2009) zwischen der Repräsentativität und der Ausgewogenheit einer Textsammlung: „I call a sample *representative* when the different parts of the population are reflected in the sample, and I call a sample *balanced* when the sizes of the parts in the population are reflected in the sample" (2009: 24). Biber bemerkt aber, dass solche Korpora oft keine angemessene Basis für linguistische Analysen bilden, da linguistische Studien nicht auf die Bestimmung der absoluten Frequenz einer Variable, sondern eher auf die Feststellung der Unterscheidungsmerkmale zwischen Strata bzw. Varietäten abzielen (1993: 347–348).

Schichtung meines Korpus' anhand der zwei Parameter „L1" und „Autor" die Durchführung einer Kontrollanalyse, die positive Ergebnisse liefern wird (vgl. Kap. 5.2.3 und ff.).

Die Texte sind teilweise im Internet verfügbar, teilweise musste ich sie aus Büchern scannen und sie durch die Technologie OCR in eine Word-Datei umwandeln.[34] Die Textsammlung findet *random* statt in dem Sinn, dass ich nach der Feststellung der Populationsgrenze bzw. -strata die Texte zufällig entnehme, ohne mich z.B. auf den Incipit der Erzählungen oder auf die bekannteren Romane zu beschränken. Das so entstandene Chamisso-Korpus enthält Texte für eine Gesamtsumme von 200.000 Wort-Tokens, die von fünfzehn verschiedenen Autoren mit zehn unterschiedlichen Nationalitäten geschrieben wurden, wobei jede Kategorie innerhalb der Untergruppen jeweils 10.000 bzw. 20.000 Wort-Tokens enthält.

Mit der Hilfe einiger Mitarbeiter des Institutes für Morphologie und Korpus-Linguistik der Humboldt-Universität zu Berlin habe ich schließlich das Chamisso-Korpus bei ANNIS2 hochgeladen, um meine quantitative Analyse durchzuführen. ANNIS2 ist „an open source, versatile web browser-based search and visualization architecture for complex multilevel linguistic corpora with diverse types of annotation": Sie gestattet die Formulierung von speziellen Suchanfragen (*Query)* auf verschiedenen linguistischen und nicht-linguistischen Ebenen, die automatische Wiederauffindung der Belege und den Export der Ergebnisse (vgl. http://www.sfb632.uni-potsdam.de/d1/annis/. Zuletzt geprüft am 10.11.2011).

Im Nachhinein sind mir jedoch einige Fehler in der Zusammensetzung meines Korpus' klar geworden. Der erste Fehler besteht in der Berechnung der Tokenanzahl, die anfänglich die Interpunktion nicht mitgezählt hat. Insofern man die Interpunktionszeichen in dem DEREKO-Korpus mitzählt (vgl. Kap. 5.1.4), basiert der Aufbau meines Vergleichskorpus' auf der totalen Anzahl der Tokens des Chamisso-Korpus' – die Interpunktion eingeschlossen (237.502 Tokens). Der zweite und der dritte methodologische Fehler sind dagegen wenig einfach behebbar. Das Chamisso-Korpus besteht aus 49 Texten mit unbeständigen Längen, wobei der Mittelwert 4.028 Tokens, die Varianz aber sehr hoch ist. Da die Wahrscheinlichkeit der Auffindung eines neuen Typs bzw. eines Hapax Legomena einer morphologischen Kategorie mit der Zahl der nacheinander auftretenden Wörter nicht konstant abnimmt (d.h. sie ist nicht nur von der Korpusgröße, sondern auch von der Textlänge abhängig), könnte die Variabilität der Letztgenannten meine Ergebnisse verfälschen.

[34] Die Internetseite der Verlage bzw. einiger digitaler Buchhandlungen bietet manchmal Leseproben der Werke der Chamisso-Schriftsteller im pdf.Format an. Unter OCR-Software (*Optical Character Recognition*) versteht man Computerprogramme für die automatische Texterkennung innerhalb von Bildern und deren Export u.a. als Textdatei. Ich nutze die Version 2.6 von FreeOCR.net, die man frei im Internet herunterladen und auf die deutsche Sprache einstellen kann (http://svsload.com/wd/freeocrnet-vsa/. Zuletzt geprüft am 10.11.2011). Trotzdem ist eine manuelle Korrektur der dadurch erstellten Word-Dokumente notwendig.

Zweitens verzichte ich infolge der Nicht-Verfügbarkeit ihrer Texte darauf, 20 verschiedene interkulturelle Schriftsteller in meiner Analyse in Betracht zu nehmen. Für fünf Muttersprachen (und zwar Italienisch, Polnisch, Arabisch, Bosnisch und Japanisch) enthält nämlich das Chamisso-Korpus Texte nur eines einzigen Autors. In diesen Fällen kann man daher einen eventuellen signifikanten Unterschied in den Produktivitätsmaßen von *-ung* nicht eindeutig auf eine Interferenz der Muttersprache bzw. eine individuelle Neigung zurückführen (vgl. Kap. 5.3).
Schließlich wäre ein größerer Korpus für die Auswertung zumindest eines Teils meiner Fragestellungen zuverlässiger gewesen. Insofern auch Baayen (1994a) die Produktivität einiger Wortbildungsmuster auf der Basis von Subkorpora von 25.000 Tokens vergleicht, ist die Rechnung der *realized* und *potential productivity* nach der Variable *Nativeness* in meinem Korpus noch veritabel (anhand 237.502 Tokens). Wegen des Mangels an Daten sind dagegen die Werte für *Muttersprache* und *Autor* mit Vorsicht zu betrachten (jeweils anhand 20.000 und 10.000 Tokens).

5.1.3 Metadaten

Wie oben erwähnt wurde, habe ich für jeden Text des Chamisso-Korpus' zusätzliche Informationen erhoben, die durch ANNIS2 als Metadaten zu jedem Beleg zugänglich sind und die man für das Filtern der Suchabfrage verwenden kann. Die Anordnung der bio- und bibliographischen Auskünfte geschieht anhand von 33 Attributen. Sie erfassen sowohl Eigenschaften der Texte, wie den Titel, das Veröffentlichungsjahr und die Gattung, als auch der Autoren, wie Geburtsjahr und -ort, die Schulbildung oder den Beruf (das Beispiel einer Metadaten-Tabelle in ANNIS2 findet sich in Anhang 1).
Die Attributswerte drücke ich manchmal in einer diskursiven Form aus, wie z.B. bei den „Literaturangaben" oder den „Anmerkungen". Die Meisten „variabilisiere" ich jedoch, indem ich einheitliche Ausprägungen feststelle, anhand derer die relevanten Informationen kategorisiert werden. Um das Gebilde von Untergruppen innerhalb meiner Population zu vereinfachen, trage ich deshalb z.B. die Profession eines Schriftstellers in einer einzigen neutrale Form ein, also unabhängig von dem Geschlecht – das übrigens ein selbständiges Attribut bildet –, während die Bestimmung der Textgattung anhand von vier Werten (Roman, Erzählung, Erzählungsband und Theaterstück) geschieht, die eine komplexere Taxonomie abkürzen. Manchmal erweist sich jedoch dieser Variabilisierungsprozess anhand systematischer Kriterien als schwierig, sodass ich im Lauf der Arbeit mehrere Male die Bezugskategorie verändern musste.

Als Beispiel kann mein Vorgehen bezüglich zweier fragwürdiger Fälle hinsichtlich des Werts des Attributs „Auswanderungsgrund“ dienen, den ich in dieser Arbeit in drei Ausprägungen (Exil, Arbeit (Gast-) und Privat) kategorisiere. Rafik Schami emigrierte mit einem Vertrag als Gastarbeiter nach Deutschland, nachdem er aus Syrien wegen seiner literarischen Tätigkeit in den Libanon geflüchtet war. Der Grund der ersten Auswanderung aus Syrien ist daher einem Exil ähnlich. Insofern aber Schami in Deutschland als Gastarbeiter wohnt bzw. eine wichtige Rolle in der Entwicklung der Gastarbeiterliteratur eingenommen hat (z.B. zusammen mit Franco Biondi als Mitbegründer der *PoLiKunst*-Vereinigung und Mitherausgeber der Anthologiereihe *Südwind Gastarbeiterdeutsch*, vgl. Kap. 1.2), trage ich für ihn „Arbeit (Gast-)“ als Auswanderungsgrund ein. Yüksel Pazarkaya dagegen zog für das Studium aus der Türkei nach Stuttgart um. Während des Universitätsstudiums zuerst als Chemiker und dann als Germanist wurde er jedoch eine Bezugsfigur für die (vor allem türkischen) Emigranten in Deutschland und deren Literatur. Als Kurator der Zeitschrift *Anadil* und Mitarbeiter im Ararat-Verlag förderte er zum Beispiel einerseits die Veröffentlichung bzw. Übersetzung von Texten von Gastarbeitern und interkulturellen Autoren, andererseits publizierte er selbst Erzählungsbände wie *Heimat in der Fremde?*, welche die typischen Themen der „Gastarbeiterliteratur“ behandeln. Trotzdem schreibe ich als Wert des Attributes *Auswanderungsgrund* Pazarkaya „Privat“ zu, wobei ich die zusätzlichen Informationen in den *Anmerkungen* annotiere.

Die Attribute meiner Metadaten übersteigen darüber hinaus die für die Auswertung meiner Hypothese hinreichenden, sodass die Möglichkeit weiterer Untersuchungen mit unterschiedlichen Fragestellungen bzw. verschiedenen unabhängigen Variablen anhand meiner Daten besteht. Obwohl ich sie im Rahmen dieser Analyse nicht in Betracht nehme, seien hier einige in den Metadaten annotierte Eigenschaften erwähnt, die auch einen gewissen Einfluss auf den Produktivitätsgrad eines Wortbildungsmusters haben könnten. Das Geschlecht ist eine davon, insofern dieses nach Säily (2011) Variationen in der *potential productivity* der nominalen Affixe -*ness* und -*ity* im modernen Englisch verursacht.

Eine andere Eigenschaft bilden die weiteren von einem Autor gesprochenen Fremdsprachen. Wenn nämlich das Deutsche wie etwa bei Tzveta Sofronieva als L5 erworben wird, könnten Besonderheiten in der Anwendung der Wortbildungsstrategie auch von der Interaktion mit den anderen Fremdsprachen und nicht nur von der Muttersprache abhängen (vgl. González Àlvarez 2004: 52–53). Schließlich beeinflussen wahrscheinlich auch der Aufenthaltsort in Deutschland und daher der Dialekt, mit dem man zuerst in Kontakt getreten ist, den Produktivitätsgrad einer morphologischen Kategorien, wie es die Studie von Keune et al. (2006) über die regionalen Variationen der Produktivität im Niederländischen beweist.

5.1.4 Vergleichskorpus

Insofern meine Hypothese einen Übergebrauch einiger produktiver Wortbildungsmuster des Deutschen bei interkulturellen Autoren annimmt, benötige ich für die Auswertung den Aufbau eines vergleichbaren Korpus' von Texten deutscher Schriftsteller, anhand dessen ich die Existenz eines signifikanten Unterschiedes hinsichtlich meiner Variable feststellen kann (vgl. Kap. 4.2). Diesbezüglich ist die Wahl des Vergleichskorpus' in einer *contrastive interlanguage analysis* besonders wichtig, vor allem wenn sie auf die Bestimmung von Frequenzunterschieden abzielt:

> „Da Korpusfrequenzen von sehr vielen Faktoren abhängen, müssen die Vergleichskorpora in möglichst vielen Designparametern übereinstimmen – im Idealfall sollten sie sich *nur* in dem zu betrachtenden Parameter unterscheiden" (Lüdeling & Walter 2010: 319).

Als Vergleichskorpus nehme ich eine Teilmenge des DEUTSCHEN REFERENZKORPUS (DEREKO), das aus literarischen, zwischen 1990 und 2001 veröffentlichten Texten von Autoren mit Deutsch als Muttersprache besteht. Das DEREKO-Korpus entsteht aus der Zusammenarbeit des *Institutes für Deutsche Sprache* (IDS) in Mannheim, des *Institutes für Maschinelle Sprachverarbeitung* (IMS) der Universität Stuttgart und des *Seminars für Sprachwissenschaft* (SfS) der Universität Tübingen. Es enthält zwischen 1956 und 2001 auf Deutsch geschriebene Texte in einem Umfang von über 4,1 Milliarden Textwörtern (Stand: 3. April 2011), die man auf der Plattform COSMASII abfragen kann (vgl. http://www.ids-mannheim.de/kl/projekte/dereko_I/. Zuletzt geprüft am 10.11.2011).
Es erweist sich aus verschiedenen Gründen für meine Untersuchung als geeignet. Erstens ist das DEREKO-Korpus nach Anmeldung im Internet frei zugänglich. Es beinhaltet darüber hinaus einen Subkorpus mit dem Titel „Belletristik des 20. und 21. Jahrhunderts" und kann nach vorbestimmten Parametern wie der gesamten Tokenanzahl, der Gattung und dem Veröffentlichungsjahr des Textes, die ich in meiner Analyse kontrollieren möchte, reduziert werden. Schließlich sind die Tokens des DEREKO-Korpus' nach morphologischer Struktur annotiert, sodass man durch die Anwendung eines Grundformoperators (&) und die Ausschaltung der verschiedenen Lemmatisierungsoptionen alle Belege von zusammengesetzten sowie mehrfach-zusammengesetzten Wörtern einer morphologischen Kategorie vollautomatisch identifizieren kann (vgl. Kap. 5.2.1).
Leider ist jedoch eine Reduktion der Tokenanzahl der schon hochgeladenen Texte nicht möglich. Dem Chamisso-Korpus (237.502 Tokens) ähnlich enthält daher mein Vergleichskorpus Texte unterschiedlicher Länge in einem Umfang von 237.501 Tokens (die Interpunktionszeichen eingeschlossen). Im Gegensatz zu Ersterem handelt es sich aber nur um zwölf Texte von zwölf verschiedenen Autoren, sodass der Mittelwert bei circa 19.792 Tokens liegt. Angesichts der oben

erwähnten Korrelation zwischen den Produktivitätsmaßen und der Textlänge stellt die größere Homogenität des Vergleichskorpus' einen nicht unbedeutenden Einwand gegen die Vergleichbarkeit meiner Daten dar. Die genaue Zusammensetzung des Chamisso- und des Vergleichskorpus' veranschaulichen Anhang 2 und Anhang 3.

5.2 Die empirische Untersuchung

5.2.1 Die deskriptive Analyse: Auswahl und Zählung der Belege

Die Auffindung der Derivate auf *-ung* in den zwei Korpora erfolgte mit unterschiedlichen Verfahren. In dem Vergleichskorpus geschieht die Suche der Ableitungen auf *-ung* auch in den Fällen von mehrfacher Affigierung bzw. bei der Anwendung weiterer morphologischer Prozesse vollautomatisch, indem die Suchabfrage „&-ung" eingegeben wird und alle Lemmatisierungsoptionen aktiviert werden. Das Chamisso-Korpus ist dagegen nicht nach morphologischen Strukturen annotiert. Durch die Suchabfrage „lemma=/.+ung.*/" erhalte ich daher zuerst die Lemmata, in denen die Zeichenkette <ung> vorkommt.[35] Nach dem Export der Belege beseitige ich manuell all jene Tokens, die kein Derivat auf *-ung* bilden (wie z.B. J*ung*, Z*ung*e oder H*ung*er).

Für die Speicherung meiner Ergebnisse erstelle ich zwei Tabellen in Excel: die eine für das Chamisso-, die andere für das Vergleichskorpus, wobei jede Zeile eine „Beobachtung" darstellt. Eine Beobachtung besteht aus einer Spalte für den Eintrag des *Tokens*, also der in dem Text auftretenden Wortform, aus einer Spalte für das *Lemma* und aus einer letzten für die Häufigkeit, mit der das Wort vorkommt. Die *Vorkommenshäufigkeit* ist in dieser ersten Tabelle immer 1, weil jeder Beleg eines Derivates auf *-ung* eine Beobachtung bildet (als Beispiel der Beobachtungstabelle vgl. Anhang 4).

Die angewandte Methode für die Tokenisierung und die Lemmatisierung der Ableitungen auf *-ung* erfordert hier zusätzliche Erklärungen, insofern sie einen wesentlichen Einfluss auf die Ergebnisse hat. Auf der Basis der Frequenz der unterschiedlichen Lemmata einer morphologischen Kategorie stellt man nämlich die Anzahl ihrer Typen bzw. ihrer Hapax Legomena fest, wobei deren Division

[35] Die Verwendung der Operatoren „.", „+" und „*" schränkt die Suche zudem nur auf diejenigen Tokens ein, bei denen zumindest ein Buchstabe (.+) dem Trigramm <ung> vorangeht, bzw. erweitert sie auf die Belege, in denen eventuell ein Buchstabe (.*) <ung> folgt (vgl. z.B. Komposita mit einem Derivat auf *-ung* als Determinans). Die Spezifizierung „lemma" schließt dagegen die Nicht-Wörter aus der Suche aus. Die lexikalische Kategorie (*part of speech*) definiere ich in meiner Suchabfrage nicht, insofern ein Derivat auf *-ung* infolge der Anwendung weiterer Wortbildungsstrategien zu verschiedenen Wortarten gehören kann.

durch die Gesamtsumme der Tokens derselben morphologischen Kategorie wiederum die *realized* und *potential productivity* des Wortbildungsmusters ergibt (vgl. Kap. 3.3.2).

Bezüglich der Tokenisierung verursachen einerseits die „heavily lexicalized Items“ bzw. einige „baseless derivative“ Bestimmungsprobleme, insofern man sie von einem synchronen und grammatikalischen Standpunkt nicht in jeder Hinsicht als Derivate betrachten kann, wohingegen sie teilweise eine erkennbare morphologische Struktur aufweisen. Andererseits zeigt sich die Mitrechnung der Ableitungen aus „inner derivational cycles“ auf einer methodologischen Ebene als fragwürdig, da sie die Übereinstimmung zwischen Wort-Tokens und Affix-Tokens in einem Korpus aufhebt, sodass man die Verteilung von zwei Affixen nicht mehr als ein statistisch unabhängiges Ereignis erfassen kann (vgl. Gaeta & Ricca 2006: 74–83).

Obwohl nach Gaeta und Ricca die Ausschließung dieser drei Fälle von der Auszählung keine signifikante Wirkung auf die Messung des Produktivitätsgrades einer morphologischen Kategorie in großen Korpora aufweist, führt mich u.a. die geringe Größe meines Korpus’ im Vergleich zu dem von den zwei Forschern ausgewerteten zu einer unterschiedlichen Entscheidung bezüglich dieser Frage. Dafür sprechen auch weitere Aspekte der genaueren Beschaffenheit meiner Fragestellung bzw. der Art meiner Untersuchung. Erstens sind auch die Derivate aus inneren Derivationszyklen durchaus bedeutend, um Schlussfolgerungen über die morphologische Kompetenz eines Sprechers zu ziehen. Diesbezüglich weisen nämlich Gaeta und Ricca darauf hin, dass „[t]here might good theoretical (and/or psycholinguistic) reasons to include these lexemes in our sample, since they might induce the activation of the respective suffixes, thus influencing their availability in the mental lexicon“ (2006: 75). Zweitens sind die Konsequenzen einer fehlerhaften Entsprechung zwischen Wort- und Affixanzahl in meiner Analyse verglichen mit anderen Studien geringer, da ich nur vier morphologische Kategorien des Deutschen beachte bzw. ihre Produktivitätsgrade nicht miteinander vergleiche (z.B. durch die Formel der *expanding productivity*). Schließlich würde eine qualitative Bestimmung der stark lexikalisierten Items in einigen Fällen nicht triviale und teilweise willkürliche Aufteilungen benötigen, insofern, wie Gaeta und Ricca erwähnen, „the boundaries in this matter are [...] rather fuzzy“ (2006: 75). Daher betrachte ich in meiner Untersuchung die Ableitungen auf *-ung* der oben erwähnten Fälle prinzipiell als Token. Ein positives Einbeziehungsskriterium ist dagegen ihre Definition als Derivat im Deutschen Universalwörterbuch Duden (2003).[36]

[36] Als weitere Erklärung bringe ich hier einige Bespiele. Obwohl sie stark lexikalisiert sind, betrachte ich einerseits Derivate wie „Wohnung“ oder „Richtung“ als Tokens. Anderseits stellt auch das Kompositum „Kinderabteilung“ bzw. das mehrfache Derivat „regungslos“ einen Beleg der Anwendung eines Wortbildungsmusters auf *-ung* dar (für deren Lemmatisierung siehe unten). Die von Gaeta und Ricca (2006) erwähnte weitere Typologie der „baseless derivative“ betrachte ich dagegen in meiner Untersuchung nicht. Sie besteht nämlich aus Wörtern, die Pseudo-Affixe enthalten und daher im Duden (2003) als Simplex definiert werden.

Hinsichtlich des Lemmatisierungsverfahrens befolge ich grundsätzlich die normalen Methoden für die Feststellung der Grundform eines Wortes, wobei z.B. das Lemma eines Substantives der Nominativ Singular bildet, während das eines Verbes den Infinitiv Präsens Aktiv. Zur Bestimmung der zweifelhaften Fälle dagegen habe ich die folgenden Richtlinien für diese Arbeit ad hoc formuliert:

1. Bei Komposita mit einem Derivat auf *-ung* bildet nur das Derivat auf *-ung*, unabhängig von seiner Rolle als Determinans oder Determinatum, das Lemma. Das Lemma der Tokens „Gehirnforschung“ und „Forschungskolloquium“ ist daher in beiden Fällen „Forschung“. Da dies die Zahl der unterschiedlichen Lemmata deutlich verkleinert, hat diese Vorgabe einen großen Einfluss auf die Berechnung der *realized* und *potential productivity* der morphologischen Kategorien in meinen Korpora.
2. Komposita mit zwei Derivaten auf *-ung* bilden zwei Beobachtungen, welche dasselbe Token, aber unterschiedliche Lemmata aufweisen. Eine Bildung wie „Störungsforschung“ trage ich daher in zwei Zeilen meiner Tabelle ein, die als Token „Störungsforschung“, aber als Lemma „Störung“ bzw. „Forschung“ haben.
3. Da *-ung* ein deverbales nominales Affix ist, geht die Feststellung des Lemmas bei mehrfachen Wortbildungen und/oder inneren derivationalen Zyklen mit der Bestimmung des Verbs einher, das als Basis des Derivates fungiert (ähnliches gilt für die anderen betrachteten Wortbildungsmuster, wobei die Basis jeweils unterschiedliche syntaktische Kategorien aufweist, vgl. Kap. 5.2.4). Falls die verbale Basis einen selbständigen Eintrag im Wörterbuch Duden (2003) bildet, ist das Lemma selbst ein mehrfaches Derivat. Insofern das Verb „einnebeln“ im Wörterbuch vorkommt, ist daher z.B. das Lemma von „Einnebelungen“ „Einnebelung“.
4. Bei Wortbildungen durch Affixoide[37] folge ich denselben Kriterien wie unter Punkt 3. beschrieben. Das Token „Selbstbeherrschung“ weist deshalb als Lemma „Beherrschung“ auf, insofern „selbstbeherrschen“ im Duden (2003) nicht vorhanden ist, während das Lemma von „Fernheizung“ „Fernheizung“ ist, da „fernheizen“ als selbständiger Eintrag im Duden zu finden ist.
5. Komposita, deren Bestandteile durch Bindestrich verbunden sind (z.B. Heim-Zahlung), betrachte ich genau wie die anderen Komposita.

[37] Unter Affixoiden versteht man Wortbildungsmorpheme, deren Form der eines freien Morphems ähnlich ist, deren Bedeutung aber einem Prozess semantischer Entleerung unterliegt (vgl. Meibauer 2007: 57–58). Der Begriff „Affixoide“ ist in der Literatur umstritten. Ich verwende hier diesen Ausdruck nicht so sehr, um eine Stellung in dieser Diskussion einzunehmen, sondern eher um meine Richtlinie in einer allgemeineren Weise zu formulieren bzw. Missverständnisse zu vermeiden.

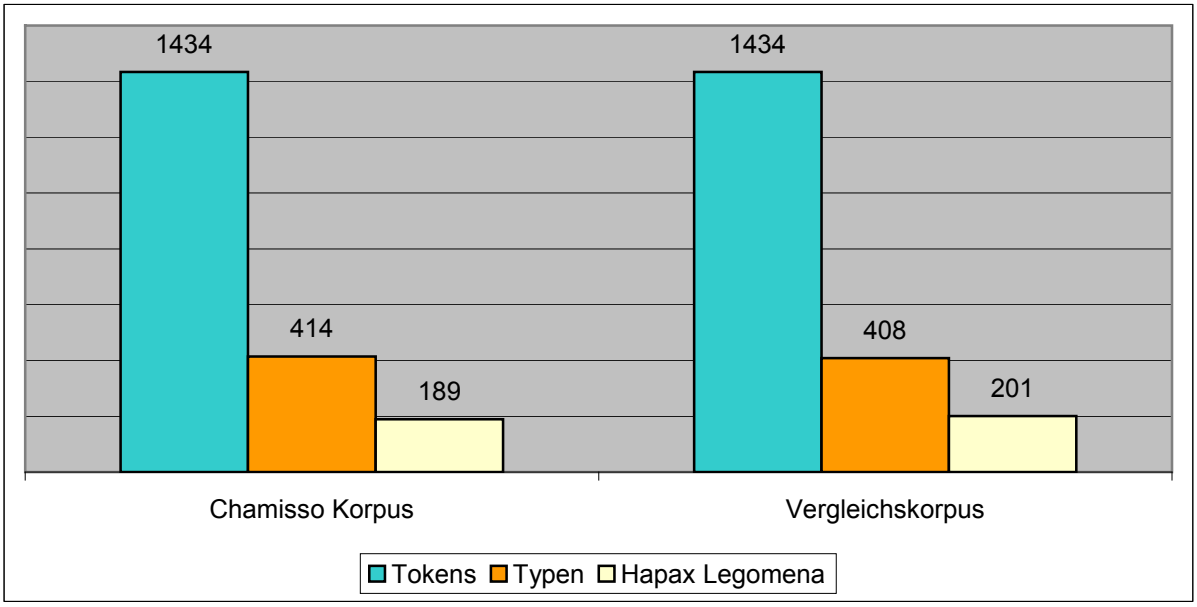
Die Verteilung der Tokens, Typen und Hapax Legomena der morphologischen Kategorie *-ung* im Chamisso- und Vergleichskorpus erhalte ich durch die Erstellung einer Pivot-Tabelle in Excel auf Basis der Daten meiner Beobachtungstabelle. Um die Maße der *realized* und *potential productivity* zu berechnen, betrachte ich dann Gaeta und Ricca (2006) folgend die Anzahl der Typen und Hapax Legomena auf *-ung* für eine normalisierte Tokenanzahl (s.u.).

5.2.2 Analytische Analysis I: die Auswertung der Belege nach der Variable *Nativeness*

Einer der heftigsten Einwände gegen die Baayen'sche Formel der *realized* und *potential productivity* betrifft die kontraintuitiven Ergebnisse, welche ihre Anwendung bei dem Vergleich von Affixen mit sehr unterschiedlicher Token-Frequenz hervorbringt. Nach Gaeta und Ricca (2006: 62–66) entkräftet diese Schwäche trotzdem nicht die Methode an sich, sondern kann durch eine Änderung in ihrer Verwendung umgegangen werden. Da die Produktivität einer morphologischen Kategorie eine „monoton fallende Funktion" der Zahl der betrachteten Tokens (N) ist, verursacht die fehlende Normalisierung der Variable (N) eine Überschätzung des Produktivitätsgrades insbesondere bei den weniger häufigen Affixen.[38] Der Fehler verschwindet jedoch bei einer Messung „for equal values of N" (Gaeta & Ricca 2006: 63).

Obwohl sich im Vergleichskorpus 1.628 Ableitungen auf *-ung* befinden, betrachte ich auf dieser Grundlage für die Rechnung der zwei Produktivitätsmaße also nur die ersten 1.434 Belege, indem ich den Wert auf die Tokenanzahl des Chamisso-Korpus' normalisiere. Die normalisierte Verteilung der Tokens, Typen und Hapax Legomena der Wortbildungsmuster auf *-ung* im Chamisso- und Vergleichskorpus stellt die Grafik in der Abbildung 1 dar (s.u.). Diesbezüglich erweisen sich die Unterschiede zwischen ausländischen und deutschen Schriftstellern in der *realized* bzw. *potential productivity* als nicht signifikant für einen p-Wert > 0,05. Eine genauere Betrachtung der Daten zeigt, dass der häufigste Typ des Chamisso-Korpus', „Wohnung", 107 Mal gegenüber den 58 Malen des häufigstem im Vergleichskorpus, „Richtung", vorkommt. Da diese Asymmetrie Topik-bedingt sein könnte, schließe ich sie von der Rechnung aus. Die Variationen in den zwei Produktivitätsmaßen bleiben trotzdem nicht-signifikant.

[38] Die Wahrscheinlichkeit des Vorkommens einer Hapax Legomena bzw. eines neuen Typs nimmt nämlich mit desto größerer Geschwindigkeit ab, je mehr Tokens einer morphologischen Kategorie man betrachtet (d.h. ihr Verhältnis ist monoton fallend). Indem sie auf einer Division zwischen nicht normalisierten Werten basieren, setzen dagegen die Baayen'sche Formeln der Produktivität ein konstantes Verhältnis zwischen Hapax Legomena bzw. Typen und Tokens voraus, als ob die Produktivität eine monoton nicht-fallende Funktion der Tokenanzahl wäre (vgl. Gaeta und Ricca 2006: 58–59).



	Chamisso-Korpus	Vergleichskorpus
Realized productivity	0,289	0,285
Potential productivity	0,132	0,14

Abbildung 1: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *Nativeness* und Berechnung der *realized* und *potential productivity*.

Um die Signifikanz meiner Ergebnisse zu prüfen, nutze ich den *Test of Equal or Given Proportions* (prop.test) der freien Software für statistische Rechnung R, den u.a. Zeldes (2009) für den Vergleich der Produktivitätsgrade von morphologischen Kategorien verwendet. Der prop.test stellt nämlich fest, ob das Verhältnis von Erfolgen und Versuchen in zwei oder mehreren Stichproben dasselbe ist, und daher, ob die Ergebnisse unter eine selbe binomiale Verteilung passen (vgl. http://stat.ethz.ch/R-manual/R-patched/library/stats/html/prop.test.html.Zuletzt geprüft 10.11.2011). Indem ich den prop.test in diesem Zusammenhang als Signifikanz-Test verwende, betrachte ich die Zahl der Typen bzw. der Hapax Legomena der Wortbildungsmuster auf *-ung* als Erfolge und die normalisierte Zahl seiner Tokens als Versuche, wobei ich die Wahrscheinlichkeit prüfe, mit der die Wahrscheinlichkeit der Erfolge in beiden Subkorpora dieselbe war. Dieser statistische Test erweist sich daher mit der Baayen'schen Auffassung der *realized* und *potential productivity* voll verträglich: Die zwei Maße werden nämlich als die Wahrscheinlichkeit definiert, mit der das nächste Token einer morphologischen Kategorie in einem bestimmten Korpus einem neuen Typ bzw. einem Hapax Legomena angehört (vgl. Kap. 3.3.2). Ein Bespiel eines prop.test auf Basis meiner Daten stellt Anhang 5 dar.

Hier sei trotzdem erwähnt, dass die Literatur sich bezüglich des zu nutzenden statistischen Tests nicht einig ist und einige Forscher andere Methoden für den Vergleich zwischen

Produktivitätsmaßen verwenden. Einerseits entwickelt Säily (2011: 125) z.B. ein Verfahren, das auf der Erstellung von „accumulation curves“ und der Anwendung des „permutation testing“ beruht, und sich der Autorin zufolge als besonders geeignet für Subkorpora mit unterschiedlicher Größe erweist. Andererseits vergleichen Keune et al. (2006) drei unterschiedliche statistische Modelle in ihrer Vorhersagefähigkeit hinsichtlich der Varianz der *potential productivity* einiger niederländischer Affixe. Interessanterweise ist das Modell, das am besten zu ihren Daten passt, „a generalized linear model with a binomial link function“, in dem die „hapax legomena were considered as successes, and all remaining words in the sub-corpus were counted as failure“ (2006: 4).

5.2.3 Analytische Analysis II: die anderen unabhängigen Variable

Eine ähnliche Vorgehensweise verwende ich für die Auswertung der drei nachfolgenden Hypothesen, die eine Variation der *realized* und *potential productivity* in Korrelation mit weiteren drei unabhängigen Variablen annehmen (*Auswanderungsgrund*, *L1* und *Autor*). Wie schon in Kapitel 4.2 erwähnt wurde, ziehe ich zu diesem Zweck das Vergleichskorpus nicht mehr in Betracht, wobei ich das Chamisso-Korpus in durch meine Variablen definierten Subkorpora unterteile und die normalisierte Anzahl der Tokens, Typen und Hapax Legomena der morphologischen Kategorie *-ung* pro Untergruppe feststelle.
Die Belege der Zeichenkette <ung> in einem bestimmten Chamisso-Subkorpus findet man diesbezüglich durch ANNIS2 vollautomatisch, indem man in dem *Query* den Wert des entsprechenden Metadatenattributes eingibt. Um z.B. die Vorkommen von <ung> bei denjenigen Autoren, die nach Deutschland als Exil ausgewandert sind, auszusortieren, ist die Suchabfrage „lemma=/.+ung.*/ & meta::Auswanderungsgrund1="Exil"“. Die Ergebnisse sind jeweils auf den Wert des Auswanderungsgrundes „Arbeit (Gast-)“, der Muttersprache „Bosnisch“ und des Autors „Catalin Florescu“ normalisiert. Da meine Hypothese bezüglich einer Variation der Wortbildungsmuster auf *-ung* für alle drei Variablen „ungerichtet“ ist, sind die Alternativen in dem prop.test "two.sided".

5.2.3.1 Die Variable Auswanderungsgrund

Die erste Variable bildet der Auswanderungsgrund, den ich in die Ausprägungen „Exil“, „Arbeit (Gast-)“ und „Privat“ aufteile (vgl. Kap. 5.1.3). Diese Ausprägungen symbolisieren etwas

vereinfacht einen unterschiedlichen Status im Empfangsland: Der Grund der Auswanderung könnte nämlich eine Rolle in der Selbst- und Fremddarstellung der Identität eines interkulturellen Schriftstellers spielen und daher auf einige linguistische Eigenschaften ihrer Sprache wirken. Insbesondere ist die Überprüfung eines eventuellen Übergebrauchs der Wortbildungsmuster auf *-ung* seitens derjenigen Autoren interessant, die als Gastarbeiter nach Deutschland ausgewandert sind: Einerseits könnte die Generalisierung der Wortbildungsstrategien auf *-ung* den Effekt einer sozial bzw. soziolinguistisch bedingten Fossilisation sein (vgl. u.a. Preston 1993), insofern die Gastarbeiter einen niedrigen bzw. stigmatisierten Status in der Empfangsgesellschaft besitzen (vgl. Kap. 1.2 und Kap. 2.2). Anderseits schreiben die Forscher des DaF Institutes in München vor allem der Gastarbeiterliteratur einen kreativen Umgang mit der deutschen Sprache zu bzw. sie stellen ihre Autoren teilweise in Anlehnung daran als literarische Gruppe vor. Die Ergebnisse meiner Analyse veranschaulicht die Grafik in Abbildung 2.

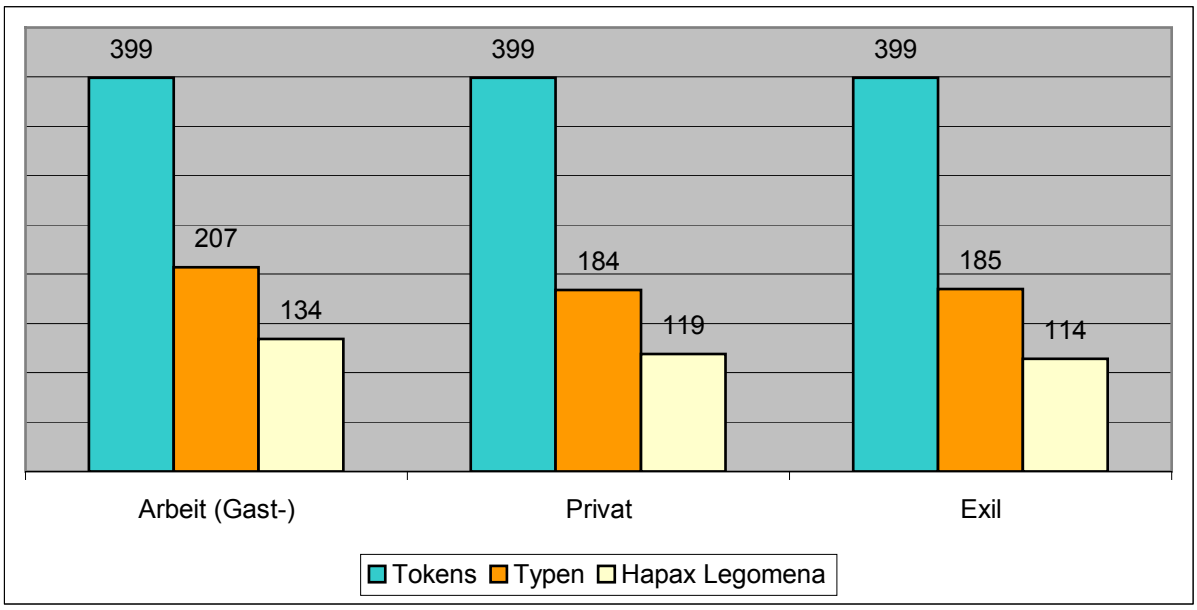


Abbildung 2: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *Auswanderungsgrund.*

Nach einem prop.test ist die Wahrscheinlichkeit, dass die Werte der *realized* und *potential productivity* pro Untergruppe zu derselben Population gehören, größer als 5 %. Daher sind die Variationen im Gebrauch der Wortbildungsmuster auf *-ung* für die Variable Auswanderungsgrund nicht signifikant.

5.2.3.2 Die Variable Muttersprache

Die zweite unabhängige Variable stellt die Muttersprache der Autoren dar, die aus zehn Ausprägungen besteht: Italienisch, Polnisch, Bosnisch, Japanisch, Bulgarisch, Tschechisch, Rumänisch, Ungarisch, Türkisch und Arabisch. Insofern die interkulturellen Schriftsteller auch L2-Lerner sind bzw. Transferphänomene die *Interlanguages* noch in ihren fortgeschrittenen Stadien kennzeichnen, sollte man einen eventuellen Einfluss der Muttersprache auf den Grad der Ausbeutung des Wortbildungssystems des Deutschen überprüfen. Die Ergebnisse meiner Analyse zeigt die Grafik in Abbildung 3.

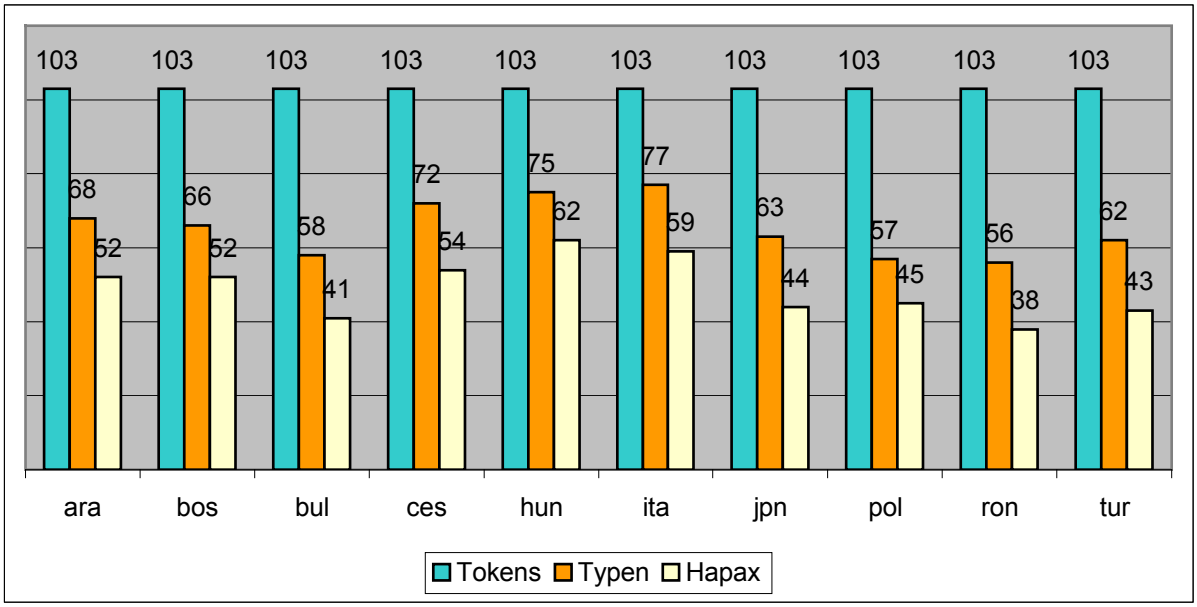


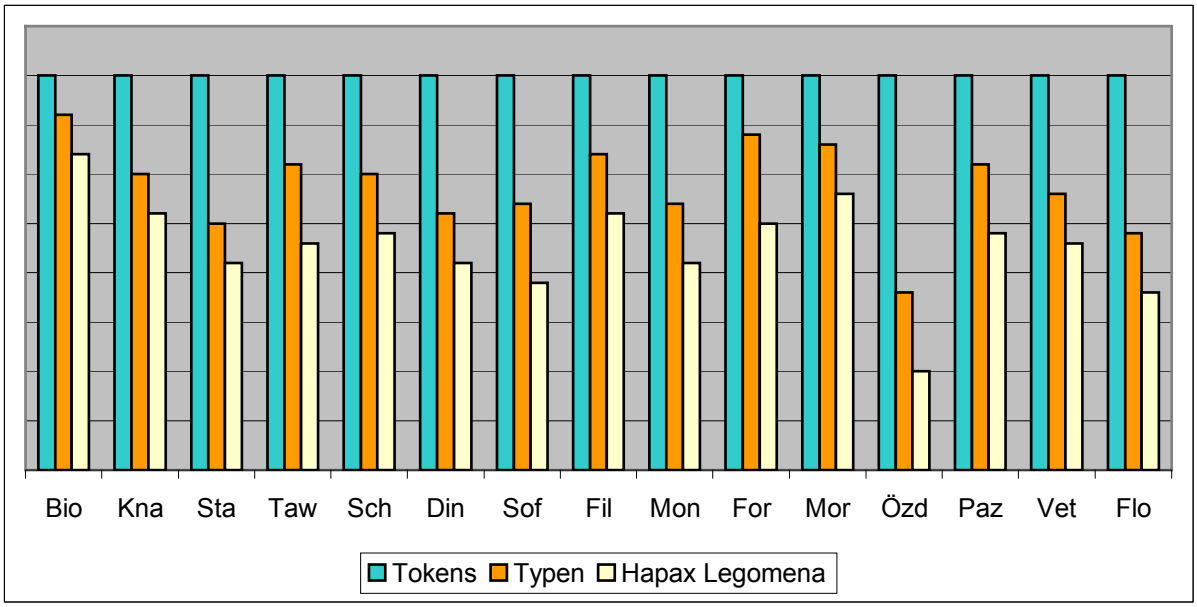
Abbildung 3: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *L1*.

Ein prop.test mit neun Freiheitsgraden stellt diesbezüglich signifikante Unterschiede sowohl für die *realized* als auch die *potential productivity* der morphologischen Kategorie *-ung* für p-Wert = 0,01140 bzw. p-Wert = 0,00785 fest.

Die Bestimmung der L1, die für die signifikanten Ergebnisse verantwortlich sind, benötigt die Durchführung zusätzlicher prop.tests, in denen man jeweils den Wert einer unterschiedlichen L1 nicht in Betracht nimmt. Da ihr Ausschluss den p-Wert in dem prop.test auf über 5 % steigert, ist L1 = Italienisch für die signifikanten Unterschiede in der *realized productivity* zwischen den Subkorpora verantwortlich. Insbesondere handelt es sich bei den italienischen Texten um einen Übergebrauch von neuen Typen auf *-ung* im Verhältnis zu der Anzahl der verwendeten Tokens. Hinsichtlich der *potential productivity* dagegen verursacht die L1 = Ungarisch eine signifikante Variation, und zwar durch einen Übergebrauch.

5.2.3.3 Die Variable Autor

Die dritte unabhängige Variable ist die der *Autor*, anhand der ich das Chamisso-Korpus in 15 Subkorpora von jeweils 10.000 Wort-Tokens aufteile (vgl. Fußnote 31). Anregungen zu einer Analyse nach dieser Variable folgen aus zumindest zwei Betrachtungen der Literatur: Einerseits beschreibt Baayen (1994a) den Produktivitätsgrad eines Wortbildungsmusters als von autorspezifischen Präferenzen stark beeinflusst bzw. verwendet ihn in einer Art stilometrischer Analyse, um Texte nach der Autorschaft zu ordnen. Anderseits kann man die Variable *Individualität* in der SLA-Forschung nach Preston (1993: 119–121) nicht außer Acht lassen, insofern sie einen großen Einfluss auf den L2-Erwerb und die Entwicklung der Lernersprache hat. Die Ergebnisse stellen die Grafik und die Tabelle in Abbildung 4 dar.



	Bio	Kna	Sta	Taw	Sch	Din	Sof	Fil	Mon	For	Mor	Özd	Paz	Vet	Flo
Tokens	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Typen	36	30	25	31	30	26	27	32	27	34	33	18	31	28	24
Hapax Legomena	32	26	21	23	24	21	19	26	21	25	28	10	24	23	18

Abbildung 4: Normalisierte Verteilung der Belege der morphologischen Kategorie *-ung* nach der Variable *Autor*.[39]

Nach einem prop.test ist die Varianz hinsichtlich der *realized* und *potential productivity* pro Autor für p-Wert = 0,001132, bzw. p-Wert = 0,001133 hoch signifikant. Die Verantwortung der

[39] Die Nachnamen der interkulturellen Autoren kürze ich in Abbildung 4 auf die ersten drei Buchstaben ab (für die ganze Name vgl. Anhang 2).

signifikanten Ergebnisse für beide Maße liegt an den Werten der türkischen Schriftstellerin Emine Özdamar, deren Texte von einem Mindergebrauch der Wortbildungsmuster auf *-ung* gekennzeichnet sind. Diesbezüglich ist zu bemerken, dass die Verteilung der Belege bzw. die Verantwortung der signifikanten Ergebnisse für die Variable *Autor* nicht mit den Ergebnissen der Variable *L1* übereinstimmt.

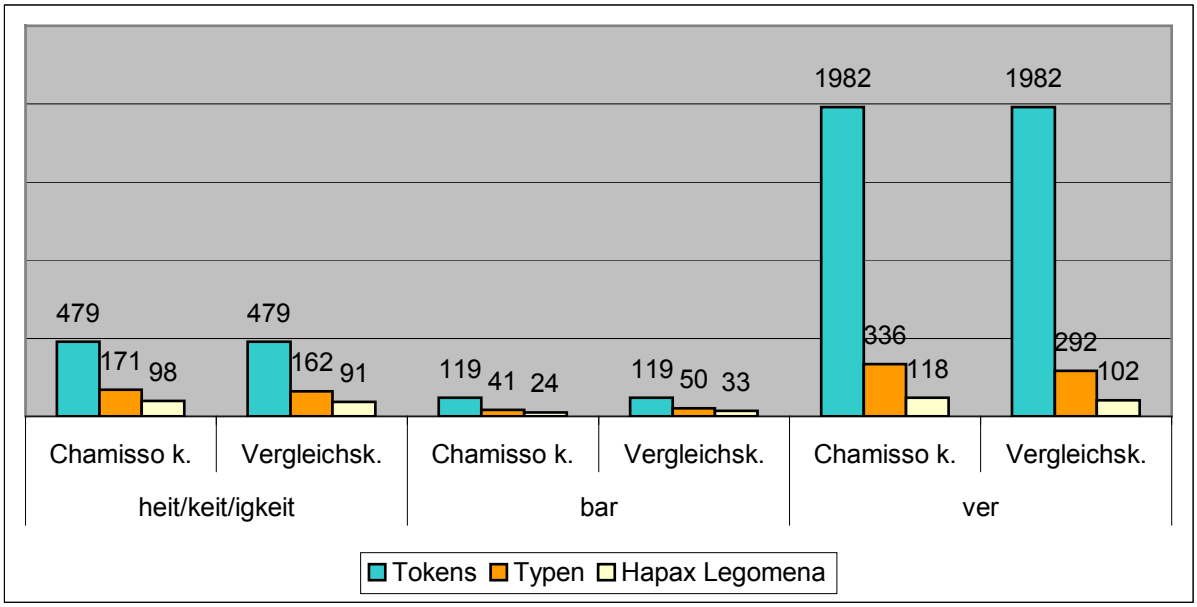
5.2.3.4 Outsider

Zum Schluss errechne ich die *realized* und *potential productivity* der Wortbildungsmuster auf *-ung* im Chamisso- und im Vergleichskorpus nochmals, indem ich jene Texte ausschließe, die jeweils für die signifikanten Ergebnisse nach *L1* und *Autor* zuständig sind. Trotzdem bleiben die Variationen im Gebrauch der morphologischen Kategorie *-ung* zwischen interkulturellen und deutschen Schriftstellern nicht signifikant.

5.2.4 Analytische Analyse III: Gebrauchsvariationen anderer Wortbildungsmuster nach der Variable *Nativeness*

Eine Methode, die den in den Kapiteln 5.2.1 und 5.2.2 vorgestellten recht ähnlich ist, wende ich für die Auffindung der Belege und die Auswertung der Unterschiede im Gebrauch dreier anderer produktiver Wortbildungsmuster des Deutschen im Chamisso- und im Vergleichskorpus an. Hier folgt eine grammatikalische Beschreibung dieser morphologischen Kategorien und die Darstellung meiner Ergebnisse, die ich in Abbildung 5 zusammenfasse (s.u.).
Die erste morphologische Kategorie bildet die der Wortbildungsmuster auf *-keit/-igkeit/-heit*, die man als Suffixvarianten für deadjektivische Nomenbildung beschreibt. Unter anderen geben mir die folgenden Bemerkungen Anlass zur Überprüfung von Variationen nach der Variable *Nativeness* hinsichtlich dieser Wortbildungsmuster: Einerseits betrachtet die Literatur sie als stark produktiv im gegenwärtigen Deutsch (vgl. u.a. Motsch 2004 und Fleischer et al. 2007). Anderseits stellt Motsch (2004: 21) sie explizit dem Wortbildungsmuster auf *-ung* gegenüber. Obwohl beide morphologischen Kategorien durch die Nomenbildung oft die Funktion der syntaktischen Umkategorisierung erfüllen, wirken die durch *-ung* zusammengesetzten Wörter auffälliger als die durch *-keit/-igkeit/-heit*, sodass die zweiten Wortbildungsmuster Motsch zufolge „nicht von Lexikoneinheiten abstrahiert sind, sondern, wie syntaktische Regeln, eine separate Existenz im System der Sprache haben“ (2004: 21 und Kap. 3.1). Daher ist es interessant, zu überprüfen, ob die

interkulturellen Schriftsteller ein unterschiedliches Verhalten bezüglich dieser Wortbildungsmuster im Vergleich zu dem auf *-ung* aufweisen. Die Varianz der *realized* und *potential productivity* der Wortbildungsmuster auf *-keit/-igkeit/-heit* pro Korpus ist aber wie die der auf *-ung* für p-Wert > 0,05 nicht signifikant.



	-keit/-igkeit/-heit		-bar		ver-	
	Chamisso k.	**Vergleichsk.**	**Chamisso k.**	**Vergleichsk.**	**Chamisso k.**	**Vergleichsk.**
Realized productivity	0,357	0,338	0,345	0,42	0,17	0,15
Potential productivity	0,205	0,19	0,202	0,277	0,06	0,051

Abbildung 5: Normalisierte Verteilung der Belege der morphologischen Kategorie *-keit/-igkeit/-heit*, *-bar*, *ver-* nach der Variable *Nativeness* und Berechnung der *realized* und *potential productivity*.

Die zweite Kategorie ist die des deverbalen adjektivalen Affixes *-bar*. Motsch (2004) stellt diesbezüglich zwei semantische Muster für diese morphologische Kategorie fest. Dasjenige, das „die Möglichkeit, Thema eines Geschehens zu sein" ausdrückt (298), gehört „zu den aktivsten der deutschen Gegenwartssprache" (301), wobei die semantischen Bedingungen für die mögliche Basis am niedrigsten sind bzw. Akzeptabilitätsbeschränkungen hauptsächlich nur aus pragmatischen Gründen folgen (vgl. auch Fleischer et al. 2007: 57). Trotz des höheren Produktivitätsgrades des Affixes *-bar* sind seine Variationen nach der Variable *Nativeness* in meiner Korpora für p-Wert > 0,05 nicht signifikant.

Zuletzt betrachte ich die Werte der *realized* und *potential productivity* der deverbalen (und teilweise denominalen) verbalen Wortbildungsmuster auf *ver-* für beide Korpora. Das Präfix *ver-* beschreibt Motsch (2004: 154) als stark polyfunktional, insofern es sechs bis 18 semantische Muster mit sehr

unterschiedlichen Produktivitätsgraden indiziert.[40] Darüber hinaus ist die morphologische Kategorie *ver-* im sprachlichen Input sehr häufig – sie ist nämlich unter den betrachteten Kategorien die mit der größten Tokenhäufigkeit: 1.982 Vorkommen im Chamisso-Korpus –, wobei die Anzahl ihrer unterschiedlichen Typen bzw. Hapax Legomena und daher ihre gegenwärtige Produktivität am niedrigsten ist (vgl. Abbildung 5 und Kap. 5.3). Aus diesen Gründen könnte gerade der Erwerb von *ver-* zusätzliche Schwierigkeiten im Vergleich zu den anderen morphologischen Kategorien aufweisen und ihre Anwendung seitens der L2-Lerner könnte erst recht Generalisierungsprozessen unterliegen. Obwohl die *potential productivity* von *ver-* sich in beiden Korpora nicht signifikant unterscheidet, ist seine *realized productivity* bei den Texten der interkulturellen Schriftsteller signifikant größer als bei den deutschen (p-Wert = 0,03072).

5.3 Diskussion der Ergebnisse

Die Ergebnisse der quantitativen Analyse lehnen die Hypothese eines signifikanten Unterschieds in der Anwendung der produktiven Wortbildungsmuster zwischen interkulturellen und deutschen Schriftstellern bezüglich der morphologischen Kategorien *-ung*, *-keit/-igkeit/-heit* und *-bar* ab (vgl. Kap. 5.2.2 und Kap. 5.2.4). Trotzdem sind in diesem Zusammenhang zwei Anmerkungen relevant.
Erstens ist die anhand des Chamisso-Korpus' festgestellte Rangliste der betrachteten Wortbildungsmuster nach dem Maße der *potential productivity* mit den Angaben in der Literatur durchaus verträglich (vgl. Motsch 2004 und Fleischer et al. 2007). Diesbezüglich gelten nämlich die Affixe *-keit/-igkeit/-heit* und *-bar* mit *potential productivity* 0,205 bzw. 0,202 als die produktivsten und werden als solche auch in Studien über das gegenwärtige Deutsch betrachtet (vgl. Kap. 5.2.4). Der Wert der potenziellen Produktivität des Suffixes *-ung* (0,132) – überprüft anhand des Chamisso-Korpus' – stimmt ferner einigermaßen mit dem überein, den Berth (2009) auf Basis von Texten deutscher Muttersprachler feststellt (0,113). Trotz seiner geringen Größe ist daher mein Korpus bezüglich der Verteilung dieser drei morphologischen Kategorien repräsentativ.
Zweitens verhalten sich die Chamisso-Autoren im Unterschied zu den fortgeschrittenen Lernern des Falko-Korpus', deren Texte nach Berth (2009) nämlich einen Mindergebrauch der Wortbildungsmuster auf *-ung* aufweisen, nicht signifikant anders als die Muttersprachler. Die

[40] Im Folgenden einige Beispiele für die unterschiedlichen semantischen Muster auf *ver-* nach Motsch (2004: 154) zusammen mit den Angaben über ihre Produktivitätsgrade (in der Motsch´schen Terminologie *Aktivitätsgrad*): „Einführung einer Thema-Stelle", *verlachen*, *verspotten*, inaktives Wortbildungsmuster; „etwas falsch tun", *sich versprechen*, schwach aktiv; „etwas vollständig tun", *verstreuen*, *versprühen*, aktiv; „Aktanten nehmen Eigenschaften an", *verarmen*, *sich verdichten*, inaktiv; „verursachen, dass ein Aktant über etwas verfügt", *vermieten*, *verpachten*, aktiv; „Bestandteil von etwas werden", *verwalden*, *verunkrauten*, aktiv; „verursachen, dass zwei Aktanten sich miteinander verbinden", *verschmelzen*, *verwachsen*, inaktiv, usw..

verschiedenen Entstehungssituationen bzw. der „Natürlichkeitsgrad“ meiner gegenüber den Berth'schen Texten erklären diese Ergebnisse teilweise. Insofern aber Frequenzunterschiede im Gebrauch sprachlicher Strukturen nicht unmittelbar als Fehler bewertet und korrigiert werden können, beweisen sie doch auch das Erreichen eines fast targetsprachlichen Niveaus bezüglich der morphologischen Wortbildungskompetenz seitens der interkulturellen Schriftsteller.

Hinsichtlich der Wortbildungsmuster auf *ver-* sind die Schlussfolgerungen zum Teil andere: Einerseits stimmt der Wert ihrer *potential productivity* mit dem in der Literatur überein, sodass die Zuverlässigkeit meiner Ergebnisse gewährleistet ist. Das Präfix *ver-* erweist sich nämlich im Verhältnis zu den anderen als am wenigsten produktiv mit einem potenziellen Produktivitätsgrad von 0,0595. Anderseits ist ihre *realized productivity* bei den interkulturellen Schriftstellern im Vergleich zu den deutschen signifikant größer. Diese Ergebnisse könnten der Effekt einer Generalisierung sein, die wiederum durch die besonders schwierige Erlernbarkeit von *ver-* als polyfunktionalem Affix erklärbar ist. Die Vielfältigkeit der Wortbildungsmuster auf *ver-*, die die Kodierung ihrer Verwendungsregeln und -beschränkungen seitens der L2-Lernern verkompliziert, wäre eine mögliche Erklärung für den festgestellten signifikanten Verwendungsunterschied zu den Muttersprachlern. Einige semantische Muster dieser morphologischen Kategorie sind nämlich im gegenwärtigen Deutschen schwach produktiv, wogegen sie aber eine höhere Vorkommenshäufigkeit aufweisen: Dies könnte einen Übergebrauch erklären, der bei der anderen Affixen nicht vorkommt.

Was die Auswertung der Korrelation zwischen der *realized* und *potential productivity* von *-ung* und den anderen drei unabhängigen Variablen angeht, so ist ein signifikanter Unterschied nach der Variable *Autor* in meinem Korpus belegt. Er war insofern zu erwarten, da die Produktivität einer morphologischen Kategorie mit autorspezifischen Präferenzen stark variiert bzw. von stilistischen Effekten abhängen kann, deren Bestimmung doch eine qualitative Analyse benötigt.

Die signifikanten Unterschiede in den Produktivitätsmaßen von *-ung* nach der Variable *L1* lassen sich dagegen mit den Ergebnissen anderer Untersuchungen schwieriger vereinbaren und gelten darüber hinaus im Rahmen dieser Analyse nicht als schlüssig. Einerseits könnte man sie als vom „morphologischen Transfer“ bedingte Phänomene erfassen, wobei der Begriff selbst umstritten ist. Andererseits bestehen die Subkorpora aus Texten von nur zwei bzw. eines einzigen Autors, sodass die signifikanten Unterschiede auch durch starke autorspezifische Tendenzen verursacht sein könnten.

Indem die Feststellungen der klassischen Kontrastiven Linguistik nach Granger (2007: 13) die Auswertung der Ergebnisse einer *contrastive interlanguage analysis* zwischen nicht-nativen Varietäten positiv ergänzen kann, seien hier doch einige Eigenschaften der L1 im Vergleich zum

Deutschen erwähnt, die sich als verantwortlich für die signifikanten Variationen im Produktivitätsgrad von *-ung* erwiesen haben.

Einerseits ist das Ungarische eine agglutinierende Sprache, die daher grammatikalische Funktionen und syntaktische Beziehungen hauptsächlich durch einzelne monofunktionale Affixe kodiert. Daraus folgt nach Brdar-Szabó (2010), dass deskriptive Erfassungen im morphologischen Bereich in der ungarischen Fachliteratur bzw. Grammatiken Morphem-zentrierte Beschreibungseinheiten verwenden, während sie z.B. im Deutschen Wort-orientiert sind. Insofern die ungarischen DaF-Lerner besondere Schwierigkeiten im Erwerb eines auf unterschiedlichen Bedingungen beruhenden morphologischen Systems aufweisen könnten, kann man eine Asymmetrie in der Anwendung des Wortbildungssystems seitens ungarischer Muttersprachler in meinen Ergebnissen weiter begründen. Dazu kommt, dass eine Morphem-orientierte Perspektive auf die Grammatik den Erwerb einer Fremdsprache teilweise beeinflusst (vgl. Kap. 3.1). Nach der kontrastiven Wortbildungsforschung besteht darüber hinaus eine der wichtigsten Differenzen zwischen den beiden Sprachen gerade darin, „dass im Deutschen die Komposition der produktivste Wortbildungsprozess ist, während im Ungarisch Suffigierung und Präfigierung dominieren“ (Brdar-Szabó 2010: 735).

Die typologischen Unterschiede zwischen dem Deutschen und Italienischen sind im morphologischen Bereich dagegen nicht so auffallend. Auer (2001) sowie Costa (2010) stellen jedoch eine Tendenz zur Komposition des Deutschen einer zur Derivation des Italienischen gegenüber, um den Bedürfnissen der lexikalischen Erweiterung nachzukommen. Diese Differenz könnte Hinweise bieten, um den signifikanten Wert bezüglich der *realized productivity* bei italienischen Muttersprachlern zu interpretieren.

Hier sei jedoch hervorgehoben, dass ähnliche Anmerkungen auch für andere der von mir untersuchten Sprachen gelten könnten. Darunter befindet sich das Türkische, das auch eine agglutinierende Sprache ist, während der Übergebrauch an Komposita ein typisches Merkmal des Deutschen ist, deren Funktion von vielen Sprachen durch andere linguistische Strategien erfüllt wird. Weitere Untersuchungen sind daher in diesem Zusammenhang erforderlich, die mehrere Daten in Betracht ziehen und sich eventuell auf für diese Fragestellungen zusammensetzte Korpora beziehen.

Schließlich ergab sich keine Korrelation zwischen der *realized* und *potential productivity* von *-ung* und der Variable *Auswanderungsgrund*. Die Variable *Auswanderungsgrund* konnte die Status-Unterschiede in meiner Population also entweder nicht wirklich erfassen, oder soziolinguistische Variationen sind im Produktivitätsgrad von *-ung* schlicht nicht vorhanden.

6. Fazit und Ausblick

Der Rahmen der vorliegenden Arbeit bildete die Frage, ob ein Übergebrauch einiger produktiven Wortbildungsmuster die Sprache der interkulturellen Autoren im Vergleich zu deutschen kennzeichnet, womit die „deutschsprachige interkulturelle Literatur“ auf der linguistischen Ebene als solche charakterisiert und gegenüber anderen literarischen Werken differenziert werden könnte. Die literarische Kritik schreibt den Texten der interkulturellen Autoren nämlich einen kreativen Umgang mit der deutschen Sprache zu und setzt diese Besonderheit mit dem Schreiben in einer Fremdsprache in Verbindung. Allerdings basiert die Feststellung der konkreten linguistischen Merkmale, in denen diese Kreativität sich manifestiert, nicht auf eindeutigen Erkenntnissen.

Die Ergebnisse dieser Untersuchung sprechen gegen die Hypothese, dass der kreative Umgang mit der deutschen Sprache bei den interkulturellen Autoren sich durch einen Übergebrauch produktiver Wortbildungsmuster zeigt. Damit ist eine Gruppierung der interkulturellen Schriftsteller auf Grundlage ihrer sprachlichen Kreativität aber dennoch nicht völlig zu widerlegen: Weitere Studien wären in diesem Zusammenhang denkbar, wobei andere linguistische Aspekte in Betracht genommen und ihre Variationen immer in Abhängigkeit von der Variable *Nativeness* untersucht werden könnten. Unter den möglichen Analyseverfahren wäre meiner Ansicht nach ein Vergleich der Kollokationen und der lexikalischen Bündel (*lexical Bundles*) in beiden Korpora interessant. Die interkulturellen Autoren verwenden nämlich häufiger linguistische Einheiten in Kombinationen, welche bei den Muttersprachlern nicht geläufig sind, gegen die zielsprachliche Gebrauchsnorm verstoßen und daher kreativ erscheinen. Für diese Untersuchung wäre jedoch eine strengere Kontrolle der Topik als die in meinem Korpus durchgeführte notwendig.

Aufbauend auf der deskriptiven Analyse dieser Studie wäre darüber hinaus eine qualitative Untersuchung der Derivate auf *-ung* interessant, um die tatsachliche Anzahl von Neubildungen in den Texten der interkulturellen und deutschen Autoren zu ermitteln. Die mittels einer statistischen Methode überprüfte Übereinstimmung in der Verteilung der Typen und Hapax Legomena auf *-ung* in beiden Korpora könnte nämlich anders interpretieren werden, wenn beispielsweise die meisten Hapax Legomena in einem Korpus tatsächliche Neologismen, in dem anderen dagegen seltene, aber doch schon lexikalisierte Wörter darstellen.

Bezüglich der Variationen der Produktivität der Wortbildungsmuster auf *-ung* nach der Variable *Autor* lassen sich im Chamisso-Korpus signifikante Unterschiede identifizieren, welche aller Wahrscheinlichkeit nach individuelle Neigungen widerspiegeln. Des Weiteren könnte auch die Muttersprache einen Einfluss auf die differente Nutzung des Wortbildungssystems des Deutschen als Fremdsprache haben. Für die Bestimmung dieser letzten Korrelation wären jedoch weitere

Untersuchungen erforderlich, die eine größere Mengen an Daten wie auch einen für diese Fragestellung sorgfältiger stratifizierten Korpus zur Verfügung hätten. Die Analyse der Wechselbeziehung eines soziolinguistischen Faktors, wie etwa der *Auswanderungsgrund,* mit der morphologischen Produktivität von *-ung* hat demgegenüber kein signifikantes Resultat generiert. Diesbezüglich könnten mehrere soziolinguistische Faktoren relevant sein und untersucht werden, um die Gruppierung der interkulturellen Autoren auf Basis ihres sozialen Status' detaillierter zu begründen.

Letztlich sind die Ergebnisse des Vergleichs für die morphologische Kategorie *ver-* zu beachten, da sie mit denen der Wortbildungsmuster auf *-ung*, auf *-keit/-igkeit/-heit* und auf *-bar* nicht übereinstimmen. Es lässt sich nämlich ein signifikanter Unterschied in der *realized productivity* von *ver-* zwischen deutschen und interkulturellen Autoren festzustellen. Insbesondere handelt es sich um einen Übergebrauch der Typen auf *ver-* im Chamisso-Korpus. Dieses Resultat könnte im Gegensatz zu den sonstigen Ergebnissen die in dieser Arbeit postulierte Hypothese teilweise stützen. Es würde sie allerdings auf einen Übergebrauch derjenigen morphologischen Kategorien präzisieren, welche als besonders schwierig im Erwerb gelten. Studien bezüglich der Produktivität anderer generalisierten und polyfunktionalen Wortbildungsmuster bei Texten von L2-Lernern sind in diesem Zusammenhang erforderlich, um die Triftigkeit dieser Überlegungen zu bestätigen.

7. Quellenverzeichnis

7.1 Literatur

Ackermann, Irmgard (1997): Der Stellenwert des Sprachwechsels in der „Ausländerliteratur“. In: S. Beer (Hrsg.): *Sprachwechsel. Eine Dokumentation der Tagung „Sprachwechsel – Sprache und Identität“, 22. bis 24. November 1996 im Deutschen Literaturinstitut der Universität Leipzig*. Esslingen: Künstlergilde, S. 16–27.

Adjemian, Christian (1976): On the nature of interlanguage systems. In: *Language Learning*, Jg. 26, H. 2, S. 297–320

Amodeo, Immacolata (2009): *Literatur ohne Grenzen. Interkulturelle Gegenwartsliteratur in Deutschland – Porträts und Positionen*. Sulzbach/Taunus: Helmer.

Androutsopoulos, Jannis und Spreckels, Janet (2010): Varietät und Stil: Zwei Integrationsvorschläge. In: P. Gilles; J. Scharloth; E. Ziegler; K. J. Mattheier (Hrsg.): *Empirische Evidenzen und theoretische Passungen sprachlicher Variation*. Frankfurt am Main: Lang, S. 197–214.

Auer, Peter (2001): Kontrastive Analysen Deutsch-Italienisch: eine Übersicht. In: G. Helbig; L. Götze; G. Henrici; H.-J. Krumm (Hrsg.): *Deutsch Fremdsprache: ein internationales Handbuch*. Berlin: de Gruyter, HSK 19.1, S. 367–374.

Baayen, Rolf Harald (1993): On Frequency, Transparency, and Productivity. In: G. Booij; J. van Marle (Hrsg.): *Yearbook of Morphology 1992*. Dordrecht: Kluwer Academic Publishers, S. 181–208.

Baayen, Rolf Harald (1994a): Derivational Productivity and Text Typology. In: *Journal of Quantitative Linguistics*, Jg. 1, H. 1, S. 16–34.

Baayen, Rolf Harald (1994 b): Productivity in Language Production. In: *Language and Cognitive Processes*, Jg. 9, H.3, S. 447–469.

Baayen, Rolf Harald (2009): Corpus linguistics in morphology. Morphological productivity. In: A. Lüdeling; M. Kytö (Hrsg.): *Corpus Linguistics. An International Handbook*. Berlin: de Gruyter, HSK 29.2, S. 899–919.

Baayen, Rolf Harald (2010): Demythologizing the word frequency effect: A discriminative learning perspective. In: *The Mental Lexikon, Methodological and Analytic Frontiers in Lexical Research (Part I): Special issue of The Mental Lexicon*, Jg. 5, H. 3, S. 436–461. Hier verwendete Quelle: http://www.ualberta.ca/~baayen/publications/BaayenML2011.pdf. Zuletzt geprüft am 10.11.2011.

Baroni, Marco und Bernardini, Silvia (2006): A New Approach to the Study of Translationese: Machine-Learning the Difference between Original and Translated Text. In: *Literary and Linguistic Computing*, Jg. 21, H. 3, S. 259–274.

Bauer, Laurie (2001): *Morphological productivity*. Cambridge: Cambridge Univ. Press.

Berruto Gaetano (1995): *Fondamenti di sociolinguistica*. Roma: Laterza.

Berth, Michael (2009): *Treffungen, Sinkung und Benützung. Korpuslinguistische Untersuchung des Erwerbs von derivationsmorphologischen Wortbildungsregularitäten bei fortgeschrittenen Lernern des Deutschen als Fremdsprache*. Unveröffentlichte Magisterarbeit. Humboldt-Universität zu Berlin.

Biber, Douglas (1993): Representativeness in Corpus Design. In: *Literary and Linguistic Computing*, Jg. 8, H. 4, S. 243–257.

Biondi, Franco und Schami, Rafik (1981): Literatur der Betroffenheit. Bemerkungen zur Gastarbeiterliteratur. In: C. Schaffernicht; S. Atasayar (Hrsg.): *Zu Hause in der Fremde. Ein bundesdeutsches Ausländer-Lesebuch*. Fischerhude: Verlag Atelier im Bauernhaus, S. 124–136.

Brdar-Szabó, Rita (2010): Kontrastive Analyse Ungarisch-Deutsch. In: H.-J. Krumm; C. Fandrych; B. Hufeisen; C. Riemer (Hrsg.): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch.* Berlin: de Gruyter, HKS 35.1, S. 732–737.

Burdorf, Dieter; Fasbender, Christoph; Moennighoff, Burkhard; Schweikle, Günther und Schweikle, Irmgard (Hrsg.) (2007): *Metzler-Lexikon Literatur. Begriffe und Definitionen*. Stuttgart: Metzler, 3. völlig neu bearb. Auflage.

Chiellino, Carmine (Hrsg.) (2000): *Interkulturelle Literatur in Deutschland. Ein Handbuch*. Stuttgart: Metzler.

Chiellino, Carmine (2000a): Interkulturalität und Literaturwissenschaft. In: Chiellino (2000), S. 387–398.

Costa, Marcella (2010): Kontrastive Analyse Italienisch-Deutsch. In: H.-J. Krumm; C. Fandrych; B. Hufeisen; C. Riemer (Hrsg.): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch*. Berlin: de Gruyter, HKS 35.1, Bd. 1, S. 586–592.

Coseriu, Eugenio (1975): System, Norm und Rede. In: U. Petersen (Hrsg. und Übers.): *Sprachtheorie und allgemeine Sprachwissenschaft. 5 Studien / Eugenio Coseriu*. München: Fink, S. 11–101. Original: *Sistema, norma y habla*, Montevideo, 1952.

Dittmar, Norbert (1997): *Grundlagen der Soziolinguistik: ein Arbeitsbuch mit Aufgaben*. Tübingen: Niemeyer.

Dittmar, Norbert (2008a): Stil und Sozialität. (Gruppe, Geschlecht, Alter). In: U. Fix; A. Gardt; J. Knape (Hrsg.): *Rethorik und Stilistik*. Berlin: de Gruyter, HSK 31.2, S. 1245–1270.

Dittmar, Norbert (2008b): Varietäten und Stil. In: U. Fix; A. Gardt; J. Knape (Hrsg.): *Rethorik und Stilistik*. Berlin: de Gruyter, HSK 31.2, S. 1669–1690.

Duden (2003), *Deutsches Universalwörterbuch*. Mannheim: Bibliographisches Institut, 5. überarb. Auflage [CD-ROM].

Esselborn, Karl (1995): Deutschsprachige Literatur von Autoren nichtdeutscher Muttersprache und der Adelbert-von-Chamisso-Preis. In: *Jahrbuch Deutsch als Fremdsprache*, Jg. 21, S. 411–427.

Esselborn, Karl (1997): Eine deutsche Literatur. AutorInnen nicht-deutscher Muttersprache. In: A. Wolff; G. Tütken (Hrsg.): *Gedächtnis und Sprachlernen. Beiträge der 24. Jahrestagung Deutsch als Fremdsprache vom 29. – 31. Mai 1996 an der Universität Göttingen*. Regensburg: Wolff, S. 326–340.

Fleischer, Wolfgang; Barz, Irmhild und Schröder, Marianne (2007): *Wortbildung der deutschen Gegenwartssprache*.Tübingen: Niemeyer.

Gaeta, Livio und Ricca, Davide (2006): Productivity in Italian word formation: A variable-corpus approach. In: *Linguistics*, Jg. 44, H. 1, S. 57–89.

González Álvarez, Elsa (2004): *Interlanguage lexical innovation*. München: LINCOM Europa, S. 1–62.

Granger, Sylviane (2007): A Bird's-eye View of Computer Learner Corpus Research. In: W. Teubert; R. Krishnamurthy (Hrsg.): *Corpus Linguistics: Critical Concepts in Linguistics*. London & New York: Routledge, Bd. 2, S. 44–72.

Granger, Sylviane (2009): Learner corpora. In: A. Lüdeling; M. Kytö (Hrsg.): *Corpus Linguistics. An International Handbook*. Berlin: de Gruyter, HSK 29.1, S. 259–275.

Gries, Stefan (2009): *Statistics for Linguistics with R: A Practical Introduction*. Berlin & New York: de Gruyter.

Handler, Peter (2008): Stilistische Phänomene der Wortbildung. In: U. Fix; A. Gardt; J. Knape (Hrsg.): *Rethorik und Stilistik*. Berlin: de Gruyter, HSK 31.2, S. 1563–1575.

Harnisch, Rüdiger (2004): Morphologie / Morphology. In: U. Ammon; P. Trudgill; K. J. Mattheier; N. Dittmar (Hrsg.): *Sociolinguistics / Soziolinguistik*. Berlin: de Gruyter, HSK 3.1, S. 522–530.

Hay, Jennifer und Baayen, Rolf Harald (2005): Shifting paradigms: gradient structure in morphology. In: *Trends in Cognitive Sciences*, Jg. 9, H. 7, S. 342–348.

Heller, Monica (2005): Language and Identity / Sprache und Identität. In: U. Ammon; P. Trudgill; K. J. Mattheier; N. Dittmar (Hrsg.): *Sociolinguistics / Soziolinguistik*. Berlin: de Gruyter, HSK 3.2, S. 1582–1586.

Holmes, David (1994): Authorship Attribution. In: *Computer and the Humanities*, Jg. 28, H. 2, S. 87–106.

Kallmeyer, Werner (1994): Das Projekt „Kommunikation in der Stadt“. In: W. Kallmeyer (Hrsg.): *Exemplarische Analysen des Sprachverhaltens in Mannheim*. Berlin: de Gruyter, Bd. 4.1 (Kommunikation in der Stadt), S. 1–38.

Kallmeyer, Werner (1995): Zur Darstellung von kommunikativem sozialem Stil in soziolinguistischen Gruppenporträts. In: I. Keim (Hrsg.): *Kommunikative Stilistik einer sozialen Welt „kleiner Leute“ in der Mannheimer Innenstadt*. Berlin: de Gruyter, Bd. 4.3 (Kommunikation in der Stadt), S. 1–25.

Keune, Karen; van Hout, Roeland und Baayen, Rolf Harald (2006): Socio-geographic variation in morphological productivity in spoken Dutsch: a comparison of statistical techniques. In: J.-M. Viprey (Hrsg.): *Proceedings of the 8th international conference on the statistical analysis of textual data*. Besançon: Presses Universitaires de Franche-Comte, Bd. 2, S. 571–581.

Klein, Wolfgang; Perdue, Clive und Carroll, Mary (1992): *Utterance structure. (developing grammars again)*. Amsterdam: Benjamins.

Konzett, Matthias (Hrsg.) (2000): *Encyclopedia of German Literatur*. Chicago: Fitzroy Dearborn Publishers.

Jannedy, Stefanie (2011): *Urban German*. Unveröffentlichter Vortrag. Kolloquium Korpuslinguistik. Humboldt-Universität zu Berlin.

Labov, William (1971): Die Logik des Nonstandard-Englisch. In: W. Klein; D. Wunderlich (Hrsg. und Übers.): *Aspekte der Soziolinguistik*. Frankfurt am Main: Athenäum Verlag, Bd. 1, S. 80–97.

Original: The Logic of Non-Standard English, 1969, *Georgetown Monograph on Languages and Linguistics* 22.

Labov, William (2004): Quantitative Analysis of linguistic variation. In: U. Ammon; P. Trudgill; K. J. Mattheier; N. Dittmar (Hrsg.): *Sociolinguistics / Soziolinguistik*. Berlin: de Gruyter, HSK 3.1, S. 6–21.

Lüdeling, Anke und Evert, Stefan (2003): Linguistic experience and productivity. Corpus Evidence for fine-grained distinctions. In: D. Archer; P. Rayson; A. Wilson; T. McEnery (Hrsg.): *Proceedings of the Corpus Linguistics 2003 Conference*. Lancaster: UCREL, S. 475–483. Hier verwendete Quelle: http://edoc.hu-berlin.de/oa/conferences/revZgYDbQIWDk/PDF/23GpExRHTedus.pdf. Zuletzt geprüft am 10.11.2011.

Lüdeling, Anke und Walter, Maik (2010): Korpuslinguistik. In: H.-J. Krumm; C. Fandrych; B. Hufeisen; C. Riemer (Hrsg.): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch.* Berlin: de Gruyter, HKS 35.1, S. 315–322.

Meibauer, Jörg (2007): *Einführung in die germanistische Linguistik*. Stuttgart: Metzler.

Motsch, Wolfgang (2004): *Deutsche Wortbildung in Grundzügen*. Berlin: de Gruyter.

Motsch, Wolfgang (2011): Grammatische und sprachpsychologische Aspekte der Wortbildung. In: H. Elsen (Hrsg.): *Wortbildung im Deutschen zwischen Sprachsystem und Sprachgebrauch. Perspektiven - Analysen - Anwendungen*. Stuttgart: Ibidem-Verlag, S. 43–71.

Nickel, Gerhard (1998): The role of interlanguage in foreign language teaching. In: *IRAL – International Review of Applied Linguistics in Language Teaching*, Jg. XXXVI, H. 1, S. 1–10.

Perdue, Clive (1984): *Second language acquisition by adult immigrants. A field manual*. Rowley, Mass.: Newbury House.

Photong-Wollmann, Pimonmas (1996): *Literarische Integration in der Migrationsliteratur anhand der Beispiele von Franco Biondis Werken*. Dissertation, Universität Siegen – Chiang Mai, Thailand.

Hier verwendete Quelle: http://www.ub.uni-siegen.de/epub/diss/photong.htm. Zuletzt geprüft am: 10.11.2011.
Plag, Ingo (1999): *Morphological productivity. Structural constraints in English derivation*. Berlin: de Gruyter.

Preston, Dennis Richard (1986): Sociolinguistics and Foreign Language Teaching and Learning. In: G. Nickel; J. Stalker (Hrsg.): *Problems of standardization and linguistic variation in present-day English*. Heidelberg: Groos.

Preston, Dennis Richard (1993): *Sociolinguistics and second language acquisition*. Oxford: Blackwell.

Rainer, Franz (2004): Produktivitätsbeschränkungen. In: G. Booij; C. Lehmann; J. Mugdan; S. Skopeteas; W. Kesselheim (Hrsg.): *Morphologie : ein internationales Handbuch zur Flexion und Wortbildung*. Berlin: de Gruyter, HSK 17.1, S. 877–885.

Rösch, Heidi (2004): Migrationsliteratur als neue Weltliteratur. In: *Sprachkunst. Beiträge zur Literaturwissenschaft. Zeitschrift der Österreichischen Akademie der Wissenschaften*, Jg. 35, H. 1, S. 89–109.

Säily, Tanja (2011): Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. In: *Corpus Linguistics and Linguistic Theory*, Jg. 7, H. 1, S. 119–141.

Schriefers, Helbert (1999): Morphologie und Worterkennung. In: A. D. Friederici (Hrsg.): *Enzyklopaedie der Psychologie. Themenbereich C.III: Sprache*. Göttingen: Hogrefe, Bd. 2, S. 117–153.

Semino, Elena und Short, Mick (2004): *Corpus stylistics. Speech, writing and thought presentation in a corpus of English writing*. London: Routledge, S. 1–18.

Selinker, Larry (1972): Interlanguage. In: *International Review of Applied Linguistics in Language Teaching*, Jg. 10, H. 1–4, S. 209–231.

Stamatatos, Efstathios (2009): A survey of modern autorship attribution methods. In: *Journal of the American society for information science and technology*, Jg. 60, H. 3, S. 538–556.

Swann, Joan (Hrsg.) (2004): *A dictionary of sociolinguistics*. Edinburgh: Edinburgh Univ. Press.

Tarone, Elaine (1994): Interlanguage. In: R. E. Asher; J. M. Y. Simpson (Hrsg.): *The encyclopedia of language and linguistics*. Oxford: Pergamon Press, Bd. 4, S. 1715–1719.

Vogel, Klaus (1990): *Lernersprache. Linguistische und psycholinguistische Grundfragen zu ihrer Erforschung*. Tübingen: Narr.

Walter, Maik und Grommes, Patrick (2008): Die Entdeckung des fortgeschrittenen Lerners in der Varietätenlinguistik. In M. Walter; P. Grommes (Hrsg.): *Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitsprachenerwerbsforschung*. Tübingen: Niemeyer, S. 3–28.

Weigel, Sigrid (1992): Literatur der Fremde – Literatur in der Fremde. In: K. Briegleb; S. Weigel; R. Grimminger (Hrsg.): *Gegenwartsliteratur seit 1968*. München: Dt. Taschenbuch-Verlag, S. 182–229.

Weinrich, Harald (1983): Um eine deutsche Literatur von außen bittend. In: *Merkur*, Jg. 37, H. 8, S. 911–920.

Yano, Hisashi (2000): Migrationsgeschichte. In: Chiellino (2000), S.1–17.

Zeldes, Amir (2009): Seminar „Statistische Methode in der Linguistik“. Unveröffentliche Folien. Sommer Semester 2009. Humboldt-Universität zu Berlin.

Züngler, Jane (1989): Identity and IL Development and Use. In: *Applied Linguistics*, Jg. 10, H. 1, S. 80–96.

7.2 Internet-Link

http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/falko/standardseite

http://www.bosch-stiftung.de/content/language1/html/14169.asp.

http://svsload.com/wd/freeocrnet-vsa/

http://www.sfb632.uni-potsdam.de/d1/annis/

http://www.ids-mannheim.de/kl/projekte/dereko_I/

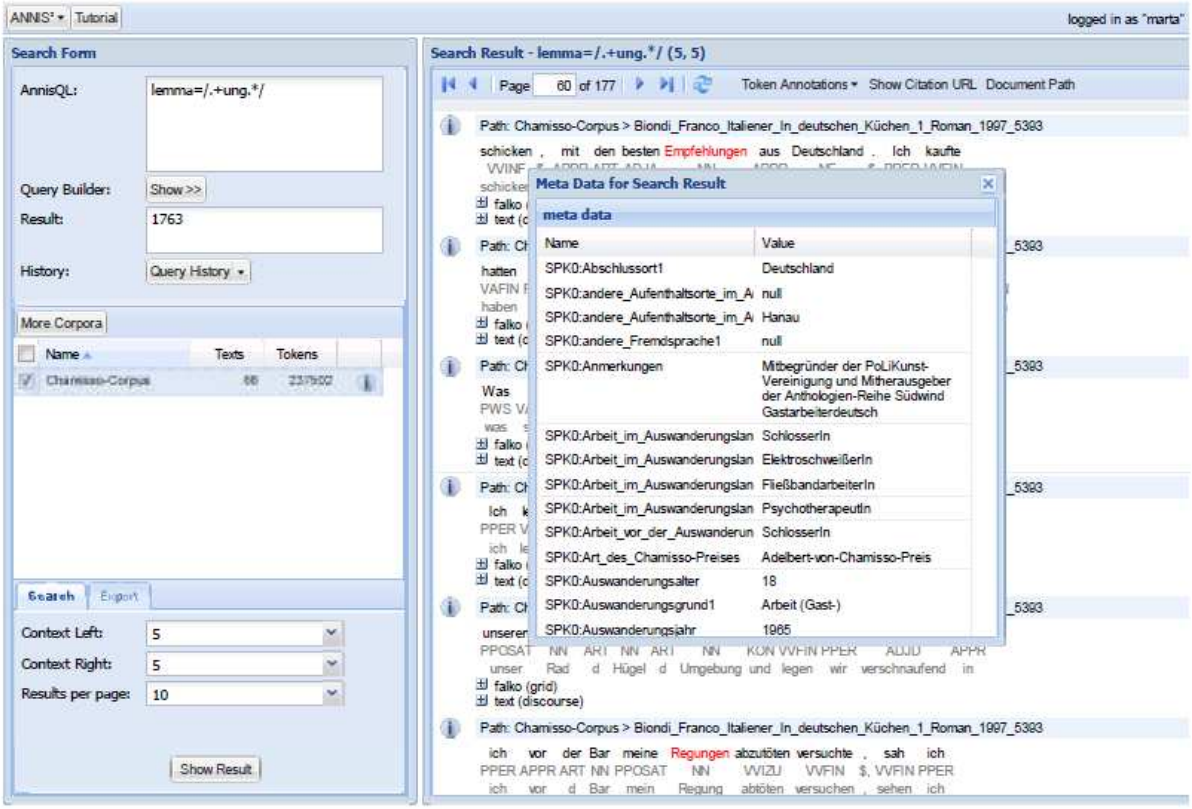
http://stat.ethz.ch/R-manual/R-patched/library/stats/html/prop.test.html

(Alle URLs wurden überprüft am 10.11.2011)

Bibliographische Angaben bezüglich der Texte des Chamisso- und des Vergleichskorpus' befinden sich in Anhang 2 und 3.

8. Anhänge

Anhang 1: Beispiel einer Metadaten-Tabelle in ANNIS2



Anhang 2: Zusammensetzung des Chamisso-Korpus'

Autor	L1	Titel	Wörter	Gattung	Jahr
Biondi Franco	italienisch	In deutschen Küchen	5393	Roman	1997
Biondi Franco	italienisch	In deutschen Küchen	3076	Roman	1997
Biondi Franco	Italienisch	In deutschen Küchen	762	Roman	1997
Biondi Franco	Italienisch	Karussellkinder	769	Roman	2007
Florescu Catalin	Rumänisch	Der blinde Masseur	1019	Roman	2006
Florescu Catalin	Rumänisch	Der kurze Weg nach Hause	3157	Roman	2002
Florescu Catalin	Rumänisch	Wunderzeit	1630	Roman	2001
Florescu Catalin	Rumänisch	Zaira	4194	Roman	2008
Knapp Radek	Polnisch	Franio	5303	Erzählungsband	1994
Knapp Radek	Polnisch	Franio	4103	Erzählungsband	1994
Knapp Radek	Polnisch	Gebrauchsanweisung für Polen	165	Erzählungsband	2005
Knapp Radek	Polnisch	Herrn Kukas Empfehlungen	429	Roman	1999
Monikova Libuse	Tschechisch	Der Taumel	1221	Roman	2000
Monikova Libuse	Tschechisch	Treibeis	4392	Roman	1992
Monikova Libuse	Tschechisch	Treibeis	1230	Roman	1992
Monikova Libuse	Tschechisch	Verklärte Nacht	3157	Roman	1996
Mora Terezia	Ungarisch	Alle Tage	5000	Roman	2004
Mora Terezia	Ungarisch	Der einzige Mann auf dem Kontinent	5000	Roman	2009
Özadamar Emine	Türkisch	Das Leben ist eine Karawanserei	1459	Roman	1992
Özdamar Emine	Türkisch	Der Spiegel im Hof	2973	Erzählungsband	2001
Özdamar Emine	Türkisch	Die Brücke vom Goldenen Horn	1448	Roman	1998
Özdamar Emine	Türkisch	Mutterzunge	2923	Erzählungsband	1990
Özdamar Emine	Türkisch	Seltsame Sterne starren zur Erde	1197	Roman	2003
Schami Rafik	Arabisch	Der ehrliche Lügner	3009	Roman	1992
Schami Rafik	Arabisch	Die Sehnsucht der Schwalbe	6410	Roman	2000
Schami Rafik	Arabisch	Schulz plant seine Entführung	581	Erzählung	2000
Sofronieva Tzveta	Bulgarisch	Berlin Sofia Berlin	3717	Erzählung	2003
Sofronieva Tzveta	Bulgarisch	Briefe einer Dichterin an einen Fahrradhändler	646	Erzählung	2005
Sofronieva Tzveta	Bulgarisch	Diese Stadt kann auch weiß sein	315	Erzählung	1995
Sofronieva Tzveta	Bulgarisch	Reise in die Einsamkeit	1249	Erzählung	1996
Sofronieva Tzveta	Bulgarisch	Trost	4073	Erzählung	2005
Stanisic Sasa	Bosnisch	Wie der Soldat das Grammofon repariert	6183	Roman	2006
Stanisic Sasa	Bosnisch	Wie der Soldat das Grammofon repariert	3817	Roman	2006
Tawada Yoko	Japanisch	Das nackte Auge	10000	Erzählung	2004
Biondi Franco	Italienisch	In deutschen Küchen	4681	Roman	1997
Biondi Franco	Italienisch	In deutschen Küchen	5319	Roman	1997
Dinev Dimitre	Bulgarisch	Engelszungen	4269	Roman	2003
Dinev Dimitre	Bulgarisch	Engelszungen	1331	Roman	2003
Dinev Dimitre	Bulgarisch	Ein Licht über dem Kopf	4112	Erzählungsband	2005
Dinev Dimitré	Bulgarisch	Haut und Himmel	388	Theaterstück	2006
Filip Ota	Tschechisch	Prag ist eine europäische Stadt	1310	Erzählung	2006
Filip Ota	Tschechisch	Der siebente Lebenslauf	8690	Roman	2001
Forgo Leda	Ungarisch	Der Körper meines Bruders	1272	Roman	2007
Forgo Leda	Ungarisch	Der Körper meines Bruders	4841	Roman	2007
Forgo Leda	Ungarisch	Vom Ausbleiben der Schönheit	3887	Roman	2010

Autor	L1	Titel	Wörter	Gattung	Jahr
Schami Rafik	Arabisch	Das Geheimnis des Kalligrafen	3808	Roman	2008
Schami Rafik	Arabisch	Das zukünftige Buch	228	Erzählung	1993
Schami Rafik	Arabisch	Die dunkle Seite der Liebe	4434	Roman	2004
Schami Rafik	Arabisch	Gesammelte Olivenkerne	1530	Erzählungsband	1997
Tawada Yoko	Japanisch	Das nackte Auge	584	Erzählung	2004
Tawada Yoko	Japanisch	Das nackte Auge	6142	Erzählung	2004
Tawada Yoko	Japanisch	Ein Gast	741	Erzählung	1993
Tawada Yoko	Japanisch	Opium für Ovid	308	Erzählungsband	2000
Tawada Yoko	Japanisch	Schwager in Bordeaux	758	Roman	2008
Tawada Yoko	Japanisch	Wie der Wind im Ei	1467	Theaterstück	1997
Stanisic Sasa	Bosnisch	Wie der Soldat das Grammofon repariert	10000	Roman	2006
Knapp Radek	Polnisch	Franio	1252	Erzählungsband	1994
Knapp Radek	Polnisch	Franio	4920	Erzählungsband	1994
Knapp Radek	Polnisch	Franio	3172	Erzählungsband	1994
Knapp Radek	Polnisch	Papiertiger	656	Roman	2003
Pazarkaya Yüksel	Türkisch	Ich und die Rose	10000	Roman	2002
Veteranyi Aglaja	Rumänisch	Das Regal der letzten Atemzüge	1373	Roman	2002
Veteranyi Aglaja	Rumänisch	Das Regal der letzten Atemzüge	5033	Roman	2002
Veteranyi Aglaja	Rumänisch	Das Regal der letzten Atemzüge	574	Roman	2002
Veteranyi Aglaja	Rumänisch	Vom geräumten Meer den gemieteten Socken und Frau Butter	1965	Roman	2004
Veteranyi Aglaja	Rumänisch	Warum das Kind in der Polenta kocht	1055	Roman	1999

Anhang 3: Zusammensetzung des Vergleichskorpus'

Autor	L1	Titel	Wörter	Gattung	Jahr
Didt Sebastian	Deutsch	Dicker Daumen	19353	Kurzgeschichten	2005
Seliger Dirk und Anke	Deutsch	Baron Münchhausens wahre Lügen	14567	Erzählung	2006
Billowie Gudrun	Deutsch	Kasimir sucht das Glück	11199	Erzählung	2006
Büttner Heidi	Deutsch	UnWahre Geschichten	13759	Kurzgeschichten	2005
Blechschmidt Jeannette	Deutsch	Karla	8586	Geschichte	2006
Pawlowski Milan	Deutsch	Meo	15646	Novelle	2005
Schneeweiß Heinrich	Deutsch	Was nun Prometheus?	20568	Erzählung	2004
Klossek Susann	Deutsch	Männer	20777	Kurzgeschichten	2005
Schoof Renate	Deutsch	In ganz naher Ferne	39021	Erzählungen	2003
Ramge Sigrid	Deutsch	Strahlenkinder	30237	Roman	2006
Kammerer Till	Deutsch	Brandnotizen	9892	Kurzgeschichten	2004
Hempel Wolfgang	Deutsch	Verlorene Jahre verlorenes Glück	33896	Roman	2005

Anhang 4: Ausschnitt aus der Beobachtungstabelle

Token	Lemma	Frequenz
Ahnung	Ahnung	1
Andeutungen	Andeutung	1
Aufregung	Aufregung	1
Begegnung	Begegnung	1
Entscheidung	Entscheidung	1
Erfüllung	Erfüllung	1
Erfüllung	Erfüllung	1
Erleichterung	Erleichterung	1
Kreuzung	Kreuzung	1
Morgendämmerung	Dämmerung	1
Nahrung	Nahrung	1
Rettung	Rettung	1
Rührung	Rührung	1
Rührung	Rührung	1
Vergebung	Vergebung	1
Verwirrung	Verwirrung	1
Selbstbedienungsautomaten	Bedienung	1
Zerstörung	Zerstörung	1
Abschiedsvorstellung	Vorstellung	1
Absperrung	Absperrung	1
Anweisungen	Anweisung	1
Bewegung	Bewegung	1
Bewegungen	Bewegung	1
Circusvorstellung	Vorstellung	1
Darbietung	Darbietung	1
Erfahrung	Erfahrung	1
Erinnerung	Erinnerung	1
Erinnerungen	Erinnerung	1
Führung	Führung	1
Handbewegung	Bewegung	1
Handlung	Handlung	1
Hoffnung	Hoffnung	1
Hoffnung	Hoffnung	1
Hoffnung	Hoffnung	1
Messung	Messung	1
Nummerierung	Nummerierung	1
Ordnung	Ordnung	1
Trennung	Trennung	1
Versammlung	Versammlung	1
Vorführung	Vorführung	1
Vorstellung	Vorstellung	1
Voruntersuchungen	Untersuchung	1
Zeitungen	Zeitung	1
Aufklärung	Aufklärung	1
Aufregung	Aufregung	1

Anhang 5: *Test of Equal or Given Proportions* in R

```
> prop.test (c (414, 408), c (1434, 1434), alternative="greater")

        2-sample test for equality of proportions with continuity correction

data:  c(414, 408) out of c(1434, 1434)
X-squared = 0.0426, df = 1, p-value = 0.4182
alternative hypothesis: greater
95 percent confidence interval:
 -0.02428944  1.00000000
sample estimates:
   prop 1    prop 2
0.2887029 0.2845188
```

In diesem Beispiel vergleiche ich die Werte der *realized productivity* der Wortbildungsmuster auf *-ung* nach der Variable *Nativeness*. Der erste Vektor enthält die Anzahl der Typen auf *-ung* pro Korpus (Erfolge), der zweite besteht aus der normalisierte Zahl der Tokens (Versuche). Da meine Hypothese „gerichtet“ ist, ist die Alternative "greater".

9. Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit mit dem Titel

Morphologische Produktivität in deutschen literarischen Texten ausländischer Schriftsteller

Das Beispiel der Ableitungen auf *-ung*

selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Berlin, den 21. November 2011 Marta Lupica Spagnolo