

Klassifizierung nominaler Anglizismen anhand von Zeitverlaufskurven

Armin Schmidt
Sabine Arnold

HS „Empirische Verfahren in Lexikologie und Lexikographie“ (Prof. Engelberg)
WS 2008/09, Germanistisches Seminar, Universität Heidelberg
April 2009

Zusammenfassung

In dieser Hausarbeit stellen wir eine Studie vor, mit der es uns gelungen ist, nominale englische Entlehnungen in der deutschen Sprache anhand ihrer und der Zeitverläufe ihrer deutschen Entsprechungen zu klassifizieren. Wir stellen verschiedene Klassifizierungen vor und diskutieren, inwiefern sie sich aufeinander abbilden lassen. Als Grundlage für unsere Untersuchung verwenden wir eine Liste nominaler Anglizismen, die wir automatisch aus großen Textkorpora extrahieren konnten. Wir besprechen die Funktionsweise der von uns entwickelten Extraktionsmethode, ihre Möglichkeiten und Beschränkungen. Im Fokus der Diskussion wird außerdem das Finden deutscher Entsprechungen für englische Entlehnungen stehen.

Inhaltsverzeichnis

1	Einleitung	3
2	Literaturübersicht	4
2.1	Erweiterung unseres Lexikons	4
2.2	Wortentlehnungsarten	5
3	Extraktion nominaler Anglizismen	8
3.1	Unsere Extraktionsmethode	9
3.2	Auswahl und Vorverarbeitung der Daten	9
3.3	Ergebnisse der Extraktionsmethode und deren Filterung	11
3.4	Grenzen des Extraktionsverfahrens	13
4	Klassifikation von Anglizismen	14
4.1	Finden deutscher Entsprechungen von Anglizismen	14
4.2	Erstellung von Zeitverlaufskurven	17
4.3	Klassifikation anhand von Zeitverlaufskurven	18
5	Fazit	26
	Literatur	27

1 Einleitung

In THIELE (2006) werden zwei Listen mit vermeintlich vom Aussterben bedrohten bzw. bereits ausgestorbenen deutschen Wörtern zusammen mit deren Begriffserklärungen bzw., falls vorhanden, ihren modernen Entsprechungen vorgestellt. Diese Listen sind nicht auf bestimmte Wortarten beschränkt und umfassen Nomen, Verben, Adjektive sowie einige Redewendungen. Aus methodischer Sicht ist problematisch, dass Thiele keinerlei Belege für die Richtigkeit seiner Angaben vorweist, d.h. nicht empirisch belegt, dass die einzelnen Einträge tatsächlich nicht mehr oder immer weniger in Gebrauch sind. Auch die Korrektheit der von ihm angebotenen modernen Entsprechungen bleibt letztlich fragwürdig, da auch hier entsprechende Belege fehlen.

Bei der Durchsicht des besagten Werks fiel uns auf, dass viele der angebotenen Entsprechungen dem Englischen entstammen, also sogenannte *Anglizismen* sind. Wie jedoch schon angedeutet wurde, widersprachen die Angaben oftmals unserer Intuition und Erfahrung; Beispielsweise war ein deutsches Wort unserem Empfinden nach keineswegs vom Aussterben bedroht oder seine englische Entsprechung deckte lediglich einen eng begrenzten seiner Bedeutungsaspekte ab.

Im Folgenden beschränkten wir unsere Betrachtung auf die ca. 60 Anglizismen in THIELE (2006, Teil "Rotbuch") und deren deutsche Entsprechungen. Es fiel uns auf, dass diese sich in Bezug auf ihre Verwendung keineswegs gleich verhielten. Durch ein primär introspektives Vorgehen teilten wir sie in sechs Gruppen ein, deren Mitglieder sich jeweils ähnlich verhielten:

1. Ersatz eines deutschen Begriffs in Bezug auf einen bestimmten Bedeutungsaspekt: In dieser Gruppe findet sich z. B. das Wortpaar 'Verabredung' - *Date*. Während eine Verabredung ein neutraler Begriff ist, der in Bezug auf viele Formen von Treffen zwischen zwei oder mehreren Personen Verwendung findet, bezieht sich das Wort *Date* lediglich auf ein romantisches Treffen zwischen zwei Menschen, ist also spezifischer als das deutsche Pendant.
2. Komplette Ablösung eines deutschen Begriffs durch einen Anglizismus: Am Beispiel des Wortes 'Kassenschlager' lässt sich zeigen, dass ein deutscher Begriff vollständig durch einen (oder mehrere) englische, in diesem Fall *Bestseller* oder auch *Blockbuster*, abgelöst werden kann.
3. Neuheitsbezeichnungen: In diese Gruppe gehören Begriffe, die beispielsweise moderne Erfindungen bezeichnen und kein deutsches Äquivalent besitzen. Der *Airbag* gehört hierhin, oder auch der *Scanner*. Zwar wurden von Thiele auch in diesen beiden Fällen deutsche Entsprechungen, 'Luftkissen' bzw. 'Einlesegerät', angegeben. Diese hielten wir jedoch für invalide. Ein 'Luftkissen' meint zumeist den namensgebenden Teil eines Luftkissenbootes, während ein 'Einlesegerät' auch für Lochkarten oder Fingerabdrücke verwendet werden kann, ein *Scanner* im allgemeinen Sprachgebrauch jedoch ausschließlich für ein Gerät zum Übertragen von Gedrucktem in Computerdateien.
4. Anglizismus ist Oberbegriff für mehrere Entsprechungen: In dieser Gruppe findet sich nur das von Thiele angeführte Wortpaar 'Butterbrot' - *Sandwich*. Während 'Butterbrot' ausschließlich eine lediglich mit Butter beschmierte Scheibe Brot bezeichnet, kann ein *Sandwich* ein mit den verschiedensten Belägen versehenes doppeltes Brot unterschiedlicher Art sein.

5. Übersetzungen: Am Beispiel von *Bodyguard* lässt sich sehen, dass ein deutsches Wort, in diesem Fall 'Leibwächter', so übersetzt werden kann, dass alle Bedeutungsaspekte des einen durch den anderen Begriff abgedeckt sind. Diese bloße Feststellung sagt allerdings noch nichts darüber aus, ob das deutsche Wort tatsächlich verdrängt wird, sich durchsetzt, oder beide Wörter nebeneinander existieren können.
6. Fehldiagnose: Diese Gruppe fungiert als Sammelklasse für alle von Thiele angegebenen Wortpaare, die unserem Empfinden nach schlicht fehlerhaft analysiert worden sind. Dazu gehört z. B. das Wort *T-Shirt*, von Thiele mit 'Unterhemd' übersetzt. Ein 'Unterhemd' hat jedoch typischerweise gar keine Ärmel und wird direkt am Körper und unter einem Hemd oder T-Shirt getragen, während ein *T-Shirt* kurze Ärmel besitzt, oft mit Aufdrucken versehen ist und als Oberteil getragen wird.

Diese Klassifizierung kann nur eine vorläufige sein und allenfalls als Orientierung für weitere Arbeit dienen. Sie beruht auf dem Sprachempfinden zweier Probanden und einer Wortliste, deren Validität zweifelhaft ist. Welche Eigenschaften sind es, die die Mitglieder einer Gruppe vereinen, sie gleichzeitig jedoch von denen einer anderen Gruppe abgrenzen? Wenn das Ziel die Klassifizierung von Anglizismen sein soll, müssen fest definierte Merkmale gefunden werden, welche einen entsprechenden Versuch operationalisierbar machen. Um dem etwas näher zu kommen, betrachten wir noch einmal Gruppe 2. Komplette Ablösung bedeutet konkret, dass ein deutsches Wort immer weniger häufig, sein englisches Pendant jedoch immer häufiger verwendet wird. Das Verhalten von Wörtern im Laufe der Zeit lässt sich durch aus großen Textkorpora berechneten Zeitverlaufskurven darstellen. Letztere zeigen die Frequenz des entsprechenden Begriffs innerhalb einer bestimmten Zeitperiode. Bei Wortpaaren der Gruppe 2 würde man nun erwarten, dass die Zeitverlaufskurven des deutschen und des englischen Begriffs sich in etwa wie in Abbildung 1 schematisch dargestellt verhielten.

Ausgehend von diesen Beobachtungen kommen wir zu folgender Hypothese als Grundlage für die vorliegenden Arbeit:

Anglizismen lassen sich in distinkte Klassen einteilen, welche sich anhand ihrer Zeitverlaufskurven definieren lassen.

2 Literaturübersicht

2.1 Erweiterung unseres Lexikons

Im Folgenden soll ein sprachwissenschaftlicher Überblick geschaffen sowie die in der Linguistik üblichen Bezeichnungen für die Klassifikation von Anglizismen vorgestellt werden. Der Wortschatz, also das Lexikon unserer Sprache, setzt sich aus ca. 300 000 bis 500 000 Erb-, Fremd- und Lehnwörtern (vgl. ZÜRN, A., 2001) zusammen. Der Erbwortschatz beinhaltet Wörter, die ihren Ursprung im Indogermanischen bzw. Germanischen haben und wird im Laufe der Zeit durch die Aufnahme von Fremd- und Lehnwörtern erweitert und verändert. Diese unterscheiden sich dadurch, dass die Fremdwörter hinsichtlich ihrer Akzentuierung, Lautung oder Schreibweise nicht in die neue Sprache integriert sind, die Lehnwörter dagegen schon. Ist es aufgrund soziokultureller Ereignisse oder moderner Erfindungen

Klasse 2: vollständige Ablösung

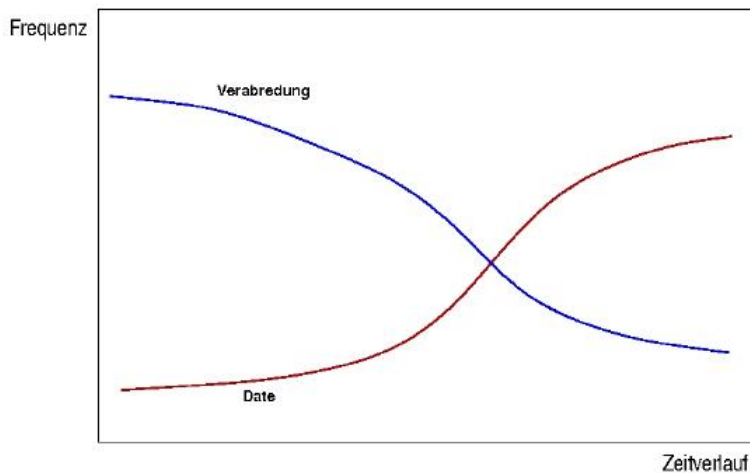


Abbildung 1: Hypothetischer Zeitverlauf für das Wortpaar *Verabredung* - *Date*

notwendig neue Wörter in unsere Sprache aufzunehmen, gibt es dafür die unterschiedlichsten Wege (vgl. ZÜRN, A., 2001, S. 21), die in Abschnitt 2.2 kurz vorgestellt werden sollen.

Die Gründe für die Aufnahme von Fremdwörtern haben in den letzten Jahrhunderten keine weitgreifende Veränderung erfahren. In erster Linie wurde die Bezeichnung eines neuen Gegenstandes aus dem jeweiligen Herkunftsland direkt übernommen, wie beispielsweise die Begriffe rund um den Weinanbau aus dem alten Rom. Desweiteren ist der Prestigewert nach wie vor ein nicht zu unterschätzender Faktor bei der Entlehnung von Fremdwörtern. Übernahm man zur Zeit des Humanismus das Lateinische oder im Barock das Französische, brachte die Christianisierung darüber hinaus griechisches Wortgut mit sich. Speziell in unserer heutigen Zeit spielt die Globalisierung und die damit einhergehende internationale Zusammenarbeit und rege Übersetzungstätigkeit eine große Rolle.

2.2 Wortentlehnungsarten

Die Unterscheidung zwischen Fremd- und Lehnwort ist nicht immer eindeutig, da sie zum größten Teil von den Sprechern abhängig ist, bei denen etymologisches Wissen und die damit verbundenen Hintergrundinformationen stets variieren können. Grob betrachtet lassen sich die unterschiedlichen Entlehnungsarten in lexikalische und semantische Entlehnungen einteilen.

2.2.1 Scheinentlehnungen

Bei den sogenannten Scheinentlehnungen, die sich dadurch auszeichnen, dass sie zwar mit den sprachlichen Mitteln der Ursprungssprache gebildet werden, in dieser jedoch unbekannt sind, unterscheidet

man darüber hinaus zwischen lexikalischen, morphologischen und semantischen Entlehnungen. Die lexikalischen Bildungen folgen meist einem englischsprachigen Vorbild und werden analog gebildet, wie beispielsweise *Showmaster* oder *Callboy*, die höchst wahrscheinlich den *quizmaster* und das *callgirl* als Ursprung hatten. Die morphologischen Scheinentlehnungen stellen Kürzungen von Wortübernahmen dar, die im Englischen so nicht existieren, wie beispielsweise *Pulli* von *pullover* oder *Deo* von *Deospray*. Auf semantischer Ebene übernimmt ein Anglizismus eine nicht-englische Bedeutung, wie man bei dem stark frequentierten Wort *Gag* sehen kann: Dieser Begriff ist als Wortwitz oder originelle Erfindung in unseren Sprachgebrauch eingegangen, im Englischen hingegen lautet die korrekte Bezeichnung *gimmick* oder *trick* (ZÜRN, A., 2001, S. 34-35).

2.2.2 Mehrfachentlehnungen

Neben diesen einzelnen Sprachschöpfungen existieren auch Wörter, die zu unterschiedlichen Zeiten und innerhalb unterschiedlicher Sprachschichten gleich mehrmals entlehnt werden. Aufgrund dieser sogenannten Mehrfachentlehnungen kann es zur Bildung von Homonymen kommen, also Lexemen mit derselben Form, aber anderer Bedeutung und teils unterschiedlicher Wortherkunft (ZÜRN, A., 2001, S. 39). Zu dieser Gruppe gehören zum Beispiel die Wörter *Single*, *Toast* oder *Killer*.

2.2.3 Semantische Entlehnungen

Parallel zu den direkten, also lexikalischen Wortübernahmen gibt es auch die indirekten semantischen Übernahmen, die oft nicht als solche zu erkennen sind, da sie wiederum mit deutschem Wortmaterial nachgebildet werden. Zum Einen existieren die Lehnbedeutungen, die zwar auch den fremdsprachlichen Ursprung besitzen, aber eine Bedeutungserweiterung oder semantische Umdeutung erfahren, also nicht das Lexem als solches übernehmen. Ein passendes Beispiel wäre hier der *Pony*, dessen etymologischer Ursprung auf das kleinwüchsige Pferd zurück zu führen ist und in unserer Sprache die Bezeichnung einer Damenfrisur übernommen hat. Zum Anderen gibt es die Lehn schöpfungen, die sich dadurch auszeichnen, dass sie komplett neue Bezeichnungen erhalten, weil sie für den Sprecher unverständlich sind und eine Übersetzung als zu umständlich angesehen wird. Gängige Beispiele sind *air condition* - 'Klimaanlage', *laser* - 'gebündelter Lichtstrahl' oder *dimmer* - 'Helligkeitsregler'. Wird das Fremdwort doch unmittelbar übersetzt, spricht man von der Lehnübersetzung, wobei die Wortbildungsregeln hier genau beachtet werden und Begriffe wie *entertainment* - 'Unterhaltung', *network* - 'Netzwerk' ihren Platz in unserer Sprache einnehmen. An dieser Stelle sollen auch die so genannten false friends erwähnt werden, die durch einen irrtümlichen Fremdwortgebrauch sowohl auf lexikalischer als auch auf semantischer Ebene zustande kommen (ZÜRN, A., 2001, S. 49-55). Zum Beispiel bedeutet *artist* 'Künstler' und nicht 'Artist', *murder* 'Mord' und nicht 'Mörder' oder *to wink* nicht 'winken' sondern 'zuzwinkern'. Dieses Phänomen der phonologischen Ähnlichkeit könnte auch auf die Etymologie des Wortes 'winken' zurückzuführen sein, das laut dem Deutschen Wörterbuch der Gebrüder Grimm seinen Ursprung nicht bloß im Althochdeutschen hat, sondern auch im angelsächsischen *wincian* und mittelenglischen *winkin* (GRIMM/GRIMM, 1854-1960)¹.

¹Die Online-Ausgabe dieses Werks wird vom Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier zur Verfügung gestellt und ist unter <http://germazope.uni-trier.de/Projects/WBB/woerterbuecher/woerterbuecher/dwb/wbgui> zugänglich.

2.2.4 Evidentes und latentes Lehnwort

Eine etwas spezifischere Klassifizierung des englischen Lehnworts findet sich bei GLAHN (2000), der sich mit dem „Einfluss des Englischen auf gesprochene deutsche Gegenwartssprache“ beschäftigt. Glahn zieht dazu ein ursprünglich von BETZ (1965) entwickeltes und von CARSTENSEN/GALINSKY (1975) erweitertes Modell heran, welches zwischen evidentem und latentem Lehnwort unterscheidet. Das evidente Lehnwort bezeichnet die direkte Übernahme eines englischen Wortes, das aufgrund der Form oder auch durch die Aussprache seinen Ursprung erkennen lässt. Unter diese Einteilung fallen somit Fremdwörter und Scheinentlehnungen als direkte Entlehnungsarten und Scheinentlehnung und Mischkomposita als indirekte Entlehnungsarten. Mischkomposita haben wir bei unserer Untersuchung ebenso aufgrund des angestrebten Umfangs der Untersuchung außen vor gelassen. Latentes Lehnwort beinhaltet die Lehnbedeutung und -bildung, die sich wiederum in Lehnwortschöpfung und -formung unterteilen lässt. Bei der Lehnformung wird erneut zwischen Lehnübersetzung und -übertragung unterschieden (GLAHN, 2000, S. 36-39). Diese Klassifikation unterscheidet sich in ihren Grundzügen nur minimal von der Zürns. Sie legt vor allem Wert auf die Unterscheidung zwischen lexikalischen und semantischen Erscheinungen, die bei Glahn unter evidentem Lehnwort zusammengefasst sind. Ausführlicher wird bei Zürn darüber hinaus das Phänomen der Bedeutungsverengung und -erweiterung dargestellt.

2.2.5 Bedeutungsverengung- und -erweiterung

Es kann passieren, dass bei der Entlehnung eines Wortes nur ein Teil der Bedeutungsvariante übernommen wird. Bei dieser Bedeutungsverengung verliert das Lehnwort viele seiner ursprünglichen Verwurzelungen auf semantischer Ebene und geht in die neue Sprache mit anderer oder relativierter Bedeutung ein (vgl. ZÜRN, A., 2001, S. 58-60), wie folgende Beispiele zeigen: Das Lexem *cover* fand unter der Bedeutung 'Hülle für CD's, Bücher oder Zeitschriften' Einzug in unsere Sprache, obwohl es im Englischen die Bedeutungen 'Decke, Hülle, Überzug, Umschlag' trägt. Der *swimming pool* ist in seiner Herkunftssprache für die Allgemeinheit bestimmt, im Deutschen dagegen wurde die Nutzung auf das eigene Grundstück beschränkt. Im Gegensatz zur Bedeutungsverengung kann es zur Bedeutungserweiterung kommen. In diesem Fall geht das Lexem die Verbindung mit anderen semantischen Ebenen ein: engl. *camp* wird zum Beispiel von seiner ursprünglichen Bedeutung 'Gefangenenlager' im Deutschen auf 'Ferienlager' erweitert, das *team* wird von der 'Sportmannschaft' zu einer allgemeinen Verbindung, die aus beliebig vielen Menschen besteht oder darüber hinaus eine 'Arbeitsgruppe' bezeichnen kann. Nach der Darlegung unterschiedlicher Entlehnungsarten bleibt zu betonen, dass die Gründe für diese Variationen in der Regel auf die Sprechergemeinschaft zurück zu führen sind, da es aufgrund mangelnder Englischkenntnisse immer wieder zu fehlerhaften Übersetzungen, Verwechslungen oder Analogiebildungen kommen kann.

2.2.6 Prestigewert und Substitutionsmöglichkeiten von Anglizismen

Ein Anglizismus wird auch in vielen Fällen dem deutschen Wort vorgezogen, weil dem Fremdwort meist ein höherer gesellschaftlicher, künstlerischer oder stilistischer Wert zugesprochen wird. In diesem Zusammenhang sollte man zwischen der Denotation, also dem Kern einer Bedeutung, und der

Konnotation, dem Gefühlswert und Nebensinn eines Wortes unterscheiden, wobei diese beiden Faktoren parallel fungieren und von Sprecher zu Sprecher variieren können (ZÜRN, A., 2001, S.127-128). So ist der *award* im Gegensatz zu der 'Auszeichnung' mit einer moderneren Bedeutung versehen. Während ein *award* eher in kulturellen Bereichen wie Musik und Film vergeben wird, spricht man vor allem im militärischen oder politischen Bereich von der 'Auszeichnung'. Nahezu alle gezeigten Entlehnungen überschneiden sich in bestimmten Kernbereichen, können sich also folglich an einer bestimmten Kontextstelle gegenseitig vertreten und gelten somit als synonym. Man spricht in diesem Fall auch von der Substitution, die für unsere Untersuchung im Speziellen eine große Rolle spielt, da wir eben dieses Phänomen näher beobachten wollen. Nun existieren aber auch Anglizismen in unserer Sprache, denen kein passendes Synonym zuzuordnen ist. Dies ist meistens dann der Fall, wenn bestimmte Erfindungen wie *Hightech* oder *Cockpit* direkt mit der englischen Bezeichnung in unserer Sprache eingegangen sind, ohne vorher einen Platz in unserem Wortschatz gehabt zu haben.

2.2.7 Grammatikalische und orthographische Assimilierung

Die Eingliederung der Fremdwörter in die morphematische Struktur unseres Sprachsystems ist unabdingbar, da diese ansonsten keine angemessene Integration erfahren können. Verben werden demnach deutschen Konjugationsformen angepasst, Adjektive der entsprechenden Flexion zugeordnet und Substantive mit Genus-, Numerus- und Kasusmerkmalen versehen. Da im Englischen nur der Artikel *the* existiert, erscheint eine Genusmarkierung oft willkürlich und ist in der Sprachwissenschaft ein umstrittenes Thema (ZÜRN, A., 2001, S. 179-181). So gibt es Lexeme, die gleich zwei Genera zu zuordnen sind, wie beispielsweise der bzw. das 'Terminal' als Abfertigungshalle am Flughafen und die Eingabeeinheit einer EDV-Anlage oder genau so 'Quiz', 'Looping' oder 'Display' mit neutraler und maskuliner Endung. Die letztendliche Zuordnung der Substantive erfolgt in erster Linie mithilfe der entsprechenden semantischen Kriterien, wie dem natürlichen Geschlecht bei Personenbezeichnungen: Die 'Diva', der 'Playboy' oder der Orientierung an dem entsprechenden deutschen Lexem: Der 'Rechner' - der 'Computer' oder das 'Zusammentreffen' - das 'Meeting'. Darüber hinaus nehmen die entlehnten Wörter in der Regel die Großschreibung an, was man bei vielen Lexemen auch als Merkmal dafür sehen kann, wie weit das Fremdwort in die neue Sprache integriert ist. Wird es groß geschrieben, hat es durch zunehmende Verwendung und Gewöhnung seine Kennzeichnung als neues Wort weitgehend verloren und geht in den alltäglichen Sprachgebrauch über. Jedoch gibt es zu dieser Behauptung auch die Gegenannahme, dass die Großschreibung keineswegs auf eine hohe Integration schließen lässt, sondern einfach nur die automatische Großschreibung im Deutschen erfährt wie jedes andere Substantiv auch und außerdem Besonderheiten mit Innenmajuskel wie 'eMail' oder 'HipHop' aufweist (vgl. ADLER, 2004, S. 70).

3 Extraktion nominaler Anglizismen

Eine zentrale Voraussetzung für die Aufgabe der Klassifikation nominaler Anglizismen im Deutschen ist eine Liste von ebensolchen. Die Einträge einer solchen Liste müssen bestimmten Voraussetzungen genügen, welche sie zur Weiterverarbeitung und näheren Untersuchung geeignet machen. Existierende

Auflistungen wurden üblicherweise manuell zusammengetragen und entsprechen unseren Anforderungen oftmals nur zu einem geringen Teil. THIELE (2006) beispielsweise macht weder Angaben über seine Methodik bei der Erstellung seiner Listen, noch wird die Richtigkeit seiner Angaben, bestimmte deutsche Wörter seien von den entsprechenden englischen Termini ersetzt worden, empirisch belegt. Der Verein für Deutsche Sprache bemüht sich in seinem Anglizismen-Index² um Vollständigkeit einer Auflistung, macht jedoch kaum Angaben zum methodischen Vorgehen bei der Erstellung derselben³, so dass davon ausgegangen werden muss, dass keine umfassende empirische Verifizierung stattgefunden hat. Andere im Internet verfügbare Sammlungen von Anglizismen⁴ sind von Nutzern zusammengetragen worden und können als Ausgangspunkt für explorative Untersuchungen dienen, sollten aber nicht als verlässliche und repräsentative Quelle valider Anglizismen genutzt werden, da kein Nachweis auf Vollständigkeit, Domänenabhängigkeit, oder Konzentration auf bestimmte Wortarten gegeben ist.

Im Folgenden werden wir eine Methode vorstellen, mit der es uns gelungen ist, eine Auflistung von Anglizismen brauchbarer Größe mittels Abgleich eines englischen und deutschen Textkorpus zu extrahieren. Schließlich werden wir diskutieren, inwiefern die so gewonnene Kandidatenliste den Voraussetzungen für eine vollständige Klassifikation entspricht.

3.1 Unsere Extraktionsmethode

Anglizismen im Deutschen können, wie schon in Abschnitt 2 besprochen, in den unterschiedlichsten Spielarten auftreten. Da wir uns in dieser Arbeit auf die Klasse der nominalen Anglizismen beschränken wollen, sehen wir uns in der glücklichen Lage, viele jener Besonderheiten ignorieren zu können, die beispielsweise auf verbale Anglizismen zutreffen (Vergleich Abschnitt 2).

Diese Beobachtungen weisen uns den Weg zu einer einfachen Methode, nominale Anglizismen mit Hilfe einer Liste englischer Nomen und einem Textkorpus deutscher Sprache automatisch zu extrahieren: Wir iterieren über das deutsche Korpus und prüfen, ob es sich dabei um ein englisches Nomen handelt oder nicht. Wenn ja, wird das entsprechende Wort als potentieller Anglizismus in die Kandidatenliste aufgenommen. Algorithmus 1 beschreibt das Verfahren formal.

3.2 Auswahl und Vorverarbeitung der Daten

Der Erfolg dieses Verfahrens hängt stark von der Auswahl der verwendeten Ressourcen ab. In Bezug auf das deutsche Korpus haben wir uns für die Textsammlung der Online-Ausgabe der Zeitung *Die Zeit*⁵ entschieden. Diese besteht aus insgesamt ca. 10000 Artikeln der Jahre 2000 und 2001 und beinhaltet gut 10 Millionen laufende Wortformen. Die Artikel sind den Themenbereichen Hochschule, Kultur, Leben, Media, Politik, Reisen, Wirtschaft sowie Wissen zugeordnet und decken somit ein

²<http://www.vds-ev.de/anglizismenindex/>

³Als Quellen werden einige Internet-Suchmaschinen und Online-Wörterbücher angegeben. Siehe <http://www.vds-ev.de/anglizismenindex/autoren.php>

⁴http://de.wikipedia.org/wiki/Liste_von_Anglizismen

http://de.wiktionary.org/wiki/Wiktionary:Deutsch/Liste_der_Anglizismen

⁵<http://www.zeit.de> - *Das Korpus selbst liegt am Seminar für Computerlinguistik der Universität Heidelberg vor.*

Algorithm 1 Basis-Algorithmus zur Extraktion nominaler Anglizismen

1. Gegeben sei eine Liste unikaler englischer Nomen $N_e = \{n_{e1}, n_{e2}, \dots, n_{en}\}$, sowie ein deutsches Textkorpus $K_d = \{w_{d1}, w_{d2}, \dots, w_{dn}\}$ mit den laufenden Wortformen w_{d1} bis w_{dn}
 2. Prüfe für jedes Wort w_{di} in K_d , ob es ein Nomen n_{en} in N_e gibt, das diesem entspricht ($w_{di} = n_{en}$).
 3. Wenn Schritt 2 ein positives Ergebnis liefert, füge das entsprechende Wort der Ergebnisliste potentieller Anglizismen K (Kandidatenliste) hinzu.
 4. Filtere K , so dass nur „echte“ Anglizismen übrig bleiben.
-

breites topikales Spektrum ab, wobei wir hier der Intuition folgen, dass verschiedene Themenbereiche unterschiedliche Anglizismen verwenden und wir auf diese Weise dem Kriterium der Abdeckung Folge leisten. Auch die Anforderung der Aktualität ist weitestgehend erfüllt, wobei darauf hingewiesen werden soll, dass noch aktuellere Zeitungskorpora für unsere Zwecke kaum vorhanden oder aus Gründen des Urheberrechtes auch für Forschungszwecke kaum zugänglich sind.

Da uns das Korpus als Sammlung von HTML-Dokumenten vorlag, welche von der Webseite der *Zeit* abgerufen worden waren, haben wir zunächst alle Dateien mit Hilfe eines Python-Skripts von HTML-Elementen befreit. Anschließend haben wir die Texte in eine einzige Datei überführt, welche den gesamten bereinigten Korpus in vertikalierter Form (d.h. ein Wort pro Zeile) beinhaltet. Um die Laufzeit unseres Algorithmus zur Extraktion nominaler Anglizismen gering zu halten sowie um Fehlern vorzubeugen, haben wir anschließend aus dem Korpus *potentielle* Nomina extrahiert. Dabei sind wir der einfachen Heuristik gefolgt, dass Nicht-Nomen im Deutschen im Allgemeinen klein geschrieben werden, es sei denn, sie stehen am Satzanfang. Der Einfachheit halber und mit dem Vertrauen auf die ausgleichende Kraft schierer Datenmenge haben wir also alle großgeschriebenen Wörter extrahiert, die nicht auf einen Punkt folgen und abweichende Fälle (z.B. vorangehende Frage- oder Ausrufesätze) ignoriert. Diese stark vereinfachende Annahme brachte gleichzeitig mit sich, dass viele falsche Positive, d.h. Wortformen, die im Korpus groß geschrieben vorkommen und zufällig auf einen Punkt folgen, aber einer anderen Wortart als der nominalen zugehören, ebenfalls extrahiert wurden. Letzteres stellt jedoch kein Problem dar, da der spätere Abgleich mit der englischen Nomenliste sowie weitere Filtermethoden diese Fälle aussortieren sollten. An dieser Stelle sei nochmals darauf hingewiesen, dass es uns bei der Extraktion nominaler Anglizismen nicht um Vollständigkeit einer entsprechenden Liste ging, sondern lediglich darum, diese mit empirischen Verfahren zu gewinnen, um so die Verlässlichkeit unserer Ergebnisse sicherzustellen. Nach allen bisher beschriebenen Vorverarbeitungsschritten erhalten wir eine Liste potentieller Nomina des Deutschen mit ca. 3.74 Millionen Einträgen. Zu beachten ist, dass diese Liste einzelne Nomen mehrfach enthält, da wir für spätere Schritte eine Angabe über die Häufigkeit eines jeden Nomens benötigen.

Mit einer Liste deutscher Nomina im Gepäck wurde nun noch eine Liste englischer Nomina benötigt um unseren Extraktionsalgorithmus anwenden zu können. Da uns keine solche vorlag und für unsere Aufgabestellung außerdem Aktualität und Vollständigkeit empirisch verifizierbar sein sollte, haben wir uns nochmals für die automatische Extraktion aus einem Korpus entschieden. Wir applizierten für

diesen Zweck das British National Corpus (BNC)⁶, ein vollständig mit Wortarten (Parts-of-Speech, POS) annotiertes Textkorpus, welches mit rund 100 Millionen laufenden Wörtern einen repräsentativen Auszug geschriebenen und gesprochenen modernen britischen Englischs beinhaltet. Die uns vorliegende Version ist im XML-Format und gibt zu jedem laufenden Wort sowohl eine allgemeine (*basic tagset*) und eine spezifische (*enriched tagset*) Wortartenklassifikation als auch das entsprechende Lemma an. Das allgemeine Tagset unterscheidet in Bezug auf Nomen zwischen Eigennamen und gewöhnlichen Nomen. Letztere werden noch einmal anhand ihres Numerus unterschieden, wobei das Tag N1 und N2 Nomen als Singular respektive Plural, und N0 Nomen auszeichnet, welche sich in Bezug auf das Numerus neutral verhalten (z.B. *committee* im Satz „The committee disagrees/disagree.“). Da Eigennamen für die von uns angestrebte Klassifizierung keine Rolle spielen, haben wir lediglich Wörter mit den Tags NN0, NN1 und NN2 betrachtet und mit Hilfe eines XSLT-Skripts zu jedem das jeweilige Lemma extrahiert. Die so gewonnene Liste umfasste ca. 20 Millionen Tokens und rund 173000 Types. Anschließend kompilierten wir diese durch das Entfernen und Zählen doppelter Einträge, Eliminierung von offensichtlichen Fehleinträgen bestehend aus Zahlen, Sonder- oder Satzzeichen sowie Filtern von Hapax Legomena (Tokens, die nur einmal im Korpus auftreten)⁷ in eine Liste von knapp 70000 Types und deren jeweiligen Vorkommenshäufigkeiten im Korpus.

3.3 Ergebnisse der Extraktionsmethode und deren Filterung

Der Anwendung der in Algorithmus 1 zusammengefassten Methode auf die in Abschnitt 3.2 beschriebenen Daten (nach der Vorverarbeitung) soll der Vollständigkeit halber noch vorausgeschickt werden, dass beim Abgleich der deutschen und englischen Nomenliste Groß- und Kleinschreibung ignoriert wurde, um den sprachspezifischen Unterschieden im Deutschen und Englischen diesbezüglich gerecht zu werden. Durch die direkte Anwendung des Algorithmus erhielten wir eine Liste mit 9087 potentiellen Anglizismen sowie deren Häufigkeiten im deutschen Korpus. Da unser Vorhaben eine teilweise manuelle Untersuchung der extrahierten Anglizismen beinhaltet, sahen wir uns gezwungen, die Größe der Kandidatenliste weiter zu begrenzen, indem wir alle Einträge eliminierten, die weniger als 30 mal im deutschen Korpus vorkamen. Die so erstellte Liste mit Kandidatenliste umfasste 960 potentielle Anglizismen, eine Menge, die manuell handhabbar schien. In einem Folgeschritt sahen wir uns diese Liste genauer an und untergliederten die Einträge in die folgenden sechs Gruppen:

3.3.1 Gruppe 1 - einfache Fehler

Diese Gruppe bezeichnet diejenigen rund 50 Einträge, die weder der einen noch der anderen Sprache zugeordnet werden konnten, nur aus einzelnen Buchstaben bestanden oder nicht zuzuordnende Abkürzungen.

Beispiele: *gent, fu, l., pr, s, v. w.*

⁶<http://www.natcorp.ox.ac.uk/>

⁷Die ursprüngliche Liste enthielt sehr viele Einträge mit einer Häufigkeit von 1. Diese waren, bei näherer Betrachtung, zum allergrößten Teil keine englischen Wörter oder, vermutlich durch Fehler bei der Tokenisierung, falsch annotiert worden. Da wir davon ausgehen, dass Hapax Legomena, auch wenn sie valide englische Nomen sein sollten, kaum als Anglizismen in die deutsche Sprache gelangen würden, sehen wir das Filtern derselben als gerechtfertigt.

3.3.2 Gruppe 2 - Eigennamen

Hier sind ca. 200 Einträge zusammengefasst, die Eigennamen bezeichnen, vor allem Firmen- und Ortsbezeichnungen. Auch Teile von phrasalen Eigennamen wurden hier aufgenommen.

Beispiele: *amazon, bbc, dean, kunsthalle, kursk, linda, mick, sir, wehrmacht*

3.3.3 Gruppe 3 - nur-deutsche Begriffe

In dieser Gruppe finden sich vor allem deutsche Begriffe, die über Zitate oder Eigennamen in englische Texte geraten sein könnten. Interessanterweise hat unsere Extraktionsmethode auch einige *Germanismen* in der englischen Sprache, d.h. deutsche Wörter, die in den englischen Wortschatz eingeflossen sind, gefunden. Zu den letzteren gehören vor allem Begriffe aus der Philosophie, der Geschichte und Politik, insbesondere aus dem Bereich des Dritten Reiches.

Beispiele: *apparat, burg, hausfrau, nazi, kulturkampf, ostpolitik, sau, weltanschauung, zeitgeist*

3.3.4 Gruppe 4 - Internationalismen

Hier sind ca. 70 Wörter zusammengefasst, welche sowohl im Deutschen als auch im Englischen existieren, in beiden Sprachen genauso geschrieben werden und die gleiche bzw. eine sehr eng verwandte Bedeutung besitzen.

Beispiele: *agent, asphalt, ballett, ego, hotel, innovation, londoner, museum, papa, psyche, sack, terror, trauma, zone*

3.3.5 Gruppe 5 - Homographien

Diese Gruppe umfasst Wörter, die im Englischen und Deutschen die gleiche Schreibweise, aber eine jeweils unterschiedliche Bedeutung besitzen. Es sei darauf hingewiesen, dass die englischen und deutschen Varianten der meisten Mitglieder dieser Gruppe nicht homophon sind, d.h. auf lautlicher Ebene durchaus unterschieden werden können, in der Schrift jedoch nicht.

Beispiele: *bin, brief, frank, list, not, prise, rat, stock, ton, welch, will*

3.3.6 Gruppe 6 - Anglizismen

Dieser Gruppe sind nun alle Wörter zugeordnet, die wir als „echte“ Anglizismen betrachten, d.h. englische Wörter, die im Deutschen verwendet werden, kein Teil eines Eigennamens sein müssen und in dieser Schreibweise nicht im Deutschen vorkommen. Auch echte Anglizismen, die jedoch in derselben Schreibweise ebenfalls in der deutschen Sprache existieren, sind hier aufgenommen worden. Ein Beispiel für letztere ist das Wort *Gang*, das zwar in der Bedeutung *kriminelle Gruppe* ein valider Anglizismus ist, in der Bedeutung *Art zu gehen* ein gewöhnliches deutsches Wort sein kann. Insgesamt umfasst jedoch die Menge der „echten“ Anglizismen ca. 180 Begriffe.

Beispiele: *award, beach, butler, center, champion, crash, fan, gang, hacker, hit, joint, layout, mix, player, teenager, user, top*

3.4 Grenzen des Extraktionsverfahrens

Wie bereits in Abschnitt 3.1 erwähnt, mussten wir unsere automatisch extrahierte Liste von Anglizismen künstlich klein halten, um eine manuelle Untersuchung zu ermöglichen. Es wurden beispielsweise Hapax Legomena aus der englischen Nomenliste gelöscht und nur diejenigen der potentiellen Anglizismen näher untersucht, die mindestens 30 mal im deutschen Korpus vorkamen. Abgesehen davon können mit dem beschriebenen Verfahren verschiedenste nominale Anglizismen aufgrund methodischer Einschränkungen nicht erkannt werden. Letztere sollen im Folgenden kurz ohne Anspruch auf Vollständigkeit beschrieben werden.

3.4.1 Amerikanische Schreibweise

Da das British National Corpus im Allgemeinen nur die britische Schreibweise englischer Wörter enthält, wurden Anglizismen in amerikanischer Schreibweise nicht extrahiert. Dies ist keine prinzipielle Einschränkung unserer Extraktionsmethode, sondern lediglich durch die Auswahl des verwendeten Korpus bedingt. Zusätzlich ist bei Verwendung eines Korpus britischer Texte zu erwarten, dass bestimmte, in der deutschen Gebrauchssprache durchaus übliche, Amerikanismen gar nicht oder so selten auftreten, dass sie entweder nicht erkannt oder spätestens in einem der Filterschritte ausgesondert werden. Zu dieser Gruppe gehören beispielsweise das Wort *Truck*, das im britischen Englisch mit *Lorry* übersetzt würde.

3.4.2 Phrasale Anglizismen und Komposita

Phrasale Anglizismen werden von unserer Methode so gut wie vollständig ignoriert. Dies hat mehrere Ursachen: Zum einen ist die Wortartenannotation des British National Corpus in dieser Hinsicht nicht vollständig, da hier lediglich einzelne Wörter, getrennt durch Leerzeichen, annotiert sind und Komposita oder phrasale Ausdrücke nicht direkt extrahiert werden können.⁸ Phrasale Ausdrücke, die nicht nur aus Nomen bestehen, würde von der einfachen, in Abschnitt 3.2 beschriebenen, Heuristik grundsätzlich nicht erkannt, da sie lediglich groß geschriebene Wörter betrachtet.

Abgesehen von den genannten Problemen stellen gerade Komposita eine ganz grundsätzliche Schwierigkeit dar, da englische Komposita im Allgemeinen getrennt, im Deutschen jedoch zusammen geschrieben oder mit Bindestrich versehen werden. Die direkte Abbildung deutscher und englischer Nomen schlägt an dieser Stelle unweigerlich fehl.

⁸Ein Ansatz, dieses Problem zu lösen, wäre die Behandlung mehrerer als Nomen gekennzeichnete direkt aufeinander folgender Wörter als potentielle Komposita, die anschließend gefiltert werden müssten.

3.4.3 Pseudoanglizismen

Ein interessante Klasse von Anglizismen sind solche, die zwar eindeutig aus dem Englischen entlehnt sind, dort aber eine andere Bedeutung besitzen. Beispiele hierfür sind das 'Handy', das im Englischen *cell phone* oder *mobile phone* genannt wird, der 'Beamer', unter englischen Muttersprachlern als *video projector* bekannt, oder der 'Showmaster', auf Englisch *host*. Diese unter dem Begriff Pseudoanglizismen gehandelten Begriffe werden im Allgemeinen von unserer Extraktionsmethode nicht erkannt. Das Wort 'Showmaster' beispielsweise existiert in der englischen Sprache schlichtweg nicht. Ähnliches gilt für den Begriff 'Beamer', der im Englischen zwar bekannt ist, entweder jedoch dem Fachvokabular der Sportart Cricket zugehört oder umgangssprachlich ein Fahrzeug der Marke BMW meint. Das Wort *Handy* sollte allein deshalb nicht erkannt werden, da es im Englischen ein Adjektiv ist (*handy*: handlich, praktisch) und im BNC als solches gekennzeichnet sein sollte. Erstaunlicher Weise hat es das 'Handy' letztendlich und nach allen Filterdurchläufen doch auf unsere Liste der „echten“ Anglizismen geschafft, was einerseits auf einige fehlerhafte Wortartenannotationen im BNC, und andererseits auf die sehr hohe Frequenz des Wortes in der deutschen Sprache zurückzuführen sein könnte.

3.4.4 Andere Wortarten

Grundsätzlich kann unsere Extraktionsmethode in ihrer derzeitigen Form nur begrenzt auf andere Wortarten angewendet werden, da die Abbildung deutscher und englischer Begriffe gerade bei stark flektierten Wortarten schwierig ist. Gerade im Bereich der Anglizismenforschung sind neben den nominalen die verbalen Lehn- und Fremdwörter interessant, welche bei Übernahme der deutschen Konjugationsmuster zu einer dem ursprünglichen englischen Verb immer entfernteren morphologischen Struktur führen und auf dieses immer schwerer automatisch abzubilden sind. Ein Beispiel ist der Anglizismus *downloaden*, der allein in seiner infinitiven Form das deutsche Suffix *-en* enthält und als Partizip zu *downgeloadet* wird. Um solche Formen automatisch extrahieren zu können, müssten weitere Schritte in die Vorverarbeitung integriert werden, unter Anderem: morphologische Analyse und Wurzelmorphemerkennung (Stemming) für alle als Verben gekennzeichnete Wörter des deutschen und des englischen Korpus. Letzteres setzt voraus, dass die Korpora beider Sprachen wortartenannotiert sind.

4 Klassifikation von Anglizismen

4.1 Finden deutscher Entsprechungen von Anglizismen

Unsere Klassifizierung setzt voraus, dass die Zeitverläufe sowohl für einen Anglizismus als auch für seine deutsche Entsprechung vorliegt. Eine gewisse Menge Anglizismen konnten wir wie in Abschnitt 3 beschrieben automatisch aus großen englischen und deutschen Textkorpora extrahieren. Woher aber die deutschen Pendants nehmen? Mehrere Möglichkeiten bieten sich an, von denen sich mehrere bei näherer Betrachtung als problematisch herausstellen. Diese sollen in den Unterabschnitten 4.1.1 bis 4.1.3 kurz diskutiert werden.

4.1.1 Automatische Übersetzungen aus wortalignierten Korpora

Die Methoden der statistischen maschinellen Übersetzung basieren naturgemäß auf der Zuordnung, im Fachjargon spricht man auch von *Alignierung*, der Wörter eines Satzes der Quellsprache zu denen des entsprechenden Satzes der Zielsprache (BROWN ET AL., 1993). Bei diesem Prozess werden Übersetzungstabellen generiert, die alle möglichen Übersetzungen eines Wortes mit den dazu gehörigen Wahrscheinlichkeitswerten beinhalten. Aus zwei wortalignierten Übersetzungskorpora der Sprachen Englisch und Deutsch ließen sich theoretisch auf diese Weise deutsche Entsprechungen für Anglizismen generieren, von denen dann die jeweils wahrscheinlichsten zur weiteren Verarbeitung ausgewählt werden könnten.

Obwohl dieses Vorgehen einer empirischen Untersuchung am ehesten entsprechen würde, bringt es auch Schwierigkeiten mit sich. Die größte ist der Mangel an brauchbaren Übersetzungskorpora. Für Systeme der statistischen maschinellen Übersetzung werden zumeist die frei und elektronisch verfügbaren Protokolle von Parlamentsdebatten verwendet (KOEHN, 2005), eine Domäne, deren Vokabular aller Voraussicht nach beschränkt ist und für welche erst untersucht werden müsste, inwiefern sie Anglizismen mit einer der Alltagssprache oder des Zeitungsjargons ähnlichen Häufigkeit benutzt.

Im Rahmen dieser Hausarbeit sähen wir uns außerdem mit praktischen Schwierigkeiten konfrontiert: Um statistische Übersetzungstabellen generieren zu können, müssten die entsprechenden Korpora zunächst bereinigt und aufbereitet sowie auf Satzebene aligniert werden - ein Vorhaben, welches allein den Rahmen dieses Projekts ausfüllen würde. Das Erstellen der Übersetzungstabellen (die sogenannte *Trainingsphase*) benötigt gewöhnlich Computer mit einer Rechenleistung, die uns nicht zur Verfügung steht, und dauert im Normalfall mehrere Tage. Wir haben daher von dieser Möglichkeit abgesehen.

4.1.2 Klassische Übersetzungswörterbücher

Die Nutzung klassischer Übersetzungswörterbücher ist ein weiterer, logischer Weg, an deutsche Entsprechungen von Anglizismen zu kommen. Die wichtigste Voraussetzung für diesen Ansatz ist die Aktualität des Wörterbuchs, da Anglizismen oftmals Dinge oder Konzepte bezeichnen, die erst vor kurzem erfunden wurden oder in das gesellschaftliche Bewusstsein gelangt sind und von älteren Wörterbuchausgaben nicht abgedeckt werden. Doch selbst in einer aktuellen Ausgabe muss ein moderner Anglizismus nicht notwendiger Weise zu finden sein. Übersetzungshandbücher haben nur eine begrenzte Kapazität und beschränken sich daher auf eine bestimmte Anzahl häufiger Wörter. Anglizismen, die lediglich in bestimmten Domänen, zum Beispiel der Wirtschaft oder in der Jugendsprache eine Rolle spielen, sind in allgemeinen Wörterbüchern oft nicht zu finden.

Ein weiteres Problem bei der Verwendung von Übersetzungswörterbüchern sind die in ihnen angebotenen Übersetzungen selbst. Diese werden, sofern der Verlag die Korpusrecherche in seine lexikographische Arbeit integriert hat, nach Häufigkeit geordnet⁹. Die „wahrscheinlichste“ Übersetzung muss jedoch nicht notwendiger Weise diejenige sein, welche der Bedeutung eines Anglizismus am ehesten

⁹Die Struktur von Wörterbucheinträgen ist äußerst komplex (KUNZE/LEMNITZER, 2007). Nichtsdestotrotz sollten Übersetzungen eines Stichwortes (nach ihrer Einteilung in semantische und syntaktische Klassen) so angeboten werden, dass derjenige Übersetzungskandidat vorne steht, welcher die größte Wahrscheinlichkeit besitzt, im Alltag angetroffen zu werden. Diesem Prinzip kann man in der Praxis jedoch nicht folgen.

entspricht. Ein Beispiel soll dies deutlich machen: Das Online-Übersetzungswörterbuch Leo¹⁰ gibt als erste Übersetzung für den Anglizismus *Kick* das Wort 'Fußtritt' an. Nun ist 'Fußtritt' zwar eindeutig eine richtige Übersetzung für den englischen Begriff *kick*, die Bedeutung des Anglizismus 'Kick', wie er im Deutschen verwendet wird, entspricht jedoch im Allgemeinen eher dem Wort 'Nervenkitzel' (welches im Übrigen von Leo gar nicht angegeben wird).

Empirisch vorzugehen würde in Bezug auf unser Entsprechungsproblem bedeuten, sich darauf verlassen zu müssen, dass die dargebotenen Reihenfolgen der Übersetzungen in Wörterbüchern den Häufigkeitsverteilungen in repräsentativen und ausbalancierten Textkorpora entsprechen. Nun werden nach unserem bestem Wissen in der modernen Lexikographie zwar Stichwortlisten anhand ihrer Vorkommenshäufigkeit in Textkorpora erzeugt. Die Aufteilung der Bedeutungsinhalte bzw. der Übersetzung eines Stichwortes wird jedoch lediglich durch das Sichten von Beispielen in Textkorpora gestützt, wobei Häufigkeitsverteilungen an dieser Stelle aufgrund methodischer Einschränkungen lediglich per Augenmaß geschätzt werden können¹¹.

4.1.3 Übersetzen ohne Hilfsmittel

Eine naheliegende Alternative zu den soeben vorgestellten Möglichkeiten ist das eigenständige, auf Erfahrung gestützte, „Erdenken“ deutscher Entsprechungen von Anglizismen ohne Hilfsmittel. Dieser introspektive Ansatz würde sich in Hinblick auf die Gültigkeit gefundener Entsprechungen vielleicht als nicht unbedingt schlechter herausstellen, ist jedoch von einem empirischen Ansatz so weit entfernt wie keiner der oben genannten. Dies allein wäre eventuell keine allzu große Schwierigkeit, würden wir sowohl alle von uns extrahierten Anglizismen als auch deren Entsprechungen in den verschiedensten Bereichen des Sprachgebrauchs *kennen*. Da dies nicht so ist, bliebe uns in vielen Fällen lediglich die korpusgestützte lexikographische Recherche, eine Arbeit, deren Umfang den Zeitrahmen dieser Hausarbeit sprengen würde.

Um trotz aller angesprochenen Umstände zu brauchbaren Ergebnissen zu kommen, haben wir uns letztendlich entschieden, einen Kompromiss aus wörterbuchbasierter Übersetzung und Introspektion anzuwenden. Wir haben zunächst mit Hilfe eines klassischen (MESSINGER ET AL., 2000) und des Online-Wörterbuchs Leo Übersetzungskandidaten für alle extrahierten Anglizismen herausgesucht und diejenigen für die weitere Verarbeitung ausgewählt, welche zuerst in beiden Ressourcen vorkommen. Anschließend haben wir alle Wortpaare aussortiert, deren Entsprechungen unserer Meinung nach fehlerhaft sind. Dieses Verfahren hatte zwar den Nachteil, dass einige, eventuell wertvolle, Anglizismen von unserer Untersuchung ausgeschlossen wurden. Auf diese Weise gelangten wir jedoch in relativ kurzer Zeit zu einer brauchbaren Menge valider Wortpaare und konnten zumindest weitestgehend sicherstellen, dass durch Introspektion verursachte negative Einflüsse auf die Objektivität und Validität der Ergebnisse gering bleiben würden.

¹⁰<http://dict.leo.org>

¹¹Das Einsetzen statistischer Maße bei Übersetzungswörterbüchern würde das Verarbeiten von wort- oder phrasenali-gnierten Übersetzungskorpora voraussetzen. Einige der Schwierigkeiten hierbei haben wir in Abschnitt 4.1.1 besprochen.

4.2 Erstellung von Zeitverlaufskurven

Die am Institut für Deutsche Sprache in Mannheim (IDS) vorliegenden und über COSMAS II recherchierbaren Textkorpora¹² bieten eine Fülle unterschiedlicher Textsorten und -domänen. Nicht alle davon können für die Extraktion von Anglizismen sinnvoll verwendet werden. So reicht das "Archiv der geschriebenen Sprache" - dies ein Ausschnitt aus dem Deutschen Referenzkorpus¹³ - bis ins 18. Jahrhundert zurück. Die meisten der vor 1960 entstandenen Texte sind belletristischer Natur und für unsere Zwecke unbrauchbar, da Anglizismen vor allem in den letzten Jahrzehnten eine zunehmende Rolle spielen und schöne Literatur zeitliche Phänomene generell nicht so zuverlässig widerspiegelt, wie dies bei Zeitungstexten der Fall ist. Ab 1985 enthält das Deutsche Referenzkorpus für jeden Jahrgang Zeitungstexte in substantieller Menge, jedoch erst ab etwa 1990 mit hinreichender Streuung (Dr. Holger Keibel, IDS, persönliche Korrespondenz). Wir haben uns daher entschieden, die Datengrundlage für die Erstellung von Zeitverlaufskurven auf Zeitungstexte ab 1990 zu beschränken und entschieden uns so für die Textgrundlage, welche für die Zeitverlaufsgrafiken in den Neologismenwortartikeln¹⁴ des IDS verwendet worden sind. Dieses umfasst zum Entstehungszeitpunkt dieser Arbeit ca. 2.83 Milliarden laufende Wörter und deckt den genannte Zeitraum lückenlos ab¹⁵.

Die Erstellung einer so großen Menge von Anglizismen und ihrer deutschen Entsprechungen, wie sie für unsere Studie notwendig war, wäre äußerst umständlich manuell zu erstellen. Dies ist besonders dann der Fall, wenn die Funktionalität der erhältlichen Software eine solche Funktionalität nicht oder nur in begrenztem Umfang zur Verfügung stellt. So lassen sich die Zeitverlaufskurven für die Einträge im Neologismenwörterbuch des IDS abrufen; Dieses umfasst jedoch lediglich 790 Einträge und deckt einen Großteil der von uns extrahierten Anglizismen nicht ab. COSMAS II stellt zwar seit Kurzem die Option zur Verfügung, absolute oder relative Häufigkeiten pro Jahr (oder andere Zeiteinheit) der Ergebnisse einzelner Suchanfragen zu exportieren. Die exportierten Angaben könnten dann mit einem anderen Programm (wie Microsoft Excel) in entsprechende Diagramme umgewandelt werden. Da dies manuell und für jeden der ca. 180 Anglizismen sowie für deren deutsche Entsprechungen einzeln hätte getan werden müssen, hätte der Aufwand den zeitlichen Rahmen dieses Projekts stark belastet. Als Notlösung konnten wir glücklicherweise von laufenden Arbeiten am IDS profitieren, in deren Rahmen auch Zeitverlaufskurven zu den von uns untersuchten Wörtern erstellt und uns für diese Hausarbeit zur Verfügung gestellt wurden¹⁶. Zur Vorbereitung erstellten wir Suchanfragen in der von COSMAS II verwendeten Anfragesyntax für alle englischen und deutschen Begriffe sowie deren Flexionsformen. Da die flexionsbezogene Vielfalt bei Nomen allgemein gering ist, war dieser Schritt überschaubar. Beispielsweise mussten für den Anglizismus 'Event' folgende zwei Suchanfragen geschrieben werden: ORS('Event Events') sowie ORS('Ereignis Ereignisse Ereignisses Ereignissen').

¹²<http://www.ids-mannheim.de/cosmas2/projekt/referenz/>

¹³<http://www.ids-mannheim.de/kl/projekte/korpora/>

¹⁴<http://www.owid.de/Neologismen/>

¹⁵Die Dokumentation unter http://www.ids-mannheim.de/kl/neoplots/owid/daten_methoden.html liefert eine ausführlichere Beschreibung der Daten.

¹⁶An dieser Stelle möchten wir uns ganz herzlich bei Dr. Holger Keibel vom IDS für seine Hilfsbereitschaft und seinen Rat bedanken, ohne die wir erheblichen Schwierigkeiten gegenüber gestanden hätten.

4.3 Klassifikation anhand von Zeitverlaufskurven

Nach nochmaliger Durchsicht der Zeitverlaufskurven standen uns 81 Wortpaare für die Klassifikationsaufgabe zur Verfügung, insgesamt also 162 Zeitverlaufskurven. Diese haben wir einzeln gesichtet und allein aufgrund des Kurvenverlaufs in sechs Klassen eingeteilt, welche im Folgende beschrieben werden sollen:

4.3.1 Klasse 1 - Ablösung eines deutschen Wortes durch einen Anglizismus

Diese Klasse enthält fünf Wortpaare. Diese zeichnen sich dadurch aus, dass die Häufigkeit des Anglizismus im Laufe der Zeit zunahm, das deutsche Pendant jedoch immer weniger häufig verwendet zu werden scheint. Ein prototypisches Beispiel für diese Gruppe ist das Wortpaar *Newsletter* - 'Rundschreiben'. Wie in Abbildung 2 zu sehen ist, wurde *Newsletter* im Jahre 2008 ca. neunmal häufiger als im Jahre 1990 verwendet. Im Gegensatz dazu sank die Frequenz des Wortes 'Rundschreiben' seit dem Beginn der Datenerhebung um einen ungefähren Faktor 4. Weitere Mitglieder dieser Klasse sind die Wortpaare *Boss* - 'Chef', *Community* - 'Gemeinschaft', *Event* - 'Ereignis' sowie *Guide* - 'Führer'.

Klasse 1 entspricht der oft in den Medien angeprangerten Gefährdung deutscher Begriffe im Sinne ihrer allmählichen Ablösung durch englische Alternativen. Erstaunlicher Weise ist diese Klasse mit nur fünf Wortpaaren gleichzeitig auch die kleinste der von uns gefundenen Gruppen. Unsere Studie unterstützt demnach die These der Gefährdung durch einen empirischen Beleg derselben. Gleichzeitig relativiert sie aber auch die Ernsthaftigkeit des Problems, da vermeintlich gefährdete Wörter sich eindeutig nicht in diese Klasse einordnen lassen. Ein Vergleich mit den Angaben aus THIELE (2006) wäre hier angebracht, lässt sich jedoch an dieser Stelle nicht durchführen, da die meisten der dort aufgeführten Wortpaare entweder von unserer Extraktionsmethode nicht erfasst werden konnten oder aufgrund fehlerhafter Übersetzungen aussortiert werden mussten. Nur das Wortpaar *Event* - 'Ereignis' kommt sowohl in THIELE (2006) als auch in Klasse 1 unserer Klassifikation vor.

4.3.2 Klasse 2 - ähnliche Entwicklung von Anglizismus und Entsprechung

Die 16 Wortpaare in dieser Klasse lassen jeweils ähnliche Verlaufskurven erkennen, d.h. sowohl der Anglizismus als auch das deutsche Pendant nehmen in ihrer Häufigkeit ähnlich zu bzw. ab, wie es am Beispiel von *Coach* - 'Trainer' in Abbildung 3 deutlich wird.

Die Interpretation des Verhaltens der Mitglieder dieser Klasse ist oft nicht einfach, da in den meisten Fällen die Gründe für die Frequenzzunahme oder -abnahme nicht klar erkennbar sind. Wenige Fälle lassen immerhin Spekulationen zu. So kann man den Anstieg der Verlaufskurve für das Wortpaar *Wellness* - 'Wohlbefinden' wohl mit einem zunehmenden gesellschaftlichen Bewusstsein für ein ganzheitliches Gesundheitskonzept erklären. Ein starker Ausschlag der Kurven der Entsprechungen *Millenium* und 'Jahrtausend' um die Jahrtausendwende ist ebenfalls leicht zu erklären. Bei einem Wortpaar wie *Award* - 'Auszeichnung' hingegen können wir die zunehmende Gebrauchsfrequenz nicht gut nachvollziehen; auch gezieltere Korpusrecherchen lassen hier kaum Rückschlüsse zu.

Weitere Mitglieder dieser Klasse sind: *Banker* - 'Bankier', *Corporation* - 'Aktiengesellschaft', *Deal* - 'Handel', *Killer* - 'Mörder', *League* - 'Liga', *Logo* - 'Markenzeichen', *Millenium* - 'Jahrtausend', *Player*

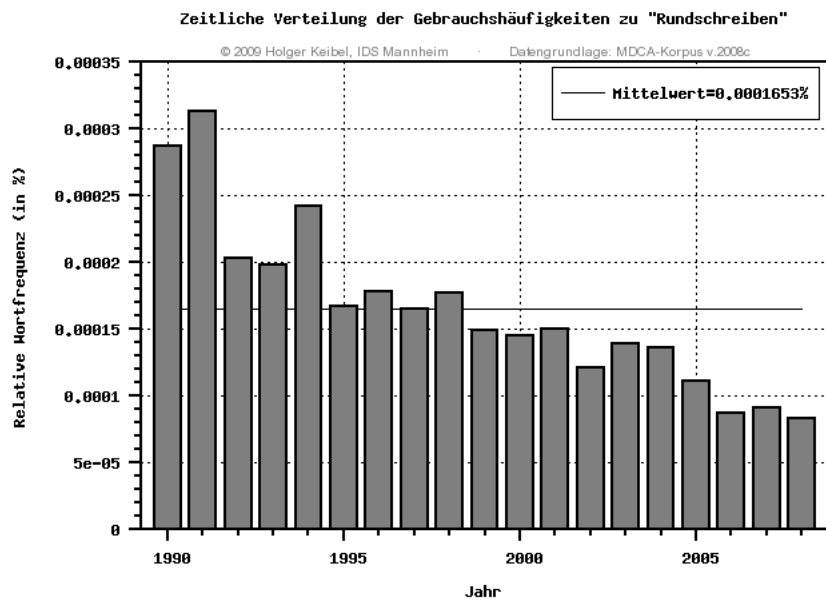
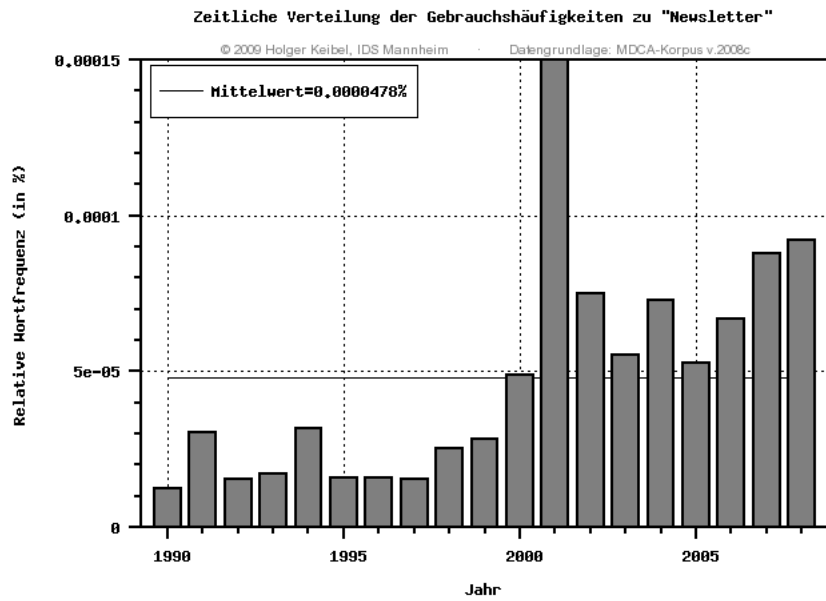


Abbildung 2: Zeitverlaufgraphiken für das Wortpaar *Newsletter - Rundschreiben*

- 'Spieler', *Report* - 'Bericht', *Song* - 'Lied', *Sound* - 'Klang', *Space* - 'Weltraum', *Ticket* - 'Karte', *Tower* - 'Turm'.

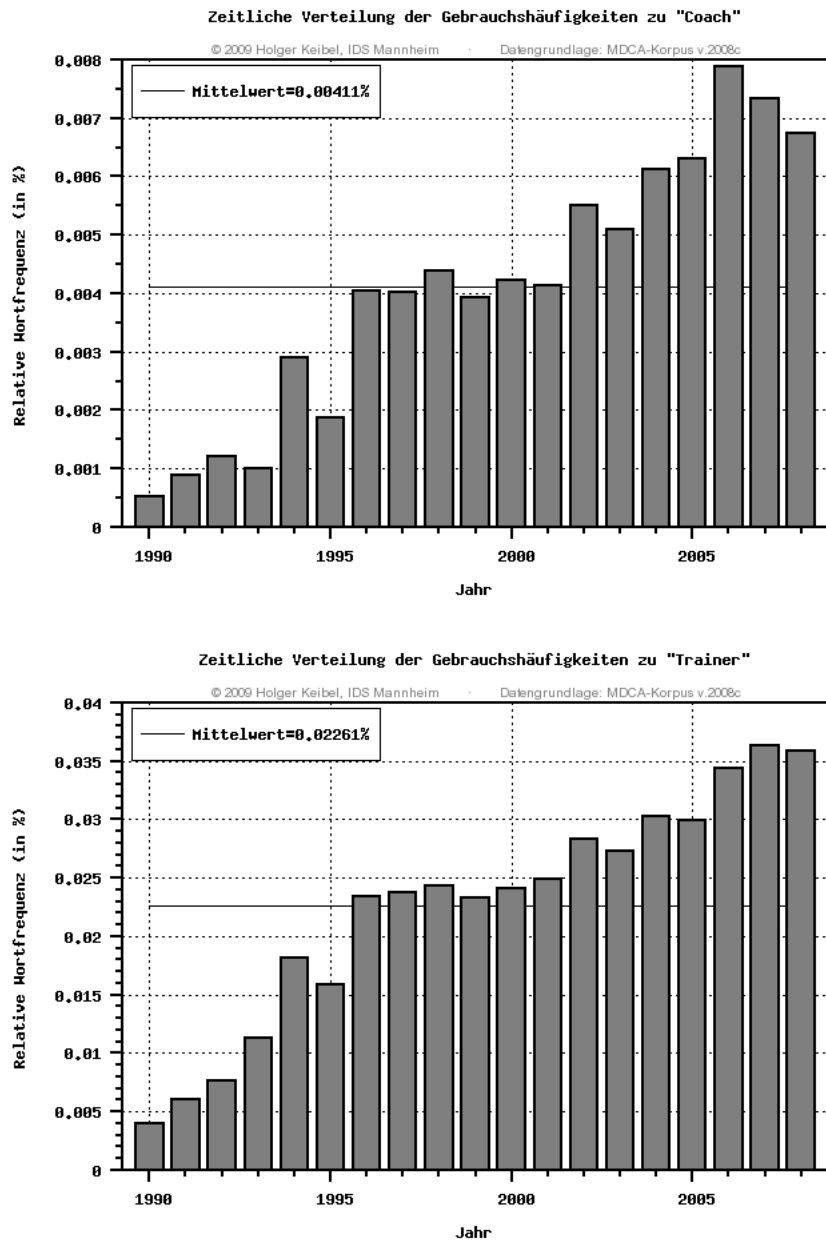


Abbildung 3: Zeitverlaufgraphiken für das Wortpaar *Coach - Trainer*

4.3.3 Klasse 3 - kein Einfluss eines Anglizismus auf seine deutsche Entsprechung

Bei dieser Klasse ist die Beobachtung, dass der Kurvenverlauf des Anglizismus kaum Einfluss auf den des deutschen Begriffs zu haben scheint. So lässt sich am Wortpaar *Mix - 'Mischung'* in Abbildung

4 erkennen, dass die Frequenz des Anglizismus zwar deutlich zunimmt, die Zeitverlaufkurve zu 'Mischung' sich davon aber unbeeindruckt gibt und einigermaßen gleichförmig um den Mittelwert 0,0031 liegt.

Weitere Mitglieder dieser Gruppe sind: *Beach* - 'Strand', *Champion* - 'Weltmeister', *Headline* - 'Schlagzeile', *History* - 'Geschichte', *King* - 'König', *Manager* - 'Verwalter', *Story* - 'Geschichte', *Vita* - 'Lebenslauf'.

4.3.4 Klasse 4 - Durchsetzen des deutschen Begriffs

Die nächste Einteilung folgt der Beobachtung, dass einige deutsche Entsprechungen sich gegenüber dem Anglizismus im Sprachgebrauch durchzusetzen scheinen. Betrachten wir die Wortpaare *Cash* - 'Bargeld', *Council* - 'Rat', *Entertainment* - 'Unterhaltung', *Foundation* - 'Stiftung', *Network* - 'Netzwerk', *News* - 'Nachrichten' und *Shopping* - 'Einkaufen', sehen wir, dass die Zeitverlaufskurven gegen Ende des Jahrzehnts ansteigen, wohingegen die der Anglizismen eher abzunehmen scheinen (siehe Abbildung 5 für das Beispielpaar *Network* - *Netzwerk*). Nicht eindeutig dieser Gruppe zuzuordnen sind die Wortpaare *Card* - 'Karte', wobei *Card* eine hohe Frequenz im Jahr 2000 aufweist, die nicht nachvollzogen werden kann, und *Trip* - 'Trip', da die 'Karte' einfach zu viele Lesarten besitzt und der *Trip* neben der Bedeutung 'Reise' ebenso einen Drogenrausch bezeichnen kann. Außerdem könnte eine hohe Frequenz ebenfalls im Jahr 2000 auf das Erscheinen des Kinofilms „Road Trip“ hinweisen. In Bezug auf die in Abschnitt 1 und 2 beschriebenen Klassifikationen ist zu sagen, dass diese Klasse in Klasse 5 (Übersetzungen) unserer ursprünglichen Klassifizierung und auf der Ebene der Entlehnungsarten den Lehnübersetzungen zugeordnet werden kann.

4.3.5 Klasse 5 - unschlüssige Kurvenverläufe

Bei den Mitgliedern dieser Klasse sind keine klaren Tendenzen zu erkennen, außer, dass die Verwendung des Anglizismus sehr unregelmäßig von Jahrzehnt zu Jahrzehnt variiert und die deutsche Entsprechung eher gleichmäßig und unverändert bleibt (siehe Abbildung 6 für das Beispielpaar *Investment* - 'Investition'). Möglicherweise kann dies wie bei Klasse 5 darauf hinweisen, dass die deutsche Entsprechung sich nicht von dem englischen Begriff hat beeinflussen lassen. Sehr ähnlich verhalten sich die Begriffe *Airline* - 'Fluggesellschaft', die besonders zur Zeit der Terroranschläge auf das World Trade Center im Jahr 2001 gleiche Verwendung zu finden scheinen. Weitere Paare wie *Code* - 'Kodex' oder *Director* - 'Direktor' verhalten sich ähnlich, sind aber aufgrund ihrer unterschiedlichen Lesarten wieder mit einem gewissen Abstand zu bewerten. Der *Code* ist vielmehr eine Verschlüsselung als ein 'Kodex', die erhöhte Frequenz im Jahr 2006 ist sicherlich auch auf den Kinofilm „The Da Vinci Code“ zurückzuführen. Der *Director* kann neben der Bedeutung des 'Direktors' auch einen Filmregisseur bezeichnen. Ausgehend von unserer ursprünglichen Klassifikation (Abschnitt 1) liegen hier wieder Übersetzungen vor. Linguistisch betrachtet ging bei *Security* - 'Sicherheit' eine Bedeutungsverengung vor sich, da *Security* in der Regel ein Sicherheitsteam bei einer Veranstaltung oder einer Diskothek bezeichnet und eher selten die Sicherheit im Allgemeinen. Ebenso bei *User*, der in erster Linie den Benutzer in Hinblick auf das Internet benennt.

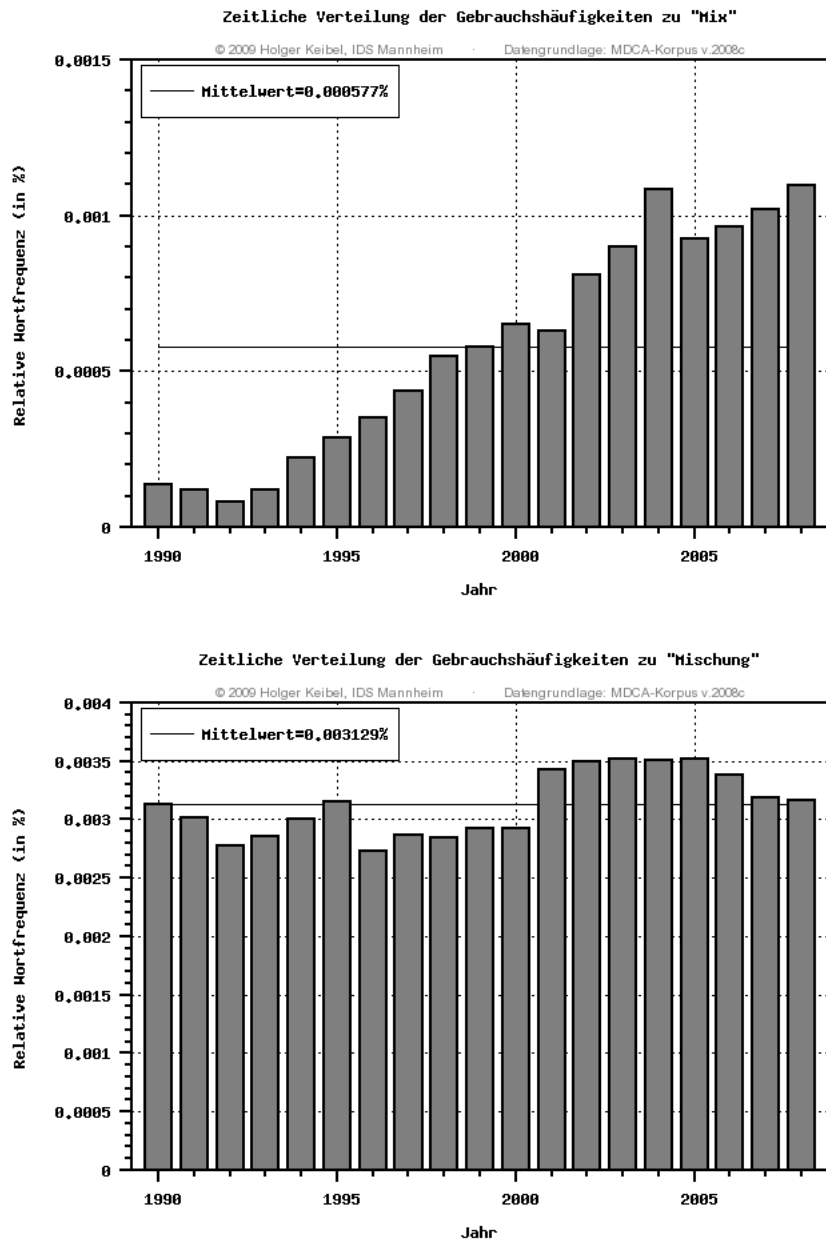


Abbildung 4: Zeitverlaufgraphiken für das Wortpaar *Mix - Mischung*

Weitere Mitglieder dieser Gruppe sind unter anderen: *Investment* - 'Investition', *Office* - 'Büro', *Security* - 'Sicherheit', *Technology* - 'Technologie', *User* - 'Benutzer' und *Trend* - 'Mode'.

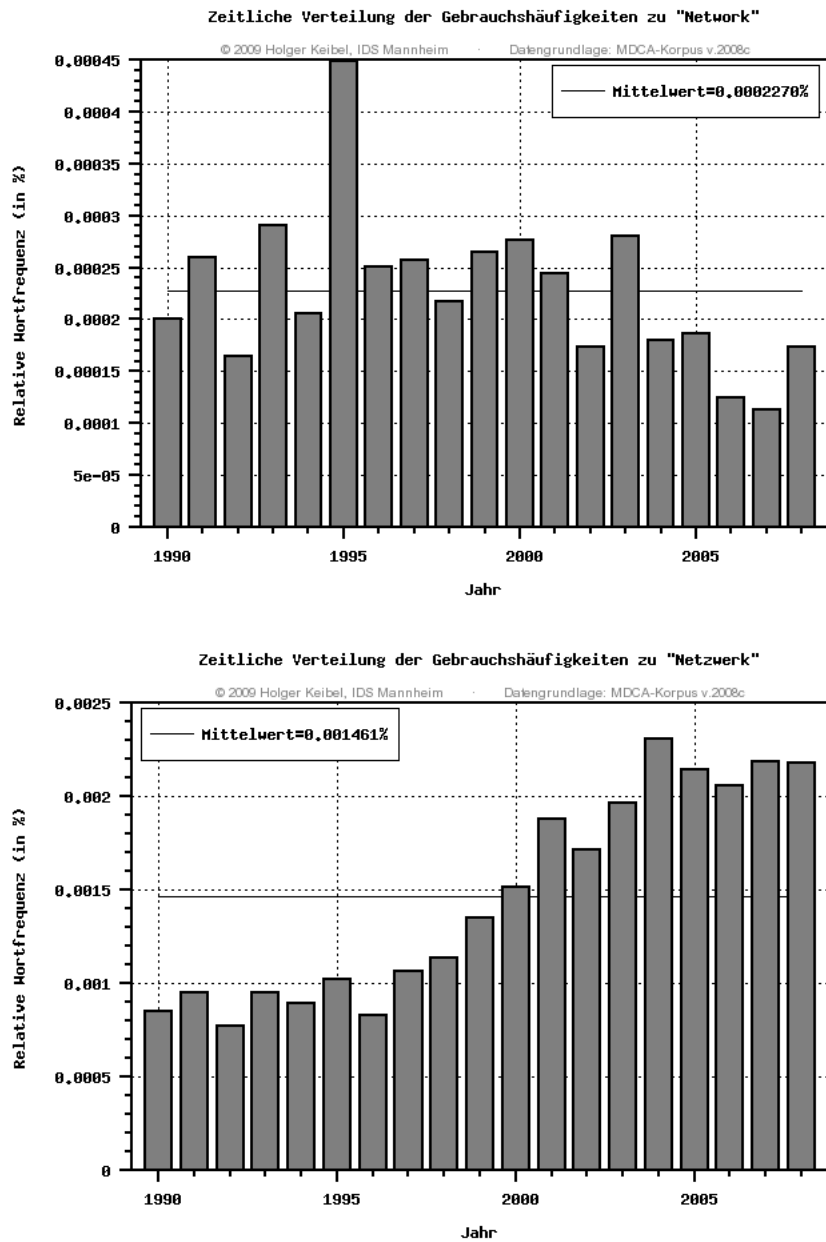


Abbildung 5: Zeitverlaufgraphiken für das Wortpaar *Network* - *Netzwerk*

4.3.6 Klasse 6 - ohne deutsche Entsprechung

Die größte Gruppe nehmen die Anglizismen ein, denen wir keine deutschen Entsprechungen zuordnen konnten. Hierzu gehören beispielsweise *CD*, *Cockpit*, *Comeback*, *Cyberspace*, *High-Tech*, *Interview*

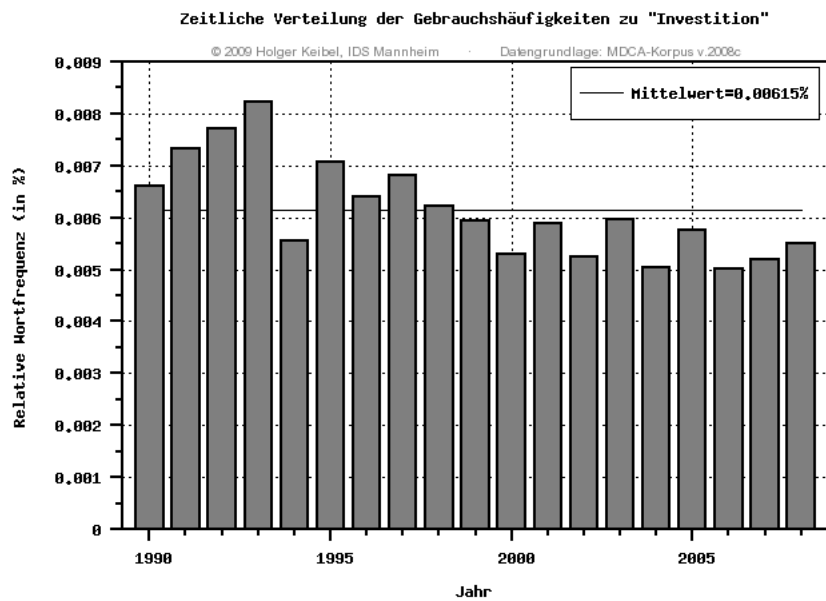
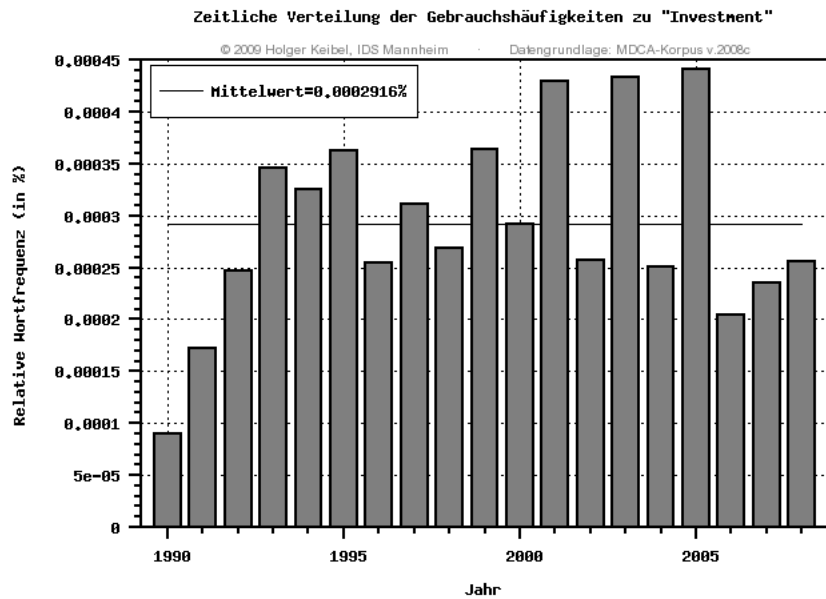


Abbildung 6: Zeitverlaufgraphiken für das Wortpaar *Investment* - *Investition*

etc., die dementsprechend zu Gruppe 3 unserer vorläufigen Klassifikation gehören. Bei *Cyberspace*, *Multimedia* und *High-tech* lässt sich sehr gut beobachten, dass die Verwendung nach dem Internetboom um das Jahr 2000 wieder stark abgenommen hat. Vermutlich hat auch die rasante Entwicklung

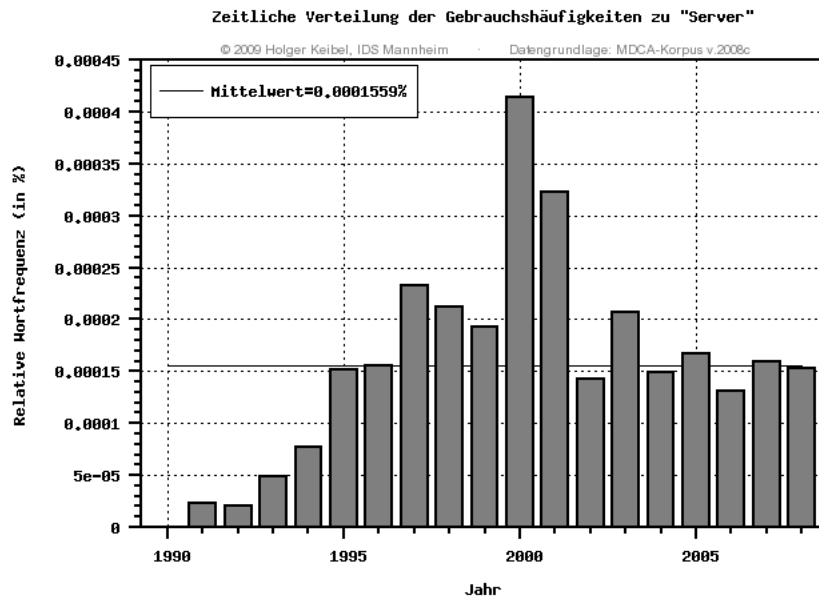


Abbildung 7: Zeitverlaufskurve für das Wort Server

der Technik der letzten beiden Jahrzehnte dazu beigetragen, dass neue Erfindungen in diesem Bereich keine besondere Erwähnung mehr finden. Nicht nachzuvollziehen fanden wir die Entwicklung von *Server*, da nach unserer Erfahrung kein derartiger Abstieg vorhanden sein sollte. Im Gegensatz dazu stehen englische Begriffe wie *Chat*, *Handy*, *Klick*, *Laptop*, *Mail*, die um die Jahrtausendwende stetig anstiegen und sich bis heute halten konnten. Wörter, zu denen es zwar deutsche Entsprechungen gibt, die jedoch zu weit von der Bedeutung des Anglizismus entfernt sind, wären: *Chat* - 'Plauderei', *Mail* - 'Postsendung', *Funk* - 'Schiss' oder *Puzzle* - 'Rätsel'. Das *Puzzle* hat eine Bedeutungsverengung erfahren, da engl. *puzzle* ein Geduldsspiel im Allgemeinen bezeichnet, die deutsche Entsprechung nur das 'Puzzle', das man aus mehreren Teilen zusammen setzt. Die Übersetzung 'Schiss' für *Funk* ist unseres Wissens nach sehr unüblich und wurde auch deshalb in diese Gruppe eingeteilt. Für *Hacker* erhielten wir die Übersetzung 'Computereindringling', die jedoch keinerlei Ergebnis bei den Zeitverlaufsanalysen lieferte und deshalb automatisch in diese Gruppe eingeteilt wurde. Daneben existieren noch Begriffe wie *Funk*, *House*, *Hip-Hop*, *Punk* und *Soul*, die im Deutschen fast nur im Bereich Musik in Erscheinung treten und keine deutsche Entsprechung in diesem Genre haben. Diese Wörter können keiner unserer Gruppen zugeteilt werden, da sie nicht in THIELE (2006) vorhanden sind. Vermeintliche Anglizismen sind *Foyer*, *Ghetto*, *Guru* und *Neon*, die jedoch etymologisch betrachtet nicht aus dem Englischen kommen und somit durch unser Untersuchungsraaster fallen.

Weitere Wortpaare und Wörter dieser Gruppe: *Budget*, *Butler*, *Cocktail*, *Computer*, *Consulting*, *Couch*, *Farm*, *Designer*, *Disco* - 'Disko', *DJ*, *Doping*, *Fan*, *Hardliner*, *Hardware* - 'Metallwaren', *Hit*, *Hobby*, *Hockey*, *Insider*, *Jeans*, *Joint*, *Know-how*, *Master* - 'Meister', *Mobile*, *Monitor*, *Monster*, *Multis*, *Musical*, *Newcomer* - 'Neuankömmling', *Partner*, *Popstar*, *Portal*, *Pullover*, *Punk*, *Ranch*, *Slogan*, *Software*,

Star - 'Stern', Surfer, T-Shirt, Teenager, Tourist.

5 Fazit

In dieser Hausarbeit haben wir einen Ansatz vorgestellt, mit welchem nominale Anglizismen anhand ihrer und der Zeitverlaufskurven ihrer deutschen Entsprechungen in distinkte Klassen eingeteilt werden können. Wir haben diese interpretiert und mit anderen, zuvor vorgestellten Klassifizierungen, verglichen. Dabei konnten wir feststellen, dass unsere Klassifizierung zwar nicht mit jenen übereinstimmt, trotzdem aber eine sinnvolle und intuitiv nachvollziehbare Gruppierung darstellt. Eine Liste nominaler Anglizismen als Ausgangspunkt unserer Studie haben wir mittels einer von uns entwickelten empirischen Methode erstellt, mit welcher es möglich war, englische Entlehnungen im Deutschen aus großen Textkorpora automatisch zu extrahieren. Diese Liste enthielt einen großen Anteil fehlerhafter Einträge, die anhand einiger Heuristiken sowie zu einem geringen Teil manuell aussortiert werden konnten.

Rückblickend können wir sagen, dass unser Extraktionsansatz zu brauchbare Ergebnissen geführt hat, auch wenn viel Spielraum für Verbesserungen existiert. So wäre denkbar, dass der Einsatz automatischer Wortklassenannotation und Lemmatisierung nicht nur die Performanz einer Extraktion nominaler Anglizismen aller Voraussicht nach erheblich verbessern würde, sondern auch die Extraktion verbaler oder adjektivischer Entlehnungen zuließe. Eine der Lehren unserer Arbeit war, dass die Domäne der zur Anglizismenextraktion verwendeten Korpora einen größeren Einfluss hat, als zunächst angenommen. Aufgrund der Beschränkung auf Zeitungskorpora konnten viele potentiell interessante und in der Alltagssprache durchaus gebrauchte Entlehnungen nicht gefunden werden. So ist uns beispielsweise bekannt, dass der Bereich des Marketing und der Werbeindustrie geradezu überfüllt ist von englischen Fachtermini. Da diese in Nachrichtentexten jedoch kaum vorkommen, konnten sie hier nicht ausgewertet werden.

Größeren Schwierigkeiten sahen wir uns beim Finden deutscher Entsprechungen gegenüber. Wir haben mehrere Möglichkeiten zur automatischen und auch manuellen Erstellung solcher diskutiert und uns schließlich für einen Weg entschieden, Introspektion und empirisches Vorgehen zu kombinieren. Dieses Verfahren erzeugte viele fehlerhafte Ergebnisse, wobei wir trotz Allem eine Menge von Wortpaaren ausreichender Größe sicherstellen konnten.

In Auswertung unserer Arbeitshypothese stellen wir fest: Die Klassifizierung nominaler Anglizismen anhand ihrer sowie der Zeitverläufe ihrer Entsprechungen ist prinzipiell möglich und führt zu sinnvollen Ergebnissen, auch wenn diese im Allgemeinen nicht denen herkömmlicher linguistischer Klassifizierungen entsprechen. Ausgehend davon bleibt zu sagen, dass man nicht von einer Bedrohung unserer Sprache durch Anglizismen ausgehen muss und der Versuch einer Klassifikation ausgewählter Nomen gezeigt hat, dass man von einem Ersatz deutscher Entsprechungen durch die englischen Wortentlehnungen nur in bestimmten Bereichen des Deutschen sprechen kann. Besonders in der Welt der Technik und der kulturellen, ständig nach Modernisierung strebenden Sektoren wie Musik und Film ist das Phänomen der Anglizismen vorrangig anzutreffen. Dabei gibt es, wie wir empirisch nachweisen konnten, tatsächlich eine Gruppe deutscher Begriffe, die offenbar von englischen Entlehnungen ersetzt werden. Diese ist jedoch verhältnismäßig klein, so dass, zumindest in Bezug auf Nomen, kaum von einer generellen Gefährdung der Integrität der deutschen Sprache gesprochen werden kann.

Eine Idee zur Erweiterung unserer Studie wäre, einen maschinellen Lernalgorithmus auf die Zeitverlaufsdaten anzuwenden. Dieser müsste so angelegt werden, dass unter bloßer Vorgabe der Anzahl der zu erwartenden Klassen Anglizismen unter Hinzuziehung der Zeitverläufe ihrer jeweiligen deutschen Entsprechungen gruppiert (Im Fachjargon wird auch von 'Clustering' gesprochen.) werden. Dieses Vorgehen würde nicht nur den Aufwand manueller Klassifizierung erheblich reduzieren und wäre somit auch auf wesentlich größere Datenmengen anwendbar. Es würde außerdem den Einfluss subjektiven Empfindens komplett eliminieren und ein deutlicheres Bild bei fragwürdigen Zeitverläufen abgeben, da Clusteringalgorithmen jeden Zeitverlauf seiner ähnlichsten Klasse zuordnen muss.

Nach Einsicht der Ergebnisse unserer Studie ist unser persönlicher Eindruck, dass die aktuelle Debatte um die vermeintliche Bedrohung der deutschen Sprache nicht nur zu großen Teilen übertrieben ist, sondern sich, gerade in den wissenschaftsfernen Medien, oftmals populistischer und unwissenschaftlicher Methoden bedient, welche vor allem auf der subjektiven Tendenz bloßen Eindrucks vermeintlicher Sprachschützer beruht. Der Versuch einer objektiven, weniger tendenziösen und auf empirischen Verfahren beruhenden Betrachtung aller Facetten sprachlicher Veränderung täte hier sicherlich gut.

Literatur

- ADLER, M. (2004): Form und Häufigkeit der Verwendung von Anglizismen in deutschen und schwedischen Massenmedien, Dissertation, Thüringer Universitäts- und Landesbibliothek Jena.
- BETZ, W. (1965): Deutsch und Lateinisch - Die Lehnbildungen der althochdeutschen Benediktinerregel, Bouvier Verlag, Bonn.
- BROWN, P./DELLA PIETRA, V./DELLA PIETRA, S./MERCER, R. (1993): The mathematics of statistical machine translation: Parameter estimation, in: Computational linguistics, 19(2), S. 263-311.
- CARSTENSEN, B./GALINSKY, H. (1975): Amerikanismen der deutschen Gegenwartssprache - Entlehnungsvorgänge und ihre stilistischen Aspekte, Winter Verlag, Heidelberg.
- GLAHN, R. (2000): Der Einfluss des Englischen auf gesprochene deutsche Gegenwartssprache: eine Analyse öffentlich gesprochener Sprache am Beispiel von Fernsehdeutsch, Dissertation, Technische Universität Darmstadt.
- GRIMM, J./GRIMM, W. (1854-1960): Deutsches Wörterbuch, Hirzel.
- KOEHN, P. (2005): Europarl: A parallel corpus for statistical machine translation, in: MT Summit, Bd. 5.
- KUNZE, C./LEMNITZER, L. (2007): Computerlexikographie. Eine Einführung.
- MESSINGER, H./TÜRCK, G./WILLMANN, H. (2000): Langenscheidts Taschenwörterbuch Englisch, Langenscheidt Verlag.
- THIELE, J. (2006): Rotbuch Deutsch - Schwarzbuch Deutsch: Die Liste der gefährdeten Wörter - Die Liste der untergegangenen Wörter, Marixverlag.

ZÜRN, A. (2001): Anglizismen im Deutschen, Dissertation, Universität Karlsruhe.