

Anja Eckelt

## Wie viele Muttersprachler braucht die Wirtschaft eines Landes?<sup>\*</sup> Einige statistische Schlaglichter auf das Thema Sprach(en)politik

### Abstract

For this study, 50 countries which pursue 5 different types of policies of official languages were selected. By means of a multiple regression analysis, the percentage of natives speakers of the official languages in these countries are checked for correlations with 8 socioeconomic factors. It can be demonstrated that (1) there is a positive correlation between the percentage of native speakers of the national official language(s) and the reading competence of people over 15 years of age and (2) that there is a negative correlation between the percentage of native speakers of the national official language(s) and the percentage of urban population.

### Sommaire

Pour cette étude, 50 pays qui poursuivent 5 différents types de politique à l'égard des langues officielles ont été sélectionnés. Par une analyse de régression multiple, on cherche des corrélations entre le pourcentage des locuteurs natifs des langues officielles dans ces pays et de 8 facteurs socioéconomiques. Il est possible de montrer (1) qu'il y a une corrélation positive entre le pourcentage des locuteurs natifs des langues officielles d'un pays et la compétence de lire (parmi les gens plus vieux de 15 ans) et (2) qu'il y a une corrélation négative entre le pourcentage des locuteurs natifs des langues officielles d'un pays et le pourcentage de la population urbaine.

### Zusammenfassung

Für diese Studie werden 50 Länder, die 5 verschiedene Typen von Amtssprachenpolitik betreiben, ausgewählt. Mittels einer multiplen Regressionsanalyse wird der Prozentsatz der Muttersprachler der Amtssprache(n) in diesen Ländern auf Korrelationen mit 8 sozioökonomischen Faktoren hin überprüft. Es kann dargelegt werden, dass (1) eine positive Korrelation zwischen dem Prozentsatz der Muttersprachler der nationalen Amtssprache(n) und der Lesekompetenz der Über-15-Jährigen in den jeweiligen Ländern besteht und (2) eine negative Korrelation zwischen dem Prozentsatz der Muttersprachler der nationalen Amtssprache(n) und dem Prozentsatz der Stadtbevölkerung besteht.

## 1. Einleitung

In seiner Zeit als Ministerpräsident des Bundeslandes Baden-Württemberg äußerte der jetzige EU-Kommissar Günther Oettinger in einem Interview für die Reportage "Wer rettet die deutsche Sprache?" des Senders SWR im November 2005: "Englisch wird die Arbeitssprache. Deutsch bleibt die Sprache der Familie und der Freizeit, die Sprache, in der man Privates liest". Vom Verein Deutsche Sprache bekam er dafür den Titel *Sprachpanscher des Jahres* 2006 verliehen. Oettingers sprachpolitischem bzw. sprachsoziologischem Konzept gegenüber stehen unterschiedliche Empfehlungen für eine EU-Sprachregelung, die der Prämisse der Sprachenvielfalt gerecht werden wollen, zugleich aber bezwecken, die Kommunikation

<sup>\*</sup> Dieser Beitrag stellt die Überarbeitung und Ergänzung eines Teils der Arbeit in Eckelt (2010) dar. Für Hinweise und Anregungen zu einer Vorversion dieses Beitrags sowie für technische Unterstützung danke ich Joachim Grzega.

zwischen den EU-Ländern zu vereinfachen (vgl. Grzega 2006: 284). Derzeit ist es offizielle EU-Politik, die derzeitigen Amtssprachen alle gleichberechtigt nebeneinander stehen zu lassen und alle EU-Bürger mindestens zu dreisprachigen Individuen zu bilden. Doch die EU stellt sich auch die Frage nach der Rolle einer Lingua Franca (vgl. Grzega 2009).

Wichtig ist, auf welcher Basis solche Ratschläge gegeben werden. Diese Studie legt dem Thema "Nicht-Muttersprache und Muttersprache als Amtssprache" wirtschaftliche Parameter zugrunde. Denn ohne Frage ist Sprache immer auch in gewisser Weise mit Wirtschaftlichkeit verbunden. Die Kenntnis einer Fremdsprache etwa macht ein Unternehmen oder ein Individuum insofern wirtschaftlicher, als dass sie die Kommunikation mit mehr Verhandlungspartnern erlaubt. Der Zusammenhang von Sprache und Wirtschaft soll jedoch aus einem anderen Blickwinkel betrachtet werden. Hierbei geht es mit Blick auf Oettingers Zitat um den möglichen nationalökonomischen Wert einer Fremdsprache als Verkehrssprache bzw. Amtssprache eines Landes. Mit anderen Worten: Stehen bestimmte sozioökonomische Ergebnisse mit dem Gebrauch von Muttersprache und Fremdsprache in Verbindung. Dabei sollen derzeit bestehende sprachpolitische Regelungen für amtliche oder öffentliche Kontexte betrachtet werden. Die Arbeit führt damit in gewisser Weise eine grundlegende Studie von Coulmas (1992: 23-26) fort, in der nur ein sozioökonomischer Parameter (das Pro-Kopf-Einkommen) mit der Anzahl der in einem Land gesprochenen Sprachen verknüpft wurde. Dabei wurden jedoch alle Sprachen als gleichwertig angesetzt: es wurden etwa nicht Amtssprachen und Nicht-Amtssprachen getrennt und auch nicht die unterschiedlichen Stärken der Sprecherzahlen einbezogen. Ähnlich wurde bei Pool (1972) das Pro-Kopf-Bruttosozialprodukt mit der Größe der größten Muttersprachengemeinschaft (als Indikator für sprachliche Homogenität) in Beziehung gesetzt – ohne Ansehen des Status einer Sprache. Dies birgt unter anderem das Problem der Abgrenzung zwischen Sprache und Dialekt. Da außerdem Amtssprachen – definiert als die Sprache, die *de jure* oder *de facto* offizielle nationale Sprache eines Landes ist – in nationalökonomischer Hinsicht in jedem Staat eine besondere Rolle spielen dürften, ist in dieser Arbeit die Zahl der Muttersprachler der Amtssprachen eines Landes als soziolinguistischer Parameter herangezogen werden.

Der Beitrag gliedert sich wie folgt. Nachdem in einem kurzen Abriss die der Arbeit zugrunde liegende Fragestellung genauer beschrieben wird, soll ein Überblick über Arbeitsmethode und Quellen darlegen, wie bei der Untersuchung vorgegangen wurde. Hierbei wird sowohl die Auswahl der sozioökonomischen Faktoren als auch die der untersuchten Länder begründet. Darauf aufbauend werden die Ergebnisse der Untersuchung präsentiert und aufgrund der erarbeiteten Grundlagen Hypothesen formuliert. Diese Hypothesen werden anhand einer statistischen Untersuchung überprüft. Die Intention dabei ist, herauszuarbeiten, ob anhand sozioökonomischer Faktoren eine Aussage darüber getroffen werden kann, dass eine bestimmte Sprachpolitik für die EU die wirtschaftlich günstigste wäre. Da das Jahr 2011 dem Kampf gegen Armut und sozialem Ausschluss gewidmet ist, sollen diese Aspekte dabei mit integriert werden.

## 2. Amtssprachen und sozioökonomische Faktoren

Zunächst zu klären war die Auswahl der sozioökonomischen Faktoren und ein sinnvolles soziolinguistisches Kriterium, mit dem diese Faktoren zu korrelieren waren. Bezüglich letzterem kann man sich vor Augen halten, dass es zum einen Länder gibt, deren Amtssprache(n) die meisten Einwohner als Muttersprache sprechen, zum anderen aber auch Länder, deren Amtssprache(n) für die Mehrheit der Einwohner Fremdsprache ist. Letzteres

gilt etwa für die kolonialisierten Länder; aber auch im europäischen Raum sind die Muttersprachen mancher Gruppen keine Amtssprachen. Als soziolinguistisches Kriterium, mit dem die sozioökonomischen Faktoren verglichen werden sollten, wurde folglich die Zahl der Muttersprachler der Amtssprache(n) bestimmt. Dabei muss zunächst unbeachtet bleiben, dass die soziopolitische Entstehungsprozesse und die Einstellungen zu diesen Amtssprachen – mit anderen Worten: die Symbolgehalte dieser Sprachen – je nach Land verschieden ausfallen können.

Amtssprache sei – wie gesagt – definiert als die Sprache, die *de jure* oder *de facto* offizielle nationale Sprache eines Landes ist. Dazu galt es, die Länder dieser Welt in unterschiedliche sprachpolitische Typen zu gliedern. Hierbei wurde auf die Kategorisierung in Grzega (2010 & Ms.) zurückgegriffen. Letztlich sollte dabei folgende Frage geklärt werden: Welche sozioökonomischen Größen könnten eine Korrelation mit dem Faktor „Anzahl der Muttersprachler der überregionalen Amtssprache“ aufweisen?

## 2.1. Auswahl der relevanten sozioökonomischen Faktoren

Bevor eine Auswahl an sozioökonomischen Faktoren getroffen werden konnte, musste die Auswahl einer geeigneten Quelle vorgenommen werden. Hierzu wurden verschiedene Webseiten und Institutionen in Betracht gezogen, konnten jedoch aus verschiedenen Gründen ausgeschlossen werden: Die *OECD* beispielsweise liefert zwar zuverlässige Daten, ihr gehören jedoch nicht alle untersuchten Länder an. Deshalb kann sie als Quelle vernachlässigt werden. Die Statistikwebseite der Vereinten Nationen (<http://unstats.un.org/unsd/default.htm>) ist zwar sehr umfangreich, doch gleichzeitig auch sehr unübersichtlich. Zudem bietet sie zu manchen Faktoren veraltete oder unvollständige Daten. Folglich ist es sinnvoll, andere Quellen zu konsultieren. Einen sehr guten Überblick bietet die Webseite *Index Mundi* (<http://www.indexmundi.com>), die Zahlen zu allen Ländern der Erde liefert. Als Quelle wird hier das *CIA World Factbook* mit dem Stand Januar 2009 angegeben. Vergleicht man diese Zahlen indes mit dem *CIA World Factbook* von 2010, fallen zum Teil erhebliche Unterschiede auf. Da das *Factbook* als Quelle angegeben ist, ist es sinnvoller, mit diesem als Basis zu arbeiten. In der Onlineversion sind jedoch die Zahlen nicht immer aktuell vorhanden und es sind nicht zu allen sozioökonomischen Faktoren für alle relevanten Länder Zahlen angegeben. Aus diesem Grund stützt sich die vorliegende Arbeit auf die Printausgabe des *CIA World Factbook 2010*.

Im *CIA World Factbook 2010* sind statistische Daten zu Geographie, Einwohnern, Regierung, Wirtschaft, Kommunikation, Transport und Militär vorhanden. Ursprünglich wurde das Buch von der *CIA* zur geheimen Verwendung durch die US-amerikanische Regierung erarbeitet; seit 1975 aber ist es gedruckt und ist frei zugänglich. Da in dieser Arbeit eine rechnerische Analyse zum Finden einer möglichen Korrelation stattfinden soll, war es sinnvoll, nur Faktoren zu auflisten, die quantitativ erfassbar sind. Deshalb wurden zunächst für jedes Land die Daten von 41 Faktoren, darunter beispielsweise Durchschnittsalter, BIP per capita und die Anzahl von Internetanschlüssen tabellarisch festgehalten. Schließlich war es unerlässlich, eine Auswahl von Faktoren zu treffen. In die Analyse nämlich sollten nur Größen eingehen, mit denen anhand ihrer Eigenschaften eine Korrelation angenommen werden kann. Es konnte nun immer noch aus einer Vielzahl sozioökonomischer Faktoren ausgewählt werden. Letzter Auswahlfaktor war das Ziel, aus den Bereichen „Einwohner“, „Bildung“, „wirtschaftliche Faktoren“ und „Kommunikation“ je mindestens eine Variable und keine komplementären Variablen (wie zum Beispiel Arbeiter in Landwirtschaft und Arbeiter in Industrie) einfließen zu lassen. Unter diesen Prämissen wurden schließlich Migrationsrate

(in 1000 Einwohner), Stadtbevölkerung (in Anteil an der Landesbevölkerung), Lesekompetenz (der Über-15-Jährigen), BIP per capita (in US\$), Arbeiter in Landwirtschaft (in Anteil an der arbeitenden Landesbevölkerung), Einwohner, die unter der Armutsgrenze leben (in Anteil an der Landesbevölkerung), Exporte (in Milliarden US\$) und Internetnutzer (in % der Landesbevölkerung) ausgewählt. In den Tabellen des nächsten Abschnitts sind die Zahlen, wenn nicht explizit anders angemerkt, in diesen Größen verwendet worden. Alle Daten wurden wie gesagt dem *CIA World Factbook* entnommen.

### 2.1.1. Demographie

Hinsichtlich des Bereichs „Demographie“ wird eine Korrelation mit der Migrationsrate untersucht. Diese Bestimmungsgröße gibt den rechnerischen Unterschied zwischen den Menschen, die in ein Land immigrieren und die aus einem Land emigrieren, an. Sie hilft unter anderem, eine Aussage darüber zu treffen, wie „attraktiv“ ein Land gegenüber einem anderen erscheint. Des Weiteren kann eine starke Immigration wirtschaftliche Probleme wie Arbeitslosigkeit nach sich ziehen, während Emigration zu einer Reduktion der Arbeitskräfte führt (CIA 2010: XXV). Besonders problematisch wirkt sich sogenannter *brain drain*, die Emigration eines hohen Anteils von hochqualifizierten Arbeitern aus (vgl. *Gabler Wirtschaftslexikon*, Stichwort „Brain Drain“). Einer Studie zufolge gibt es eine negative Korrelation zwischen der Emigration und der Größe eines Landes; für kleine Länder mit einer Gesamtbevölkerung von unter 2,5 Millionen ist die Emigrationsrate von Hochqualifizierten fast siebenmal höher als bei großen Ländern mit über 25 Millionen Einwohnern (vgl. Ehm 2009: 41). Hier stellt sich die Frage, ob die Sprachpolitik eines Landes mit Bildung und beziehungsweise oder der Migrationsrate korreliert ist und sich beispielsweise eine solche Abwanderung voraussagen ließe.

### 2.1.2. Bildung

Aus dem Bereich Bildung wird die Lesekompetenz der Über-15-Jährigen betrachtet. Im *CIA World Factbook* ist der prozentuale Anteil der Über-15-Jährigen in den jeweiligen Ländern aufgeführt, der über Lesekompetenz verfügt. Eine hohe Analphabetenrate kann in wirtschaftlichen Maßstäben insofern negativ sein, als dass in der technologisch geprägten und sich schnell veränderten heutigen Welt nur durch eigene Lernfähigkeit auch Wettbewerbsfähigkeit gegeben ist (CIA 2010: XXII).

### 2.1.3. Wirtschaftliche Faktoren

Die analysierten wirtschaftlichen Faktoren sind das Bruttoinlandsprodukt per capita, Arbeitskräfte in der Landwirtschaft, Arbeitslosenrate, der Bevölkerungsanteil unter der Armutsgrenze sowie Exporte. Normalerweise wird das Bruttoinlandsprodukt, kurz BIP, als die wichtigste Größe bei der Bestimmung der wirtschaftlichen Leistungskraft eines Landes bezeichnet. Allgemein ist es definiert als „Wert aller Güter, die in einem Land in einem Jahr hergestellt werden“ (Wildmann 2007: 29). Dieses Kriterium bemisst die Güter nicht nach den Herstellungskosten, sondern nach den tatsächlichen Marktkosten. Güter werden so nur durch den Endverbrauch relevant. Wichtig ist also nicht der Herstellungsprozess, sondern der tatsächliche Konsum. Der Faktor errechnet sich nach dem Inlandsprinzip, das besagt, dass alle in einem Land erstellten Güter, egal ob von einem Ausländer oder von einem Inländer gefertigt, in das BIP des jeweiligen Landes eingehen. Nur so werden Länder anhand dieses Wertes untereinander vergleichbar. Die Bestimmungsgröße macht jedoch meist unpräzise Angaben bezüglich der Lebensqualität und des Wohlstandes der unterschiedlichen Länder

(Wildmann 2007: 28ff.). Deshalb wurde in dieser Arbeit das BIP per Kopf, per capita also, gewählt, das „als wesentlicher Leistungsindikator im Vergleich verschiedener Volkswirtschaften“ gilt (Thome/Birkel 2007: 100). Auch dieses Kriterium ist ein Indikator für die Wirtschaftsentwicklung eines Landes während eines bestimmten Zeitraums. Es ist der Quotient aus dem BIP und der Bevölkerungszahl eines Landes und berücksichtigt somit auch die unterschiedlichen Landesgrößen. Selbstverständlich muss, um einen Vergleich erzielen zu können, auch das unterschiedliche Preisniveau der Länder einbezogen werden, was durch die sogenannte Bereinigung der Kaufkraft erzielt wird (Podziech 2008: 54). Um den Grad der wirtschaftlichen Entwicklung eines Landes zu bestimmen, bedient man sich also dieses Faktors.

Die Anzahl der Arbeitskräfte nach Beschäftigung in Landwirtschaft, Industrie und Dienstleistungen wurden ausgewählt, da anzunehmen ist, dass für die Arbeit in Industrie und Dienstleistungen ein höheres Bildungsniveau Voraussetzung ist als für die in der Landwirtschaft. Es ist festzuhalten, dass die Theorie besteht, dass je fortschrittlicher und entwickelter ein Land ist, desto größer und wichtiger der Dienstleistungssektor in diesem Land ist. Dies lässt sich durch die Tatsache begründen, dass die Landwirtschaft für die physische Existenz des Menschen absolut unabdinglich und somit der Grundstock der Wertschöpfung ist. Die mittlere Ebene ist die Industrie, ohne deren Produkte das moderne Leben kaum mehr vorstellbar ist. Dienstleistungen, die dritte Ebene, ist die Ebene der „zwischenmenschlichen Wechselwirkungen“ (Ludwig/Gabriel 2000: 306). Durch den Reichtum in fortgeschrittenen Nationen wurde im Lauf der Geschichte die Anzahl der Arbeiter im landwirtschaftlichen Sektor durch den Einsatz von Maschinen gesenkt. Diese freigesetzten Arbeiter fanden in der Industrie einen neuen Platz. Im Laufe der Zeit wurden die in der Industrie angestellten Arbeiter durch Maschinen ersetzt, was die heutige Größe des Dienstleistungssektors; und parallel dazu das Schrumpfen des Landwirtschaftssektors, erklärt (vgl. Ludwig/Gabriel 2000: 304ff.). In Bezug auf die Korrelation mit Sprachpolitik ist besonders interessant, ob in Ländern mit einer geringen Muttersprachleranzahl der Amtssprache die gleiche Verteilung der Arbeitskräfte in diesen Bereichen festzustellen ist wie in einem Land mit vielen Sprechern der Amtssprache. Die Arbeitslosenrate soll diesen Bereich noch komplettieren.

Der Faktor Bevölkerung unter der Armutsgrenze gibt den prozentualen Anteil der Anwohner der Gesamtbevölkerung an, der unter der Armutsgrenze lebt. Die von der Weltbank festgesetzte Armutsgrenze ist 1 US\$ pro Tag. Selbstverständlich sollte man bei diesem Faktor im Kopf behalten, dass die Gültigkeit des Konzeptes „absoluter Armut“ kritisiert werden kann. Dies bedeutet, dass die Armut einer Person selbstverständlich auch von deren jeweiligen Umfeld, sprich dem Einkommen anderer Mitglieder der Gesellschaft bestimmt wird (vgl. Durth/Körner/Michaelowa 2002: 15).

Exporte werden als US\$-Betrag von Exporten, die auf einer FOB-Basis (*free on board*) ablaufen, verzeichnet. Dies bedeutet, dass bei Schiffstransporten die Güter „frei an Bord“ gehandelt werden und der Haupttransport nicht vom Verkäufer bezahlt wird (vgl. CIA 2010: XXI; ICC Deutschland o.J.).

#### 2.1.4. Kommunikation

Zum Bereich „Kommunikation“ zählen die Nutzer von Festnetztelefonen, Handys, Fernsehern und Internet. In diesem Bereich wird die absolute Anzahl der Einwohner in allen Ländern angegeben, die zum Beispiel Zugang zum Internet haben. Diese Einträge sind als wichtig

einzustufen, da anzunehmen ist, dass durch Kommunikation nicht nur Sprache vermittelt wird, sondern sie auch für diese notwendig ist. Des Weiteren ist zu vermuten, dass durch diverse Fernsehprogramme und Internet Bildung positiv beeinflusst werden kann.

## 2.2. Auswahl der Länder

Wie bereits kurz erwähnt, liegt den hier ausgewählten Ländern eine Studie von Grzega (Ms.; vgl. auch 2010) zu Grunde. In dieser Studie werden die 197 durch die UN anerkannten Nationen in zwölf verschiedene Sprachpolitiktypen unterteilt. Anschließend wird untersucht, wie oft jeder einzelne sprachpolitische Typ einen der weltweit obersten 10 (Top-Ten) Plätze auf 72 sozioökonomischen Skalen belegt. Anhand des Chi-Quadrat-Tests wird eine statistisch relevante Korrelation zwischen eben diesen Bestimmgrößen und einer Sprachpolitik nachgewiesen. Der Test ergibt, dass positive Ergebnisse in den einzelnen sozioökonomischen Faktoren am ehesten bei Ländern mit dem Sprachpolitiktyp C zu erwarten sind, d.h. die Verwendung von 1-2 supraregionalen Amtssprachen und mehreren regionalen Amtssprachen. Dieser Sprachpolitikgruppe sind die folgenden 15 Länder zuzuordnen: Bolivien, China, Deutschland, Guatemala, Indien, Irak, Italien, Kolumbien, Mexiko, Nicaragua, Österreich, Peru, Russland, Spanien und die Vereinigten Staaten (vgl. Grzega 2010). Diese sind in den Analysen in dieser Arbeit vollständig eingegangen.

Die folgende Tabelle stellt exemplarisch die Regelungen zu den europäischen und einigen weiteren Ländern des sprachpolitischen Typs C dar:

Land/ Regelung	Amts- sprache	% Mutter- sprachler	Regional sprachen	Sonstiges
Bolivien	Spanisch	37,9%	36	wichtige Regionalsprachen (gesprochen von 2/3 der Bevölkerung) - Regionalsprache durch Gesetze geschützt und unterstützt, Umsetzung mangelhaft
China	Mandarin- Chinesisch	70,0%	8	- 55 weitere Minderheitensprachen - Ziel: Bevölkerung spricht Standardchinesisch und kann lateinische Schrift <i>pinyin</i> nutzen
Deutschland	Deutsch	91,1%	4	- viele Immigrantensprachen - Deutsch als dominante Sprache
Guatemala	Spanisch	36,7%	4	-23 <i>Maya</i> -Sprachen, nur 4 gesetzlich unterstützt -Umsetzung mangelhaft
Indien	Hindi	15,9%	17	- Zwiespältige Beziehung zwischen Englisch und <i>Hindi</i> - Autochthone Sprache vs. Kommunikationsmedium - Gute Minderheitensprachpolitik
Italien	Italienisch	93,8%	4	- Anerkennung von 12 Minderheiten - Gute gesetzliche Regelungen nur für größere Minderheiten
Mexiko	Spanisch	82,1%	0	- zwar viele Minderheitensprachen, aber wenige Sprecher (8%) - Spanisch als dominierende Sprache
Österreich	Deutsch	90,4%	6	- offizielle Anerkennung und Schutz von 6 Volksgruppen - Deutsch als dominierende Sprache
Spanien	Spanisch	65,0%	3	- offizielle Anerkennung von nur 3 Sprachen - Spanisch als einzige mit Namen in der Verfassung festgehaltene Sprache

Land/ Regelung	Amts- sprache	% Mutter- sprachler	Regional sprachen	Sonstiges
Russland	Russisch	80,8%	21	- Titularnationen dürfen eigene Sprache als kooffizielle Sprache haben - Russisch dominierende Amts- und Kommunikationssprache

*Abb. 1: Sprachpolitik ausgewählter Länder*

Oberflächlich betrachtet würde man wohl zustimmen, dass die fünf Länder, die den geringeren Anteil Muttersprachler der nationalen Amtssprache haben, nämlich Bolivien, China, Guatemala, Indien und Spanien, schlechtere sozioökonomische Ergebnisse verzeichnen als die anderen Länder. Spanien wird zwar als europäisches Land besser abschneiden als Indien, China oder die lateinamerikanischen Länder, im europäischen Vergleich ist aber die Vermutung noch immer aufrecht zu erhalten. Hieraus ergibt sich die Annahme, dass je geringer der Anteil der Muttersprachler der Amtssprache ist, desto schlechter ist das Abschneiden eines Landes bei sozioökonomischen Faktoren einzuschätzen. Diese Vermutung soll nun überprüft werden. Zunächst müssen Hypothesen den genauen Untersuchungsraum festhalten. Anschließend wird mit einer statistischen Analyse der Zusammenhang zwischen einzelnen sozioökonomischen Faktoren mit dem Faktor „Anzahl der Muttersprachler der überregionalen Amtssprache“ untersucht.

### 3. Hypothesen

Folgende Hypothesen sollen überprüft werden:

Hypothese 1: Es existiert eine positive Korrelation zwischen dem Prozentsatz der Muttersprachler der überregionalen Amtssprache eines Landes und den Faktoren Migrationsrate, Lesekompetenz, BIP per capita, Exporte und Internetnutzer.

Hypothese 2: Es existiert eine negative Korrelation zwischen dem Prozentsatz der Muttersprachler der überregionalen Amtssprache eines Landes und den Faktoren Stadtbevölkerung, Arbeiter in Landwirtschaft und Bevölkerung unter der Armutsgrenze.

### 4. Statistische Methoden

#### 4.1. Datensätze

Ziel der statistischen Analyse ist, wie bereits erwähnt, die Darstellung von Zusammenhängen zwischen der Zielgröße „Anteil Muttersprachler der überregionalen Amtssprache“ an der Gesamtbevölkerung (in %) und anhand einer Reihe von sozioökonomischen Faktoren. Da eine Vorstudie (Eckelt 2010) gezeigt hat, dass die Länder der Sprachpolitikategorie C eine zu kleine Datenbasis für eine sinnvolle Analyse mit Bezug auf die genannten Hypothesen darstellen, wurde ein erweiterter Datensatz aus 50 untersuchten Ländern gewählt. Das Ziel war, mit einer größeren Länderauswahl Zufallstreffer auszuschließen. Die Länderauswahl basierte auf der von Grzega (2010; Ms.) vorgenommenen Einteilung der Länder in verschiedene Sprachpolitikgruppen. Die Sprachpolitikgruppe C stellt das Hauptinteresse der vorliegenden Arbeit dar. Daher flossen alle 15 in ihr vertretenen Länder in die neue Analyse ein. In der genannten Studie (Grzega Ms.) wird dargestellt, dass sich bei den gebildeten Sprachpolitikgruppen die bei zufälliger Verteilung erwarteten Werte für Treffer im Kriterium „unter den besten 10 oder schlechtesten 10 Länder der Erde im sozioökonomischen Faktor x“ von den tatsächlichen Treffern unterschiedlich stark abweichen. Von dieser Grundlage ausgehend, die aufzeigt, dass die Erwartungswerte in Sprachpolitikgruppe C größer sind als

die tatsächlichen Treffer, wurde folgende Länderauswahl getroffen: Von den Sprachpolitikgruppen, deren bei zufälliger Verteilung erwarteten Werte geringer waren als die tatsächlichen Treffer wurden die mit den größten Unterschieden ausgewählt: zwei Länder aus Gruppe J, die beiden Länder aus Gruppe A und die Länder aus Sprachpolitikgruppe B. Zuzüglich zu diesen Ländern wurden Länder gesucht, deren bei zufälliger Verteilung erwarteten Werte größer sind als die tatsächlichen Treffer. Da die bisherige Grundlage aus Sprachpolitikgruppe C bestand, bildeten die 15 Länder aus dieser Gruppe den ersten Teil. Sprachpolitikgruppe D weist den signifikantesten Unterschied auf, beinhaltet aber 34 Länder. Deshalb wurden aus ihr 15 Länder ausgewählt. Hierzu wurde die Gruppe alphabetisch geordnet. Die 15 ersten Länder, zu denen auch alle Informationen über alle relevanten sozioökonomischen Faktoren bekannt waren, flossen in die Untersuchung mit ein. In einem gesonderten Anhang stellt eine Liste die ausgewählten Länder, ihre Sprachpolitikgruppe, sowie eine Beschreibung der Sprachpolitikgruppe dar.

Eine besondere Schwierigkeit war es, verlässliche Informationen über Amts- und Regionalsprachen und die Prozentanzahl der Sprecher zu finden. Da als Quelle für sozioökonomische Kennzahlen das *CIA World Factbook 2010* dient, war es zunächst beabsichtigt, auch die Sprecheranzahlen der einzelnen Sprachen aus diesem Buch zu entnehmen (CIA 2010). Für einige Länder fehlten jedoch Daten, für andere war die Liste der Amtssprachen sichtlich unvollständig. Aus diesem Grund kann das *CIA World Factbook* als Quelle für die Sprecheranzahl ausgeschlossen werden. Als nächste Quelle wurde Jacques Leclercs (1999ff.) Webseite *L'aménagement linguistique dans le monde* in Betracht gezogen. Diese Webseite beschreibt alle sprachpolitischen Gegebenheiten aller auf der Welt anerkannten Länder. Zwar sind dort für alle untersuchten Länder Zahlen vorhanden, dennoch war es nötig (nicht zuletzt angesichts Leclercs sehr unterschiedlicher Quellen), diese mit einer zweiten Quelle zu überprüfen. Hierfür wurde Stegner (2006) *TaschenAtlas Völker und Sprachen* des Klett Verlages zur Hand genommen. Der Atlas führt für einige Länder ebenfalls keine Zahlen an, kann jedoch trotzdem sinnvoll genutzt werden: Um zu kontrollieren, wie aussagekräftig die Webseite der Universität Laval ist, wurden die hier angegebenen Kennzahlen für die prozentuale Menge der Muttersprachler der jeweiligen Amtssprache eines Landes anhand des *TaschenAtlas* überprüft. Zwar weichen die Werte nur um durchschnittlich 0,32 Prozentpunkte ab, in manchen Fällen lassen sich aber große Unterschiede feststellen. So gibt es in Indien beispielsweise laut der Universität Laval 50% *Hindi*-Muttersprachler, laut dem *TaschenAtlas* jedoch lediglich 18,13%. Ein Vergleich mit dem oben zu Rate gezogenen *CIA World Factbook* zeigt, dass auch hier andere Zahlen angegeben werden. Für Indien listet dieses nämlich 41% *Hindi*-Muttersprachler auf. Nachdem es nicht nur generell ein Problem darstellte, überhaupt Aussagen über die Anzahl Muttersprachler von verschiedenen Ländern zu finden, sondern auch die vorhandenen Quellen unterschiedliche Angaben machen, war es nötig, nach weiteren Informationen zu forschen. Sowohl auf der Webseite der Universität aus dem Québec als auch im *TaschenAtlas* ist als Quelle die Webseite *Ethnologue: Languages of the World* angegeben, deren Aufzeichnungen auch als Printversion erhältlich sind (Lewis 2009). Nach einer Überprüfung konnte festgestellt werden, dass hier für alle Länder Kennzahlen vorhanden sind. Diese sind zwar zum Teil unterschiedlich zu den zuvor gefundenen Prozentzahlen, zum Beispiel werden hier lediglich 15,8% *Hindi*-Muttersprachler angegeben, dennoch soll diese Webseite als Grundlage für die vorliegende Arbeit dienen. Dies liegt unter anderem darin begründet, dass die beiden oben genannten Werke *Ethnologue* als Quelle anführen und somit als verlässlich charakterisieren. Des Weiteren kann vermutet werden, dass die Webseite aktuellere Zahlen vorzuweisen hat als der 2006 publizierte *TaschenAtlas*. Der *Ethnologue* wird seit 50 Jahren veröffentlicht und soll deshalb auch gegenüber der von Jacques Leclerc entwickelten Webseite der Universität Laval, die erst 1999



eingerrichtet wurde, bevorzugt werden. Die Daten für die Prozentanzahl der Muttersprachler stammen also von Lewis (2009). Eine zusätzliche Prämisse musste jedoch eingeführt werden: Da in den untersuchten Ländern nun zum Teil nicht nur eine nationale Amtssprache gesprochen wird, musste eine Entscheidung getroffen werden, wer in den Berechnungen als „Muttersprachler der überregionalen Amtssprache“ zählt. Es wurde entschieden, dass im Falle mehrerer Amtssprachen alle Muttersprachler aller Amtssprachen in dem Faktor beachtet werden. Dies ist wie folgt zu verstehen: Angenommen unter den 100 Einwohnern eines Landes sprechen 40 Amtssprache A und 20 Amtssprache B, geht dieses Land mit 60% Muttersprachler der Amtssprache in die Berechnungen ein. Außerdem wurde im Sonderfall der Amtssprache „Standard-Arabisch“, das deutlich von nationalen muttersprachlichen Arabischvarietäten abweicht (vgl. Bassiouney 2009), so verfahren, dass hier keine Muttersprachler für diese Sprache angesetzt werden.

## 4.2. Auswertungsmethoden

Generell folgen die statistischen Analysen Sachs (2004). Das Ergebnis der angegebenen statistischen Testverfahren ist jeweils der aus den Daten berechnete empirische Signifikanzwert  $p$  oder der  $p$ -Wert. Je kleiner der  $p$ -Wert ausfällt, desto größer ist die Wahrscheinlichkeit, dass die Nullhypothese „es besteht kein Zusammenhang“ nicht aufrecht erhalten werden kann. Üblicherweise wird ab einer Schwelle von  $p \leq 0,05$  von einem statistisch signifikanten Ergebnis ausgegangen, das heißt dass der ermittelte Zusammenhang nicht zufallsbedingt ist.

Daneben wird in der folgenden Berechnung der Korrelationskoeffizient eine Rolle spielen. Er bewegt sich zwischen  $+1$  und  $-1$ . Je näher der berechnete Wert eines Zusammenhangsmaßes bei  $+1$  oder  $-1$  liegt, desto ausgeprägter ist der Zusammenhang. Diese Größe kann nach Schwarze (1994: 193) folgendermaßen typisiert werden:

Korrelationskoeffizient $r$	Bestimmtheitsmaß $r^2$	Bedeutung der Korrelation
$r = 0$	$r^2 = 0$	kein Zusammenhang $r^2 =$ kein Erklärungsbeitrag
$0 < r \leq 0,4$	$0 < r^2 \leq 0,16$	niedriger oder schwacher Zusammenhang
$0,4 < r < 0,7$	$0,16 < r^2 < 0,49$	mittlerer Zusammenhang
$0,7 \leq r < 1$	$0,49 \leq r^2 < 1$	hoher oder starker Zusammenhang
$r = 1$	$r^2 = 1$	vollständig linearer Zusammenhang $r^2 =$ vollständige Erklärung

Abb. 2: Korrelationstypen

Sämtliche Berechnungen und statistischen Analysen erfolgten mit Microsoft Excel2007 im Betriebssystem WindowsVista und dem Excel-AddIn XLStat, Version 2010 (Fahmy 2010).

## 4.3. Anzahl der Faktoren

Die Empfehlung des Statistikers Backhaus (2005: 113) lautet, dass in einer Regressionsgleichung die Zahl der Datensätze, also der Länder, mindestens doppelt so hoch sein sollte wie die Anzahl der Variablen. Andere Empfehlungen in der Statistikk-literatur, wie beispielsweise Miles/Shelvin (2001: 119), sprechen von 20 Datensätzen pro Variable. Voraussetzung einer multiplen Regressionsanalyse ist außerdem die Unkorreliertheit der Variablen untereinander (Backhaus 2005: 113). Es wurden 50 Länder und 8 sozio-ökonomische Faktoren ausgewählt.

## 5. Statistische Berechnung

Sowohl die Zielgröße als auch die meisten Einflussfaktoren können ebenso bei 50 Datensätzen aufgrund der Schnelltests auf Normalverteilung nicht als normalverteilt gelten, da der Quotient Median dividiert durch Mittelwert nicht zwischen 0,9 und 1,1 liegt und das Dreifache der Standardabweichung größer als der Mittelwert ist (siehe Abb. 3):

Deskriptive Statistik	Anteil Muttersprachler Amtssprache(n)	Migrationsrate (pro 1000)	Anteil Stadtbevölkerung	Lesekompetenz (>15 J.)	BIP per capita (US\$)	Arbeiter (Landwirtschaft)	Anteil unter Armutsgrenze	Exporte (Milliarde US\$)	Internetnutzer
Mittelwert	0,60	0,93	0,59	0,85	16798,76	0,29	0,32	178,71	0,27
Median	0,74	0,00	0,65	0,93	7400,0	0,15	0,25	32,54	0,20
Standardabweichung	0,33	3,92	0,23	0,19	18640,26	0,30	0,21	357,52	0,25
Anzahl	50,00	50,00	50,00	50,00	50,00	50,00	48,00	50,00	50,00
Median / Mittelwert	1,22	0,00	1,11	1,09	0,44	0,50	0,80	0,18	0,74
3s < Mittelwert	0,976	11,753	0,698	0,562	55920,785	0,895	0,628	1072,571	0,749
normalverteilt	Nein	Nein	Nein	Ja	Nein	Nein	Nein	Nein	Nein

Abb. 3: Deskriptive Statistik der Einflussgrößen

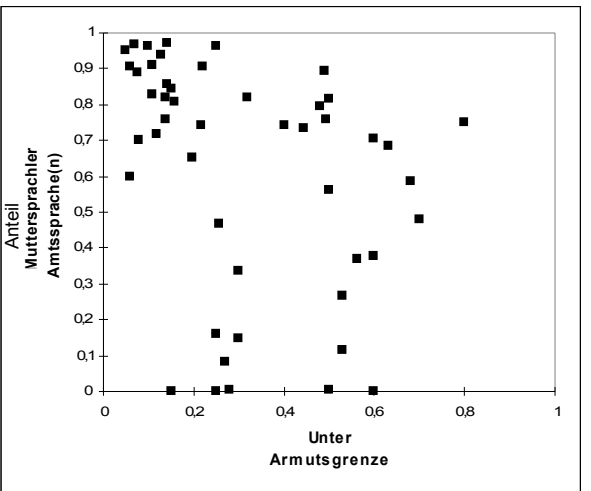
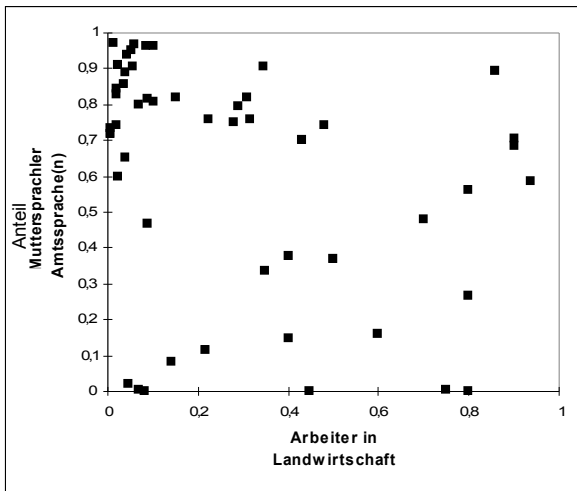
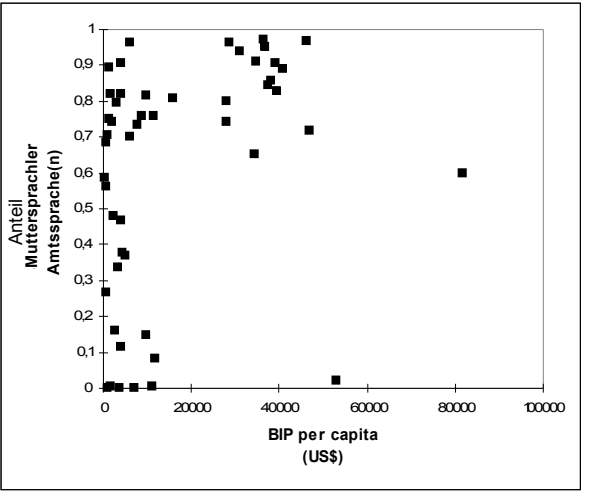
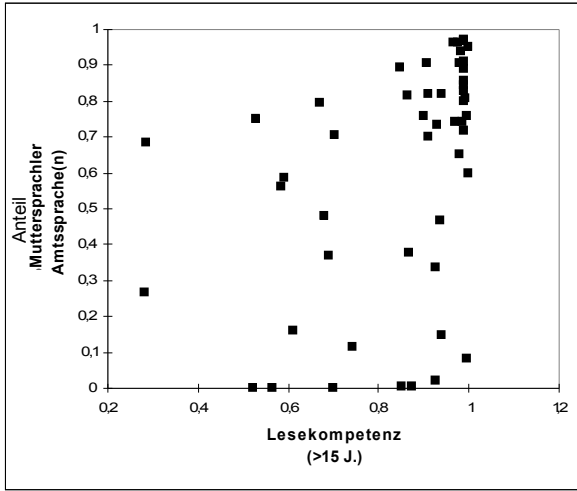
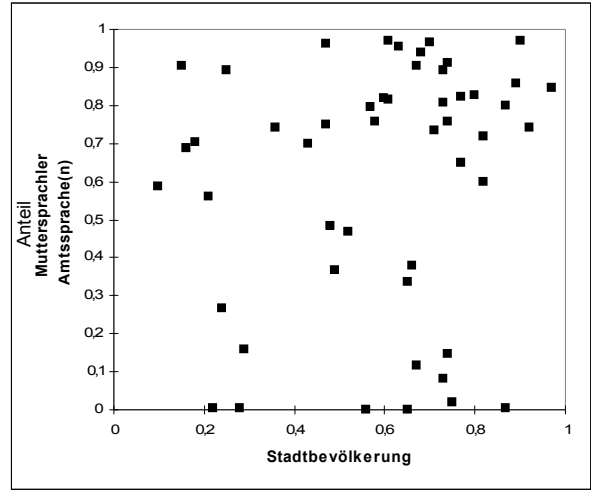
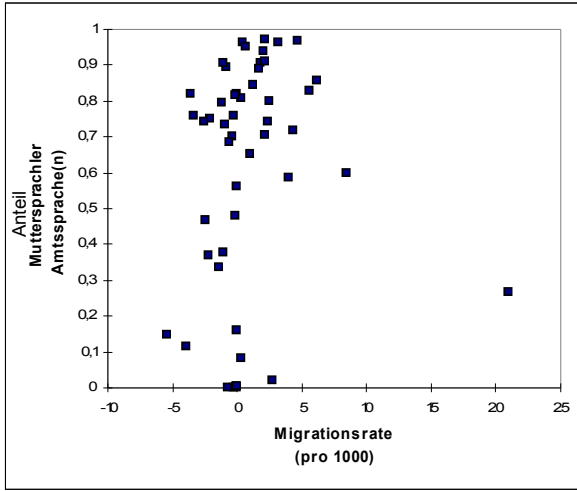
Daher wurde zuerst, wie in der Vorstudie (Eckelt 2010), die Zusammenhangsanalyse mit dem Rangkorrelationskoeffizienten nach Spearman durchgeführt, der für nicht-normalverteilte Daten geeignet ist (vgl. Sachs 2004: 511). Wie die folgende Tabelle zeigt, weisen einige Einflussfaktoren einen positiven, einige einen negativen signifikanten univariaten Zusammenhang mit der Zielgröße auf (fett gedruckte Variablen). Keine signifikante univariate Rangkorrelation wurde für die Variable Stadtbevölkerung ermittelt.

R: Rangkorrelationskoeffizient nach Spearman  
 Zielgröße % Muttersprachler Nationalsprache  
 Werte in Fettdruck sind von Null verschieden mit einem Signifikanzniveau von  $\alpha=0,05$   
 $r^2$  : Bestimmtheitskoeffizienten (Spearman)

Variablen	Korrelationskoeffizient (Spearman):	p-Werte:	Bestimmtheitskoeffizienten (Spearman):
Anteil Muttersprachler Amtssprache(n)	1	0	1
Migrationsrate	0,2828	0,0469	0,0800
Stadtbevölkerung	0,2061	0,1506	0,0425
Lesekompetenz (>15 J.)	0,5269	0,0001	0,2777
BIP per capita (US\$)	0,4069	0,0036	0,1656
Arbeiter in Landwirtsch.	-0,4273	0,0021	0,1826
Bev. unter Armutsgrenze	-0,5026	0,0002	0,2526
Exporte (Milliarde US\$)	0,3739	0,0078	0,1398
Internetnutzer	0,4553	0,0010	0,2073

Abb. 4: Rangkorrelationskoeffizienten

Die hier angefügten x-y-Diagramme, auch Scatterplots genannt, visualisieren die univariaten Zusammenhänge für alle untersuchten Variablen mit der Zielgröße.



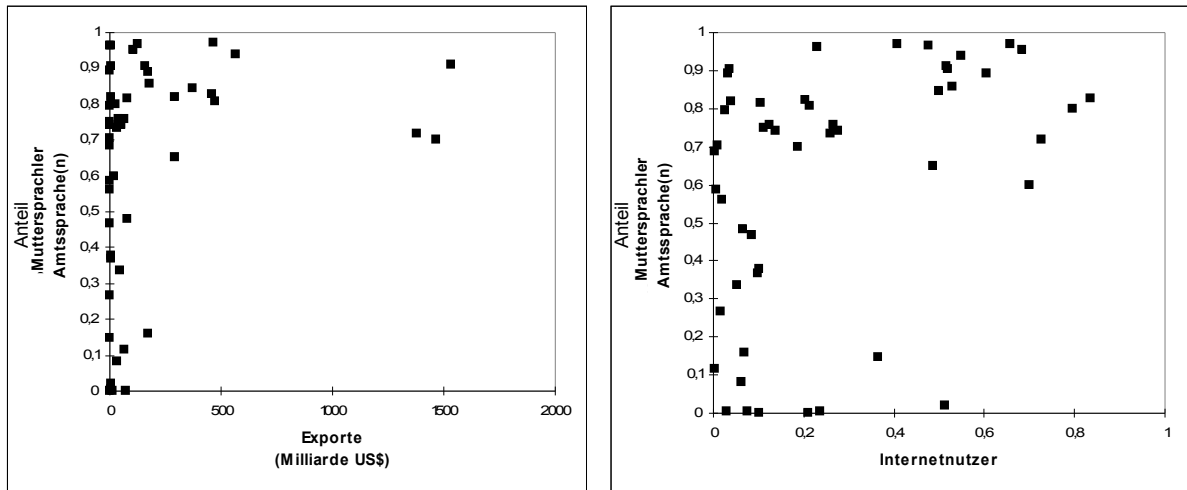


Abb. 5: Scatterplots - Zusammenhänge

### 5.1. Prüfung der Voraussetzungen

Zunächst gilt es festzustellen, ob auch im Datensatz von 50 Ländern (wie im Datensatz aus 10 Ländern in der Vorstudie) eine Multikollinearität der Einflussfaktoren besteht. Folgende Ergebnisse sind zu beobachten: Die Einflussfaktoren sind teilweise untereinander korreliert (vgl. die nachstehenden Tabelle). Jedoch gibt es auch Variablen, die nur wenige signifikante Korrelationen zu anderen Einflussfaktoren aufweisen (z.B. Migrationsrate). Das nächste Ziel ist deshalb, festzustellen, ob diese Zusammenhänge unter den Variablen ähnlich gravierende Einschränkungen der statistischen Aussagekraft mit sich bringen wie im ersten Datensatz, der nur 10 Länder untersuchte.

Variablen	Mutter-sprachler Amtssprache(n)	Migra-tionsrate	Stadt-Bevölke- rung	Lese-Kom- petenz	BIP per capita	Arbeiter (Land-wirt- schaft)	Unter Armuts- grenze	Exporte	Inter- net- nutzer
Mutter-sprachler Amtsspra- che(n)	1	0,0922	0,1601	0,4186	0,3170	-0,3184	-0,3199	0,2750	0,3967
Migra-tionsrate	0,0922	1	0,0029	-0,1585	0,3945	-0,0037	-0,1220	0,1034	0,2947
Stadtbevöl- kerung	0,1601	0,0029	1	0,6550	0,6302	-0,8673	-0,5673	0,2513	0,6856
Lese-kom- petenz	0,4186	-0,1585	0,6550	1	0,5614	-0,7303	-0,6775	0,2894	0,5965
BIP per capita	0,3170	0,3945	0,6302	0,5614	1	-0,6524	-0,6510	0,3072	0,8712
Arbeiter (Landwirt- schaft)	-0,3184	-0,0037	-0,8673	-0,7303	-0,6524	1	0,6868	-0,2791	-0,6806
unter Armuts- grenze	-0,3199	-0,1220	-0,5673	-0,6775	-0,6510	0,6868	1	-0,4376	-0,6657
Exporte	0,2750	0,1034	0,2513	0,2894	0,3072	-0,2791	-0,4376	1	0,3916
Internet- nutzer	0,3967	0,2947	0,6856	0,5965	0,8712	-0,6806	-0,6657	0,3916	1

Abb. 6: Korrelationsmatrix der Einflussfaktoren

Da eine multiple Regressionsanalyse den Zweck hat, die bedeutsamen Variablen von den nicht bedeutenden zu unterscheiden, sollten wie bekannt in einer multiplen Regressionsanalyse möglichst nur voneinander unabhängige Variablen eingehen (Janssen/Laatz 2007: 421 & 454). Ob die Korrelationen der Variablen untereinander toleriert werden können, oder ob die Korrelationen der Einflussfaktoren untereinander zu stark sind, kann, wie oben, mit den Maßen Toleranz und VIF geprüft werden.

Die folgende Tabelle zeigt, dass die Toleranzen bei 50 Datensätzen nicht nahe Null sind. Somit bestehen bei der Analyse mit 50 Ländern und 8 sozioökonomischen Faktoren wenig oder kaum Kollinearitätsprobleme. Dies bedeutet, dass die entsprechenden Variablen alle einen eigenständigen Erklärungswert haben (Norusis 2006: 267). Auch liegen die Werte des Maßes VIF deutlich unter 20-30, das heißt, sie brauchen nicht als problematisch betrachtet werden (Norusis 2006: 267).

Statistik	Migrationsrate (pro 1000)	Stadtbevölkerung	Lesekompetenz (>15 J.)	BIP per capita (US\$)	Arbeiter in Landwirtschaft	unter Armutsgrenze	Exporte (Mrd. US\$)	Internetnutzer
Toleranz	0,6131	0,2186	0,3427	0,1937	0,1773	0,3563	0,7690	0,1926
VIF	1,6310	4,5739	2,9176	5,1614	5,6408	2,8065	1,3003	5,1919

Abb. 7: Toleranzen und VIF der Korrelationen

Eine multiple Regressionsanalyse setzt außerdem voraus, dass parametrische Regressionsanalysen sinnvoll und die Daten weitgehend normalverteilt sind (Sachs 2004: 511; Janssen/Laatz 2007: 420, 450ff.). Diese Richtlinie ist nicht vollständig erfüllt. Die Bedingungen einer multiplen Regressionsanalyse sind weiter auch in Bezug auf die Voraussetzung der Varianzgleichheit verletzt, wie Abbildung 8 zeigt. Trägt man die standardisierten Residuen gegen die Zielgröße auf, so sollte sich kein Muster ergeben; im vorliegenden Fall nehmen jedoch die Residuen mit zunehmender Zielgröße zu, das heißt je höher der Prozentsatz Muttersprachler, desto größer ist auch die Variabilität beziehungsweise die Abweichung, gemessen durch die standardisierten Residuen, von der Regressionsgeraden (Janssen/Laatz 2007: 453). Andererseits ist die Anpassung des Modells gelungen, da bei Auftragung der vorhergesagten Werte (% Muttersprachler) kein Muster mehr zu erkennen ist und die Punkte ohne Muster oder Trend streuen (siehe rechte Grafik).

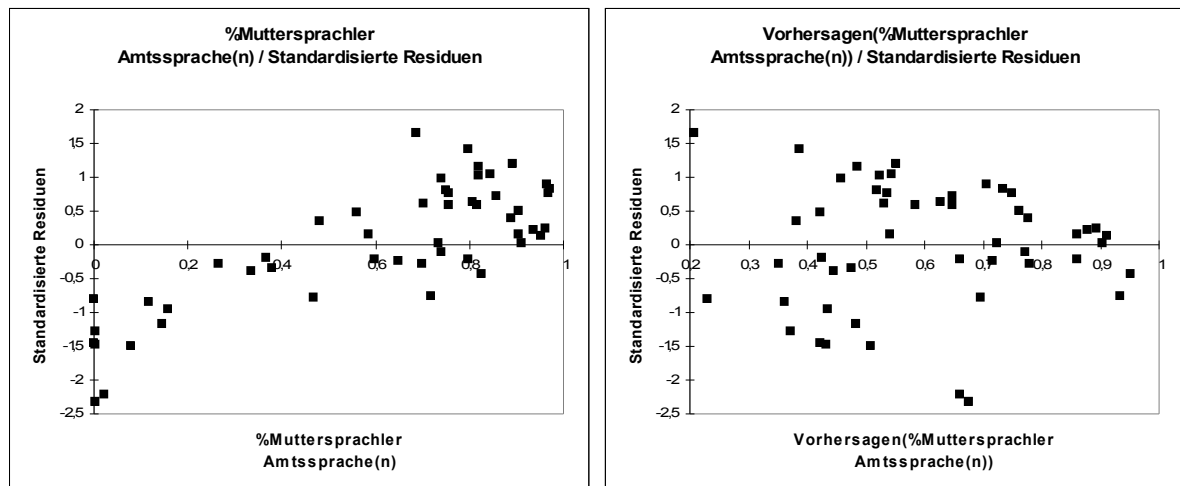


Abb. 8: Vorhersagen

## 5.2. Entwicklung des Modells – Originaldaten

Das Bestimmtheitsmaß ist bei 50 Datensätzen niedrig, das angepasste  $r^2=0,21$ , das bedeutet, dass 21% der Variabilität der y-Achse (Zielgröße) durch die Variablen der x-Achse (Einflussfaktoren) erklärt wird. Das angepasste  $r^2$  berücksichtigt die Tatsache, dass in eine multiple Regressionsanalyse mehrere Variablen eingehen und ist daher das hier richtigerweise zu verwendende Maß (Janssen/Laatz 2007: 427).

	FG	41
Bestimmtheitsmaß	$r^2$	0,3433
Bestimmtheitsmaß	Angepasstes $r^2$	0,2151

Abb. 9: Bestimmtheitsmaß

Auch hier ist eine varianzanalytische Prüfung des Regressionsmodells notwendig, die, wie bereits erwähnt, prüft, ob das Regressionsmodell rein zufällig ist oder ob eine signifikante Anpassung der Daten an die Zielgröße gelingt, das heißt, ob die multiple Regression überhaupt einen signifikanten Zusammenhang ermittelt. Das Regressionsmodell ist sehr viel signifikanter als das erste ( $p=0,0182$ ), was bedeutet, dass eine signifikante Anpassung an die Zielgröße erreicht worden ist. Nach der Interpretation von Janssen und Laatz kann dies dahingehend verstanden werden, dass der Test, ob das Bestimmtheitsmaß  $r^2$  gleich Null ist, deutlich signifikant ausfällt, also dass hier eindeutig ein Unterschied zu Null ermittelt werden kann (Janssen/Latz 2007: 429). Gegenüber 10 Datensätzen ist somit eine deutliche Verbesserung erreicht worden.

Quelle	FG	Summe der Quadratwerte	Mittel der Quadratwerte	F	p-Wert
Modell	8	1,7795	0,2224	2,6787	0,0182
Fehler	41	3,4046	0,0830		
Total	49	5,1841			

Abb. 10: Varianzanalytische Überprüfung

Bei gleichzeitiger Betrachtung mit weiteren Variablen zusammen in der Regressionsanalyse ergeben sich drei signifikante beziehungsweise schwach signifikante Einflussgrößen. Dies sind die Variablen Stadtbevölkerung, Lesekompetenz und Internetnutzer. Die Variable Stadtbevölkerung wird hier als signifikant ermittelt, auch wenn sie bei univariater Betrachtung als nicht signifikant bezeichnet wurde. Dies liegt daran, dass in der multiplen Regressionsanalyse die Zusammenhänge der unabhängigen Variablen untereinander berücksichtigt und diese Korrelationen miteinander bei univariater Analyse nicht in Betracht gezogen werden (Janssen/Laatz 2007: 419).

Quelle	FG	Summe der Quadratwerte	Mittel der Quadratwerte	F	p-Wert	Bemerkung
Migrationsrate (pro 1000)	1	0,0363	0,0363	0,4368	0,5123	nicht signifikant
Stadtbevölkerung	1	0,5094	0,5094	6,1348	0,0175	signifikant
Lesekompetenz (>15 J.)	1	0,3499	0,3499	4,2143	0,0465	signifikant
BIP per capita (US\$)	1	0,0433	0,0433	0,5218	0,4742	nicht signifikant
Arbeiter in Landwirtschaft	1	0,1951	0,1951	2,3497	0,1330	nicht signifikant
Unter Armutsgrenze	1	0,0607	0,0607	0,7306	0,3977	nicht signifikant
Exporte (Milliarde US\$)	1	0,0601	0,0601	0,7236	0,3999	nicht signifikant
Internetnutzer	1	0,2496	0,2496	3,0061	0,0905	schwach signifikant

Abb. 11: Signifikanz der Einflussgrößen

Die obige Tabelle bezieht sich auf die einzelnen Variablen mit ihren jeweiligen Einheiten (Anteil, BIP, Jahre, etc.), während die unten stehende Tabelle die standardisierten Regressionskoeffizienten darstellt. Das heißt, sie zeigt dimensionslose Maße nach Standardisierung der Variablen. Hierdurch wird es möglich, Einflussfaktoren mit unterschiedlichen Einheiten miteinander zu vergleichen, um herauszufinden, welche Größen wichtiger sind. Allgemein gilt: je größer der Wert des standardisierten Regressionskoeffizienten, desto stärkerer Einfluss auf die Zielgröße (Kohn 2004: 146).

Wie aus der folgenden Tabelle hervorgeht, weisen die signifikanten Variablen Lesekompetenz und Internetnutzer annähernd gleich hohe positive Werte des standardisierten Regressionskoeffizienten auf (Spalte „Wert“: zwischen 0,44 und 0,5), und die Variable Stadtbevölkerung weist den größten Regressionskoeffizienten (-0,67) auf, hat also den stärksten Einfluss. Hier liegt ein negativer Zusammenhang vor: je niedriger der Anteil der Muttersprachler der Amtssprache, desto größer der Anteil der Stadtbevölkerung.

Quelle	Wert	Standardfehler	T	p-Wert	Untergrenze (95%)	Obergrenze (95%)
Migrationsrate (pro 1000)	0,1068	0,1616	0,6609	0,5123	-0,2196	0,4333
Stadtbevölkerung	-0,6704	0,2707	-2,4769	0,0175	-1,2171	-0,1238
Lesekompetenz (>15 J.)	0,4438	0,2162	2,0529	0,0465	0,0072	0,8804
BIP per capita (US\$)	-0,2077	0,2875	-0,7224	0,4742	-0,7884	0,3730
Arbeiter in Landwirtschaft	-0,4608	0,3006	-1,5329	0,1330	-1,0678	0,1463
Unter Armutsgrenze	0,1812	0,2120	0,8547	0,3977	-0,2470	0,6094
Exporte (Milliarde US\$)	0,1228	0,1443	0,8506	0,3999	-0,1687	0,4142
Internetnutzer	0,5000	0,2884	1,7338	0,0905	-0,0824	1,0824

Abb. 12: Standardisierte Regressionskoeffizienten der Einflussgrößen

Von den 8 potenziellen Einflussfaktoren verbleiben damit zwei signifikante und ein schwach

signifikanter Einflussfaktor, die mit der Zielgröße in einem statistischen Zusammenhang stehen.

### 5.3. Entwicklung des Modells – Transformierte Daten (Winkeltransformation)

Im Datensatz befinden sich mehrere Variablen, die Prozentwerte darstellen. Prozentwerte sollten bei Vergleichen umgewandelt werden, empfehlenswert ist hierbei die Winkeltransformation. Werden die Variablen mit Prozentzahlen mit der Winkeltransformation geändert, so bessern sich bei einigen Variablen ihre Verteilungseigenschaften, was bedeutet, dass sie nach der Modifizierung als normalverteilt gelten können (grau markierte Felder in Abbildung 13). Damit werden die Ausgangsbedingungen für die multiple Regression besser (Sachs 2004: 355f.). Zwecks einer optimierten wissenschaftlichen Darlegung soll nun anhand der erlangten winkeltransformierten Daten eine erneute Überprüfung stattfinden. Das statistische Vorgehen bleibt unverändert.

Deskriptive Statistik	Migrationsrate	Stadtbevölkerung	Lesekompetenz	BIP per capita	Arbeiter in Landwirtschaft	unter Armutsgrenze	Exporte	Internetnutzer
Mittelwert	0,93	50,52	70,88	16798,76	29,87	31,64	178,71	28,35
Median	0,00	53,73	74,44	7400,00	22,42	30,00	32,54	26,35
Standardabweichung	3,92	14,78	15,06	18640,26	20,93	14,84	357,52	17,74
Anzahl	50,00	50,00	50,00	50,00	50,00	50,00	50,00	50,00
Median / Mittelwert	0,00	1,06	1,05	0,44	0,75	0,95	0,18	0,93
<b>3s &lt; Mittelwert</b>	11,753	44,329	45,188	55920,785	62,792	44,510	1072,571	53,212

Abb. 13: Deskriptive Statistik der winkeltransformierten Einflussgrößen

Die Variablen sind, abgesehen von der Migrationsrate, nach wie vor untereinander korreliert. Die Migrationsrate allerdings ist kaum von den anderen Variablen abhängig. Bei den Faktoren „Bevölkerung unter der Armutsgrenze“ und „Internetnutzer“ ist das Dreifache der Standardabweichung jeweils größer als der Mittelwert. Dies bedeutet, dass hier die Normalverteilung fraglich ist, da nur einer von beiden Schnelltests eine Normalverteilung liefert.



Variablen	Muttersprachlicher Amtssprache(n)	Migrationsrate	Stadtbevölkerung	Lesekompetenz	BIP per capita	Arbeiter (Landwirtschaft)	Unter Armutsgrenze	Exporte	Internetnutzer
Muttersprachlicher Amtssprache(n)	1	0,1047	0,1550	0,4636	0,3146	-0,3156	-0,3297	0,2649	0,3755
Migrationsrate	0,1047	1	0,0129	-0,0540	0,3945	-0,0501	-0,1408	0,1034	0,2490
Stadtbevölkerung	0,1550	0,0129	1	0,6654	0,6282	-0,8628	-0,5520	0,2506	0,7223
Lesekompetenz	0,4636	-0,0540	0,6654	1	0,6529	-0,7522	-0,7316	0,3109	0,7161
BIP per capita	0,3146	0,3945	0,6282	0,6529	1	-0,6938	-0,6787	0,3072	0,8530
Arbeiter (Landwirtschaft)	-0,3156	-0,0501	-0,8628	-0,7522	-0,6938	1	0,6871	-0,3112	-0,7628
Unter Armutsgrenze	-0,3297	-0,1408	-0,5520	-0,7316	-0,6787	0,6871	1	-0,4453	-0,7151
Exporte	0,2649	0,1034	0,2506	0,3109	0,3072	-0,3112	-0,4453	1	0,3955
Internetnutzer	0,3755	0,2490	0,7223	0,7161	0,8530	-0,7628	-0,7151	0,3955	1

Abb. 14: Korrelationsmatrix der winkeltransformierten Einflussgrößen

Das Bestimmtheitsmaß ist niedrig (angepasstes  $r^2 = 0,2068$ ). Das heißt, dass auch dieser Test beweist, dass circa 21% der Variabilität der y-Achse durch die Variablen der x-Achse erklärt wird.

$r^2$	0,3363
Angepasstes $r^2$	0,2068

Abb. 15:  $r^2$

Das Regressionsmodell ist auch bei winkeltransformierten Daten signifikant ( $p=0,0214$ ), was bedeutet, dass eine signifikante Anpassung an die Zielgröße erreicht worden ist. Schlussfolgernd ergibt sich hieraus, dass der Test, ob das Bestimmtheitsmaß  $r^2$  gleich Null ist, also eindeutig ein Unterschied zu Null ermittelt werden kann, sich als deutlich signifikant erweist.

Quelle	FG	Summe der Quadratwerte	Mittel der Quadratwerte	F	Pr > F
Modell	8	9327,5979	1165,9497	2,5973	0,0214
Fehler	41	18405,4630	448,9137		
Total	49	27733,0609			

Abb. 16: Varianzanalytische Überprüfung (winkeltransformierte Variablen)

Der nächste Punkt ist die Überprüfung der Signifikanz der Einflussfaktoren. Bei gleichzeitiger Betrachtung mit weiteren Variablen in der multiplen Regressionsanalyse ergeben sich zwei signifikante Einflussgrößen; und zwar die Variablen Stadtbevölkerung und Lesekompetenz. Die Variable Internetnutzer ist nicht mehr signifikant. Dies sollen die folgenden Tests zeigen:

Quelle	FG	Summe der Quadratwerte	Mittel der Quadratwerte	F	p-Wert
Migrationsrate (pro 1000)	1	303,1774	303,1774	0,6754	0,4159
Stadtbevölkerung	1	2038,0100	2038,0100	4,5399	0,0392
Lesekompetenz	1	2903,4224	2903,4224	6,4677	0,0149
BIP per capita (US\$)	1	150,7535	150,7535	0,3358	0,5654
Arbeiter in Landwirtschaft	1	511,4373	511,4373	1,1393	0,2920
Unter Armutsgrenze	1	382,0398	382,0398	0,8510	0,3617
Exporte	1	328,8282	328,8282	0,7325	0,3970
Internetnutzer	1	498,6451	498,6451	1,1108	0,2981

Abb. 17: Signifikanz der Einflussgrößen (winkeltransformierte Daten)

Nun soll die Ermittlung des standardisierten Regressionskoeffizienten der Einflussfaktoren (winkeltransformierte Daten), parallel zur zuerst durchgeführten Analyse, eine Erklärung über die unterschiedliche Stärke des Zusammenhangs zwischen den erklärenden und den zu erklärenden Variablen geben. Die Daten hierzu finden sich in Abbildung 18.

Quelle	Wert	Standardfehler	t	p-Wert	Untergrenze (95%)	Obergrenze (95%)
Achsenabschnitt	18,4569	43,0608	0,4286	0,6704	-68,5062	105,4200
Migrationsrate (pro 1000)	0,7867	0,9573	0,8218	0,4159	-1,1466	2,7201
Stadtbevölkerung	-0,9103	0,4273	-2,1307	0,0392	-1,7732	-0,0475
Lesekompetenz	0,9395	0,3694	2,5432	0,0149	0,1934	1,6855
BIP per capita	-0,0002	0,0004	-0,5795	0,5654	-0,0009	0,0005
Arbeiter (Landwirtsch.)	-0,3691	0,3458	-1,0674	0,2920	-1,0674	0,3292
Unter Armutsgrenze	0,3637	0,3943	0,9225	0,3617	-0,4325	1,1599
Exporte	0,0083	0,0096	0,8559	0,3970	-0,0112	0,0277
Internetnutzer	0,4180	0,3966	1,0539	0,2981	-0,3830	1,2190

Abb. 18: Standardisierte Regressionskoeffizienten der Einflussfaktoren (winkeltransformierte Daten)

Wie aus der folgenden Tabelle der standardisierten Regressionskoeffizienten hervorgeht (Abbildung 19), weist die signifikante Variable Stadtbevölkerung einen negativen Regressionskoeffizienten (Spalte Wert: -0,56) auf, hat also einen inversen Einfluss. Dies bedeutet, wie oben erwähnt, dass ein negativer Zusammenhang besteht: je größer der Anteil Stadtbevölkerung in einem Land, desto niedriger der Anteil Muttersprachler der Amtssprache. Die signifikante Variable Lesekompetenz (+0,59) weist einen geringfügig höheren, also einen standardisierten Regressionskoeffizienten auf, und zeigt damit einen positiven Zusammenhang an.

Quelle	Wert	Standardfehler	t	Pr >  t	Untergrenze (95%)	Obergrenze (95%)
Migrationsrate	0,1296	0,1576	0,8218	0,4159	-0,1888	0,4479
Stadtbevölkerung	-0,5654	0,2654	-2,1307	0,0392	-1,1013	-0,0295
Lesekompetenz (>15 J.)	0,5948	0,2339	2,5432	0,0149	0,1225	1,0672
BIP per capita (US\$)	-0,1607	0,2773	-0,5795	0,5654	-0,7207	0,3993
Arbeiter in Landwirtschaft	-0,3247	0,3042	-1,0674	0,2920	-0,9391	0,2897
Unter Armutsgrenze	0,2037	0,2208	0,9225	0,3617	-0,2423	0,6497
Exporte (Milliarde US\$)	0,1240	0,1449	0,8559	0,3970	-0,1686	0,4166
Internetnutzer	0,3117	0,2957	1,0539	0,2981	-0,2855	0,9089

Abb. 19: Signifikanzprüfung

Die Voraussetzungen einer multiplen Regressionsanalyse sind nach wie vor in Bezug auf die Varianzgleichheit verletzt, wie die folgende Grafik (Abb. 20) zeigt. Wie bereits erwähnt, sollte sich beim Auftragen der standardisierten Residuen gegen die Zielgröße kein Muster ergeben (Janssen, Laatz 2007: 453). In der vorliegenden Abbildung nehmen jedoch die Residuen mit zunehmender Zielgröße zu. Das bedeutet in diesem Fall: je höher der Prozentsatz der Muttersprachler der Amtssprache, desto größer ist auch die Variabilität beziehungsweise die Abweichung von der Regressionsgeraden. Diese Abweichung wird durch die standardisierten Residuen gemessen. In der Abbildung wird dies links durch einen Trend von links unten nach rechts oben angezeigt. Andererseits ist die Anpassung des Modells gelungen, da bei Auftragung der vorhergesagten Werte (% bzw. Anteil Muttersprachler) kein Muster mehr zu erkennen ist und die Punkte ohne Muster oder Trend streuen. Dies lässt sich in Abbildung 20 rechts ablesen.

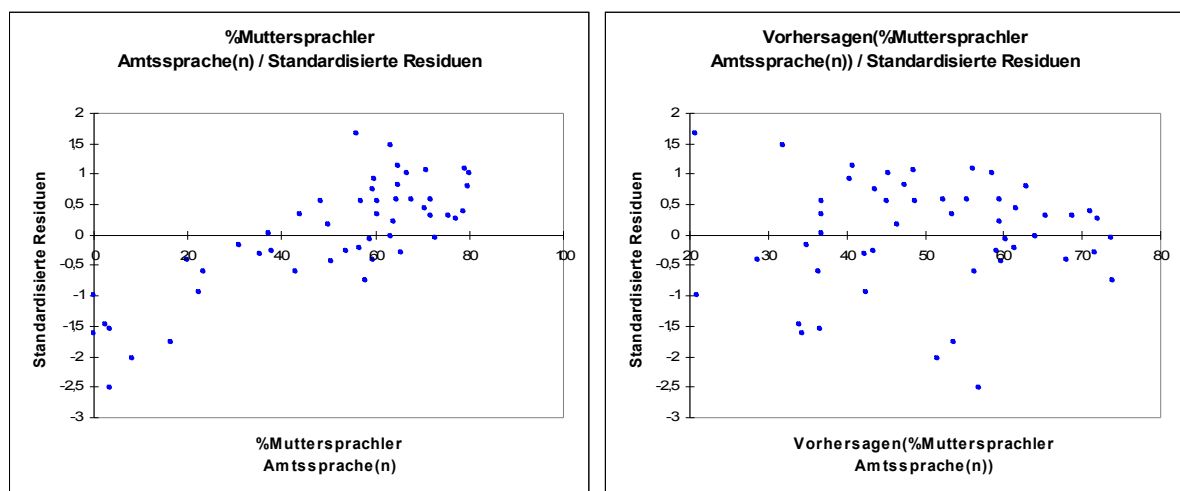


Abbildung 20: Vorhersagen (winkeltransformierte Daten)

#### 5.4. Fazit

Aus diesen Ergebnissen lässt sich folgendes Fazit ziehen: Lesekompetenz und Anteil Stadtbevölkerung konnten bei der multipler Regressionsanalyse sowohl von Rohdaten als

auch von winkeltransformierten Daten als signifikante Einflussfaktoren ermittelt werden, dabei weist der Faktor Stadtbevölkerung einen negativen Regressionskoeffizienten auf und Lesekompetenz einen positiven. Die Einflussvariable Internetnutzer, die bei der Analyse der Rohdaten als schwach signifikant ermittelt wurde, ist nicht stabil, das heißt, sie fällt nach der Winkeltransformation der Daten als nicht signifikant aus und kann nicht mehr als schwach signifikante Variable bezeichnet werden.

Aufgrund der bestehenden Probleme wie der Verletzung der Voraussetzung der Varianzgleichheit und der Normalverteilungsannahme ist der Datensatz nicht als optimal anzusehen; dies schränkt die Verallgemeinerungsfähigkeit der Zusammenhangsanalyse ein. Das bedeutet, dass der Rückschluss von der hier analysierten Stichprobe auf die Gesamtheit aller Länder schwierig ist. Auch die Gültigkeit der Signifikanzprüfungen der Regressionskoeffizienten ist begrenzt. Interpretationen, insbesondere Verallgemeinerungen, sind aber dennoch zulässig, da das Regressionsmodell signifikant ist. Sie werden aber, unter Beobachtung obiger Einschränkungen, vorsichtig durchgeführt (Janssen, Laatz 2007: 419 ff.).

## 6. Interpretation

Es waren folgende Hypothesen aufgestellt:

Hypothese 1: Es existiert eine positive Korrelation zwischen dem Prozentsatz der Muttersprachler der überregionalen Amtssprache eines Landes und den Faktoren Migrationsrate, Lesekompetenz, BIP per capita, Exporte und Internetnutzer.

Hypothese 2: Es existiert eine negative Korrelation zwischen dem Prozentsatz der Muttersprachler der überregionalen Amtssprache eines Landes und den Faktoren Stadtbevölkerung, Arbeiter in Landwirtschaft und Bevölkerung unter der Armutsgrenze.

Anhand der statistischen Berechnung wurden beide Hypothesen mit dem vorhandenen Datensatz zumindest teilweise unterstützt; außerdem wurden Teile beider Hypothesen widerlegt. Es konnte dargelegt werden, dass eine positive Korrelation zwischen dem Prozentsatz der Muttersprachler der nationalen Amtssprache(n) eines Landes und der Lesekompetenz der Über-15-Jährigen in den jeweiligen Ländern besteht. Außerdem existiert eine negative Korrelation zwischen dem Anteil Muttersprachler und dem Anteil der Stadtbevölkerung eines Landes. Wie kann dieses Ergebnis im Bezug auf die europäischen Länder des sprachpolitischen Typs C gedeutet werden? Die folgende Tabelle (Abb. 21) zeigt die beiden Faktoren noch einmal für die europäischen Länder und einige andere Länder des Typs C auf.

	Öster- reich	Deusch- land	Italien	Mexiko	Russ- land	China	Spanien	Boli- vien	Guate- mala	Indien
Mutter- sprachler Amtsspra- chen	90,4%	91,1%	93,8%	82,1%	80,8%	70,0%	65,0%	37,9%	36,7%	15,9%
Lese- kompetenz (>15 J.)	98,0%	99,0%	98,4%	91,0%	99,0%	90,9%	97,9%	86,7%	69,1%	61,0%

Abb. 21: Muttersprachler in der Amtssprache und Lesekompetenz

Die Regressionsanalyse hatte Folgendes ergeben: Je geringer der Anteil der Muttersprachler in den Amtssprachen, desto geringer die Lesekompetenz. Wie man an der Abbildung sieht, kommt diese Wirkung aber erst bei einem sehr niedrigen Anteil der Muttersprachler zum

Tragen. Es wäre falsch, den Schluss zu ziehen, dass (autochthone oder allochthone) Minderheiten sofort zu einem Abfall der allgemeinen Lesekompetenz führten. Es wäre auch falsch, den Schluss zu ziehen, dass eine zu große Zahl von Minderheiten automatisch zu einem Abfall der allgemeinen Lesekompetenz führt – dies ist nur dann der Fall, wenn man nicht dafür Sorge trägt, dass Mitglieder von Minderheiten echt bilingual werden – das betrifft vor allem die Erziehung von Kindern, für die Amtssprache wie eine Muttersprache werden sollte. Ohnehin wäre es falsch, aus sozioökonomischen Korrelationen Schlüsse für eine Sprachpolitik zu ziehen, ohne sozialpsychologische Parameter (mit Blick auf den Symbolcharakter von Sprachen) zu berücksichtigen zu haben.

Mindestens ebenso interessant wie dieses Ergebnis ist, dass keine positive oder negative Korrelation zwischen der Anzahl der Muttersprachler und den Faktoren Migrationsrate, BIP per capita, Arbeitern in Landwirtschaft, Exporte, Internetnutzer oder Bevölkerung unter der Armutsgrenze bewiesen werden konnte. Mit anderen Worten: wenn viele Mitglieder einer Gesellschaft eine Amtssprache nicht als Muttersprache oder wie eine Muttersprache spricht, kann man daraus nicht schließen, dass die Mitglieder oder die Gesellschaft insgesamt wahrscheinlich arm sei. Dies sollte auch bei den immer wiederkehrenden Diskussionen über Wohl und Weh von Migration und Integration berücksichtigt werden.

## 7. Schlussbemerkungen

Die zu klärende Frage dieser Studie war, ob ein Zusammenhang mit einigen sozioökonomischen Faktoren und der Anzahl der Muttersprachler der Amtssprache vermutet werden kann. Das Ergebnis der Untersuchungen ist, dass mit zwei Faktoren durchaus eine Korrelation existiert. Mit Blick auf die Lesekompetenz (und damit die viel beachteten PISA-Ergebnisse) könnte man aus den Ergebnissen etwa folgern, dass es für Länder, in denen die Amtssprachenmuttersprachler die Minderheit bilden, ratsam ist, Programme zur Förderung echter Bilingualität zu entwickeln.

Für die Mehrheit der Faktoren (z.B. BIP per capita, Armutsrate) kann jedoch kein Zusammenhang mit dem Anteil der Muttersprachler der Amtssprachen nachgewiesen werden. Auch dies ist für die Diskussion um sprachpolitische Modelle interessant, da diese Erkenntnis bei manchen Befürwortern einer "Englisch-Only"-Konzeption vielleicht allzu schiefe ökonomische Argumente gerade rücken kann. Die Frage nach der "besten" Lösung bleibt mit dieser Studie gewissermaßen offen.

Außerdem können etliche weitere Erkenntnisse aus dieser Arbeit für die zukünftige Forschung dienlich sein. Folgende Ergebnisse nämlich sind unerlässlich für die Weiterarbeit auf diesem Gebiet: Es ist wichtig, verlässliche Quellen zu finden. Schwierig erweist sich hier besonders die Tatsache, dass zwischen Autoren über die Sprecheranzahl in unterschiedlichen Sprachen Uneinigkeit herrscht. Die in dieser Arbeit zu Rate gezogenen Werke wiesen zum Teil große Differenzen in ihren Angaben auf. Außerdem zeigte die Darstellung der ausgewählten Länder, wie unterschiedlich doch vor allem in verschiedenen geographischen Gebieten der Erde die Entwicklung der Sprachpolitik zu sehen ist. Viele dieser Abweichungen hängen mit der Kolonialisierung und dem Zeitraum, in dem eine Sprache schon in einem Land beheimatet ist, zusammen. In Lateinamerika beispielsweise herrschen völlig andere Voraussetzungen als in Europa. Deshalb empfiehlt sich für die weitere Forschung eine Konzentration auf einzelne Kontinente oder eine anderweitige Berücksichtigung der Unterschiede. Des Weiteren können statistisch relevante Aussagen nur bei einem relativ großen Datensatz getroffen werden. Die an dieser Stelle durchgeführten Berechnungen konnten veranschaulichen, dass vor allem auf

dem Gebiet „Bildung“ ein Zusammenhang zu vermuten ist, der weiter erforscht werden sollte. Auch wenn, zumindest bisher, keine Korrelation mit den hier aufgezeigten Faktoren aus dem Wirtschaftsbereich bewiesen werden konnte, ist dies eine sehr wichtige Erkenntnis. Denn es ist folgende Kette zu vermuten: Je höher die Alphabetisierungsrate, desto höher die Schulabgänger, desto höher die Universitätsabgänger, desto besser die technologische und wirtschaftliche Entwicklung eines Landes, desto größer der wirtschaftliche Erfolg.

Was bedeutet dies nun im Bezug auf die Europäische Union? Es muss natürlich bedacht werden, dass, um aber eine endgültige Aussage treffen zu können, die Untersuchung aller Länder der Erde, eventuell mit einer Konzentration auf die Faktoren aus dem Bereich „Bildung“ zu empfehlen ist. Dies ist damit zu begründen, dass die Aussage aufgrund von 50 Ländern nur vorsichtig getroffen werden kann. Trotzdem scheinen folgende Anmerkungen zur Frage, nach dem Nutzen einer Lingua Franca für die Verwaltung der EU-Länder erlaubt: mit Vorsicht zu genießen sind Konzepte zur Einführung einer alleinigen überregionalen EU-Amtssprache bzw. Vorstellungen, dass die Muttersprache kurz- oder mittelfristig nur noch für den privaten Bereich gut seien – wie dies etwa EU-Kommissar Günther Oettinger in einem Interview mit dem SWR im November 2005 äußerte. Denn in der Tat: die etwaige Einführung von Esperanto oder Englisch als verpflichtender Sprache in Bildung, Verwaltung, Justiz und Medien könnte mit einem Rückgang der Alphabetisierungsrate einhergehen. Die vorliegende Arbeit konnte zeigen, dass, je geringer der Anteil der Muttersprachler in der Amtssprache ist, desto geringer auch die Lesekompetenz der Über-15-Jährigen ist. Würde in der Grundschule europaweit etwa allein auf Englisch unterrichtet werden, stünde folgendes Szenario zu befürchten: Die Kinder würden langsamere Fortschritte, machen, die Eltern könnten diese nicht immer bei den Hausaufgaben unterstützen und nicht alle Lehrer würden die notwendigen Kompetenzen aufweisen. Vor allem in ländlichen Gegenden, in denen das Leben zum größten Teil in der Muttersprache stattfindet, wäre ein Unterricht in einer (fremden) Lingua franca von Schüler- und Lehrerseite als kritisch anzusehen. Dies muss auch im Zusammenhang mit der Tatsache gesehen werden, dass der Anteil der Muttersprachler der Amtssprache negativ mit dem Anteil der Stadtbevölkerung korreliert. Dies hängt vermutlich mit der Tatsache zusammen, dass in Städten die Akzeptanz der (neuen) Amtssprache durch eine größere internationale Bevölkerung schneller voranschreitet. Vermutlich würde vor allem die Anfangsphase nach der Einführung einer EU-weiten Amtssprache große Probleme bringen. Nicht zu vergessen ist auch die emotionale Verbindung mit der regionalen Muttersprache, wegen der eine große Abneigung gegen eine auferlegte Amtssprache zu befürchten ist. Dies beweisen die indigenen Völker Lateinamerikas, sowie das Beispiel Indien. Aus diesem Grund wäre es sinnvoll, eine Mischform anzudenken. In der Grundschule wäre dabei in jedem Fall in der regionalen/nationalen Muttersprache zu unterrichten, der Unterricht in einer übernationalen Amtssprache sollte allenfalls in einem Schulfach stattfinden. So hätten alle Kinder die gleiche Chance, Grundlagen in ihrer Muttersprache zu erlernen. Anzudenken wäre die Einführung einer Kommunikationssprache als verpflichtendes Unterrichtsfach, das EU-weit ab dem gleichen Alter und nach dem gleichen Lehrplan unterrichtet wird, damit alle Schüler dieselben Voraussetzungen haben.

*Anja Eckelt  
Marienstr. 22  
DE-90592 Ochenbruck  
anja.eckelt@web.de*

## Literatur

- Backhaus, Klaus (ed.) (2005), *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*, Berlin: Springer Verlag.
- Bassiouney, Reem (2009), *Arabic Sociolinguistics*, Washington: Georgetown University Press.
- CIA (2010), *The CIA World Factbook 2010*, New York: Skyhouse Publishing.
- Coulmas, Florian (1992), *Language and Economy*, Oxford: Blackwell.
- Durth, Rainer / Körner, Heiko / Michaelowa, Karlheinz (2002), *Neue Entwicklungsökonomik*, Stuttgart: UTB.
- Eckelt, Anja (2010), *Sprachpolitik und sozioökonomische Faktoren*, unveröffentlichte Masterarbeit, Univ. Eichstätt.
- Ehm, Stephanie (2009), *Neuere Erkenntnisse der Auswirkungen des sogenannten Brain Drains auf die Entwicklungsländer*, München/Ravensburg: Grin Verlag.
- García Ruiz, Jesús (1997), "L'accord sur l'identité et les droits indigènes au Guatemala", *Volcans* 4/1997. Auch online: <http://pauillac.inria.fr/~maranget/volcans/04.97/accord.html>, 01.07.2010.
- Grzega, Joachim (2006), *Eurolinguistischer Parcours: Kernwissen zur europäischen Sprachkultur*, Frankfurt am Main: IKO.
- Grzega, Joachim (2010), "Can Language Make You Economically Human and Happy? Some Statistical Views on Socioeconomic Performance, Language Policies and Patterns of Communication", *Journal for EuroLinguistiX* 7: 1-22.
- Grzega, Joachim (Ms.), "On the Correlation between Socioeconomics and Policies of Languages in Official Contexts", noch unveröffentlichtes Manuskript, Univ. Eichstätt-Ingolstadt.
- ICC Deutschland (o.J.), "ICC-Incoterms Regeln", <http://www.icc-deutschland.de/index.php?id=35>, 15.08.2010.
- Ikonomu, Demeter Michael (2008): *Mehrsprachigkeit und ihre Rahmenbedingungen: Fremdsprachenkompetenz in den EU-Ländern*, Frankfurt am Main: Europäische Hochschulschriften.
- Janssen, Jürgen / Laatz, Wilfried (2007), *Statistische Datenanalyse mit SPSS für Windows*, Berlin: Springer.
- Kohn, Wolfgang (2004), *Statistik. Datenanalyse und Wahrscheinlichkeitsrechnung*, Berlin: Springer.
- Leclerc, Jacques (1999ff.), *L'aménagement linguistique dans le monde*, Québec: TLFQ, Université Laval, <http://www.tlfq.ulaval.ca/axl/index.html>, 25.06.2010.
- Lewis, M. Paul (ed.) (2009), *Ethnologue: Languages of the World*, Dallas, Tex.: SIL International. Auch <http://www.ethnologue.com/>, 29.06.2010.
- Ludwig, Heiner / Gabriel, Karl (2000), *Gesellschaftliche Integration durch Arbeit: Über die Zukunftsfähigkeit sozialkatholischer Traditionen von Arbeit und Demokratie am Ende der Industriegesellschaft*, Münster: LIT Verlag.
- Miles, Jeremy / Shevlin, Mark (2001), *Applying Regression & Correlation: A Guide for Students and Researchers*, London: SAGE Publications.
- Norusis, Marija (2006), *SPSS 15.0 Statistical Procedures Companion*, Upper Saddle River, NJ: Prentice Hall.
- OECD (2001), *Lernen für das Leben: Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000*, Paris: OECD. Auch online: <http://www.oecd.org/dataoecd/44/31/33691612.pdf>, 12.11.2010.
- Pool, Jonathan (1972), "National Development and Language Diversity", in: Fishman, Joshua (ed.), *Advances in the Sociology of Languages Vol. 2*, 213-230, The Hague: Mouton.
- Posdziech, Sascha (2008), *Einfluss von private Equity in der Bundesrepublik Deutschland unter besonderer Berücksichtigung volkswirtschaftlicher Faktoren*, München/Ravensburg: Grin Verlag.
- Sachs, Lothar (2004), *Angewandte Statistik*, 11. Aufl. Berlin: Springer Verlag.
- Schwarze, Jochen (1994), *Grundlagen der Statistik I*, Herne/Berlin: Verlag neue Wirtschafts-Briefe.
- Stegner, Willi (2006), *TaschenAtlas Völker und Sprachen*, Stuttgart: Ernst Klett Verlag.
- Thome, Helmut / Birkel, Christoph (2007), *Sozialer Wandel und Gewaltkriminalität: Deutschland, England und Schweden im Vergleich 1950 bis 2000. Analysen zu gesellschaftlicher Integration und Desintegration*, Wiesbaden: VS Verlag.
- United Nations (o.J.), "Statistics Division", <http://unstats.un.org/unsd/default.htm>, 20.05.2010.
- Wildmann, Lothar (2007), *Makroökonomie, Geld und Währung*, Oldenburg: Oldenburg Wissenschaftsverlag.
- Winter, Eggert / Mosenca, Riccardo / Roberts, Laura (eds.) (2010): *Gabler Wirtschaftslexikon*, Wiesbaden: Gabler. Auch online: <http://wirtschaftslexikon.gabler.de>.
- World Bank (2008), *World Development Indicators*, Washington DC: World Bank Publications.

Anhang: Datensatz aus 50 Ländern (nach *Ethnologue 2009*, CIA 2010)**Sprachpolitische Typen nach Grzega (2010)**

sprachpol. Typ	Beschreibung
A	several official languages, regionally not strictly separated (officially)
B	several official national languages (the number is given in brackets), regionally strictly separated (officially)
C	one or two supraregional national languages, several regional official languages
D	one or more native national languages + 1 or more non-native official language(s) (each spoken as a mother-tongue by less than 20% of population); numbers indicate the languages (1 = English, 2 = French, 3 = Portuguese, 4 = Swahili, 5 = Danish, 6 = German, 7 = Arabic, 8 = Spanish, 9 = Urdu, 0 = Tok Pisin)
J	one national language (with minorities covering less than 50% of the population) (incl. those countries where this holds only unofficially – this latter case is marked by brackets)

Land	spr.pol. Typ	% Muttersprachler Amtsspr.	Migrat.-rate (je 1000)	Stadtbevölk.	Lese-komp. (>15 J.)	BIP pc (US\$)	Arbeiter in Landwi.	% in Armut	Exporte (Mrdl. US\$)	Internetnutzer
Afghanistan	A	26,7%	21	24%	28,1%	800	80,0%	53%	0,327	1,73%
Algerien	D2	0,0%	-0,29	65%	69,9%	7000	8,1%	25%	75,06	10,24%
Australien	J	85,7%	6,23	89%	99,0%	38100	3,6%	14,30%	178,9	52,87%
Belarus	A	8,0%	0,38	73%	99,6%	11800	14,0%	27,10%	31,81	6,22%
Belgien	B	84,4%	1,22	97%	99,0%	37500	2,0%	15,20%	372,9	50,12%
Bolivien	C	37,9%	-1,05	66%	86,7%	4500	40,0%	60%	6,384	10,23%
Bosnien und Herzegovina	A	96,3%	3,17	47%	96,7%	6200	10,2%	25%	5,092	22,87%
Brunei Darussalam	A	2,1%	2,69	75%	92,7%	53100	4,5%	n.a.	8,25	51,40%
Burundi	D2	58,6%	4,04	10%	59,3%	400	93,6%	68%	0,047	0,67%
China	C	70,0%	-0,39	43%	90,9%	6000	43,0%	8%	1465	18,90%
Deutschland	C	91,1%	2,19	74%	99,0%	34800	2,4%	11%	1530	51,62%
Dominica	D1	14,7%	-5,45	74%	94,0%	9700	40,0%	30%	0,094	36,47%
Eritrea	A	56,1%	0	21%	58,6%	700	80,0%	50%	0,014	2,12%
Fidschii	D19	46,8%	-2,47	52%	93,7%	3900	8,9%	25,50%	1,202	8,47%
Finnland	D	95,2%	0,68	63%	100,0%	36700	5,4%	4,80%	104,3	68,57%
Großbritannien und Nordirland	J	97,0%	2,16	90%	99,0%	36500	1,4%	14%	468,7	65,78%
Guatemala	C	36,7%	-2,21	49%	69,1%	5200	50,0%	56,20%	8,028	9,94%
Haiti	D2	74,9%	-2,07	47%	52,9%	1300	28,0%	80%	0,491	11,07%
Indien	C	15,9%	-0,05	29%	61,0%	2800	60,0%	25%	175,7	6,86%
Irak	C	11,7%	-3,9	67%	74,2%	4000	21,6%	53%	66,1	0,19%
Irland	A	96,8%	4,71	61%	99,0%	46200	6,0%	7%	128	40,64%
Israel	A	74,0%	2,37	92%	97,1%	28200	2,0%	21,60%	54,16	27,65%
Italien	C	93,8%	2,06	68%	98,4%	31000	4,2%	13%	566,1	55,05%
Kanada	A	82,7%	5,63	80%	99,0%	39400	2,0%	10,80%	461,8	83,61%
Kasachstan	A	75,7%	-3,3	58%	99,5%	11500	31,5%	13,80%	66,57	12,34%
Kenia	A	0,4%	0	22%	85,1%	1600	75,0%	50%	4,882	7,69%
Kolumbien	C	75,7%	-0,26	74%	90,0%	8900	22,4%	49,20%	41,08	26,51%
Komoren	D27	0,2%	0	28%	56,5%	1000	80,0%	60%	0,032	2,80%
Kyrgyzstan	A	74,2%	-2,57	36%	98,7%	2000	48,0%	40%	1,676	13,80%
Lesotho	D1	89,3%	-0,78	25%	84,8%	1500	86,0%	49%	1,06	3,23%
Libanon	D2	0,4%	0	87%	87,4%	11100	7,0%	28%	3,5	23,65%



Land	spr.pol. Typ	% Mutter- sprachler Amtsspr.	Migrat.- rate (je 1000)	Stadt- bevölk.	Lese- komp. (>15 J.)	BIP pc (US\$)	Arbeiter in Landwi.	% in Armut	Exporte (Mrdl. US\$)	Internet- nutzer
Luxemburg	D26	59,9%	8,44	82%	100,0%	81600	2,2%	6%	20,78	70,15%
Marokko	D1	0,0%	-0,72	56%	52,3%	3800	44,6%	15%	16,14	20,94%
Mexico	C	82,1%	-3,61	77%	91,0%	1538	15,1%	13,80%	294	20,52%
Neuseeland	A	79,8%	2,46	87%	99,0%	27900	7,0%	n.a.	29,53	79,75%
Nicaragua	C	79,6%	-1,11	57%	67,2%	2900	29,0%	48%	3,183	2,63%
Niger	D2	68,6%	-0,57	16%	28,7%	700	90,0%	63	0,428	0,26%
Nigeria	D1	48,1%	-0,1	48%	68,0%	2200	70,0%	70%	83,09	6,70%
Österreich	C	90,4%	1,85	67%	98,0%	39200	5,5%	5,90%	163,3	52,09%
Paraguay	A	81,9%	-0,07	60%	94,0%	4100	31,0%	32%	6,901	4,00%
Peru	C	73,3%	-0,95	71%	92,9%	7800	0,7%	44,50%	33,27	25,84%
Phillipinen	D1	33,6%	-1,34	65%	92,6%	3300	35,0%	30%	49	5,41%
Ruanda	D12	70,3%	2,17	18%	70,4%	900	90,0%	60%	0,219	0,95%
Russland	C	80,8%	0,28	73%	99,4%	15800	10,2%	15,80%	476	21,42%
Schweiz	B	89,0%	1,76	73%	99,0%	40900	3,9%	7,60%	172,7	60,62%
Spanien	C	65,0%	0,99	77%	97,9%	34600	4,0%	19,80%	292,8	48,59%
Sri Lanka	A	90,4%	-1,09	15%	90,7%	4100	34,7%	22%	8,1	3,62%
Südafrika	A	81,5%	-0,13	61%	86,4%	9800	9,0%	50%	81,47	10,40%
USA	C	71,7%	4,31	82%	99,0%	46800	0,6%	12%	1377	72,59%
Zypern	A	96,4%	0,42	70%	97,6%	28600	8,5%	10%	0,153	47,70%

*erste Version eingereicht am 19. November 2010  
überarbeitete Version eingereicht am 21. Dezember 2010*