Julia Sosnizka

The Collocational Behavior of Anglicisms in German and American Business and News Magazines

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

The Collocational Behavior of Anglicisms in German and American Business and News Magazines

The Collocational Behavior of Anglicisms in German and American Business and News Magazines

Inaugural-Dissertation

zur Erlangung der Doktorwürde

der

Philosophischen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität

zu Bonn

vorgelegt von

Julia Sosnizka

aus

Geldern

Bonn 2014

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2014
Zugl.: Bonn, Univ., Diss., 2014

Gedruckt mit der Genehmigung der Philosophischen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Zusammensetzung der Prüfungskommission

Prof. Dr. Uwe Baumann
(Vorsitzender)

Prof. Dr. Jürgen Esser
(Betreuer und Gutachter)

Prof. Dr. Klaus Peter Schneider
(Gutachter)

Prof. Dr. Marion Gymnich
(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 21. März 2014

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2014
Gedruckt auf umweltfreundlichem, säurefreiem Papier aus nachhaltiger Forstwirtschaft.

ISBN 978-3-95404-712-3
eISBN 978-3-7369-4712-2

Meinen Eltern

Danksagung

An erster Stelle gilt mein persönlicher Dank meinem Doktorvater, *Prof. Dr. Jürgen Esser*, der mir die Anfertigung dieser Dissertation ermöglichte und mir über die Dauer der Erstellung größtmöglichen Freiraum ließ. Daneben danke ich *Prof. Dr. Klaus Peter Schneider* für die Erstattung des Zweitgutachtens sowie *Prof. Dr. Uwe Baumann* für die Übernahme des Vorsitzes der Prüfungskommission.

Besonderen Dank schulde ich den *Doktoranden des Forschungskolloquiums* für die kontinuierliche Lektüre einzelner Abschnitte dieser Arbeit und insbesondere *Laura Göttmann* für die zahlreichen anregenden Diskussionen und das mehrfache zeitintensive Korrekturlesen des Manuskripts.

Gewidmet ist die Arbeit meinen Eltern, *Sabine* und *Dr. Herbert Sosnizka*, mit größtem Dank für die allumfassende Unterstützung ohne die ich mein Promotionsvorhaben nicht hätte verwirklichen können.

Bonn, im Mai 2014

Julia Sosnizka

Contents

Abbreviations and symbols

ALD	Advanced Learner's Dictionary
APR	April
COBUILD	Collins Birmingham University International Language Database
DEC	December
DS	Der Spiegel
FEB	February
JAN	January
JUL	July
JS	Julia Sosnizka
LC	Lexical collocate
LDAL	Longman Dictionary of Applied Linguistics
NOV	November
NYT	The New York Times
OALD	Oxford Advanced Learner's Dictionary
SEP	September
SL	Source language
TDB	The Daily Beast
TL	Target language
T	Total
←	Indicates that the following word is derived from the one preceding the arrow
[…]	Indicates the omission of parts of a quote or a corpus example
✓	Marks collocates which prove to be identical or equivalent according to the quantitative and qualitative corpus analysis

*	Precedes incorrect phrases or words
≡	Identical collocates are marked with the mathematical ‘identical to’ sign
≙	Equivalent collocates are marked with the mathematical ‘corresponds to’ sign
x	Nodes which according to the qualitative corpus analysis do not display similar collocational behavior are marked with a small ‘x’
NODE	Nodes are given in capitalized, bold print
COLLOCATE	Collocates are given in capitalized print

List of figures

1 Introduction

In general terms, collocations can be defined as highly frequent and relatively fixed syntagmatic combinations of two or more words. The present study is concerned with the likelihood of a source language collocate to become a target language collocate. This can occur when the central word (node) of a given collocation is an English word-form which has entered the German language as an Anglicism. English, in this study, is the source language and German is the target language. The study explores whether collocations with identical nodes in English and German possess identical or equivalent collocates. This study relies on corpora of German and American business and news magazine articles for linguistic analysis. These corpora are considered representative of general and specialized journalistic writing. The two subsets of language are contrasted particularly with regard to the collocational behavior of Anglicisms.

Traditionally, a higher penetration of Anglicisms occurs in specialized text. However, countless Anglicisms are acceptable outside this category in everyday language. The present study claims that the frequent use of identical and equivalent source and target language collocates does not depend on the degree to which a subset of language is infiltrated with Anglicisms alone. In addition, other principles must apply. One of them is the 'idiom principle' which claims that a language user has available a large number of semi-preconstructed phrases that constitute single choices, i.e. words appear to be chosen in pairs or groups quite frequently (cf. Sinclair 1991: 110, 115). Therefore, the present study argues that the use of identical collocates in English and German is not a rare occurrence, but rather that the idiom principle is "far more pervasive and elusive than we have allowed so far" (ibid.: 111).

It should be noted that a study of this nature does not fit neatly into one area of linguistics. In examining collocations across languages and language varieties, the present study applies corpus-based computer linguistics. In addition, it refers to concepts and methodology specified in contrastive linguistics, e.g. *tertium comparationis* and translation studies (e.g. equivalence relations cf. chapter 3). More generally, the present study is embedded in British traditions of text analysis (cf. Stubbs 1996: 22 ff.), following the approaches of John Rupert Firth, M.A.K. Halliday, and John Sinclair in particular.

Already in 1957 Firth recognized that words combine not only according to grammatical rules, but that they also display collocational properties. He coined the term 'collocation' for the habitual or customary places of a word (cf. Firth [1951] 1957a: 181). Collocations

add decisively to the naturalness of language. Proficient native language users are intuitively aware that certain words in their language in some unspecified way tend to co-occur in relatively fixed and recurrent combinations. Consequently, collocations can also be described as psychological associations between words (cf. Hoey 2005: 5). Many collocations are so frequent that the choice of one of their constituents automatically triggers the selection of one or more other constituents in their immediate co-text. Collocations can be "evidenced by their [co-]occurrence in corpora more often than is explicable in terms of random distribution" (ibid.).

In the area of syntagmatic relations, linguistic investigations on word co-occurrences have not been very consistent. "[P]erhaps because its proper province is the rather ill-defined area of linguistic patterning that is neither clearly syntactic nor clearly semantic (Clear 1993: 271)." Collocations remain a complex, multifaceted linguistic phenomenon, which complicates the construction of distinct definitions and their thorough analysis. In recent times though, mainly due to the growing possibilities of corpus analysis, the number of investigations concentrating on word co-occurrences has increased significantly and some new and rewarding insights into their functioning have been gained (cf. for example Sinclair 1991, Kjellmer 1984, 1987, Hausmann 2004).

English functions both as a local and global medium of communication. It is an essential part of communication in multinational settings, often involving exclusively non-native speakers. The English language, therefore, can no longer be regarded as belonging to the native speakers of English as it is widely used all over the world. Over the past years, this dominance of the English language has led to an unprecedented influx of English words in other languages.

As a central phenomenon of languages in contact, 'borrowing' has secured a firm place in linguistics. Numerous studies on borrowing have been published. Many of these studies center their attention on Anglicisms and the language of the press. As a result of the influence of Anglicisms, journalists in all fields face difficulties in their practical work of writing. Some of these difficulties concern collocations which are at the heart of the present study.

Due to their extensive presence in language, collocations play a crucial role in text production. Although collocations are at first sight semantically transparent, not all of them can be translated literally. In rendering source-language collocations into any target language, a translator ideally aims at producing a collocation which is typical in the target language while, at the same time, preserving the meaning associated with the source language collocation. In the case of Anglicisms this ideal cannot be achieved, because one constituent

of the collocation (the Anglicism) is already untypical, i.e. not originating, in the target language. Source language collocation patterns which are untypical of the target language should not be carried over (cf. Baker 1992: 55). The corpus analysis will show though that this can be observed repeatedly and is usually owed to one of two strategies the journalist pursues to implement Anglicisms. Consider the following examples. The node is always identical in the source and target language.

(1)

NODE	
CASH	
COLLOCATE SL[1]	COLLOCATE TL[2]
MANAGEMENT	MANAGEMENT

(2)

NODE	
RESEARCH	
COLLOCATE SL	COLLOCATE TL
FIRM	FIRMA

When producing text, the contemporary German journalist uses Anglicisms. This can be considered state-of-the art in journalistic writing. Because of the status of the English language, journalists are most likely educated in English and have encountered the employed Anglicisms several times before in their source language (and perhaps also in the target language, i.e. the journalists mother tongue). It is presumed that consciously or unconsciously journalists will revert to previously learned source language (collocational) structures of the Anglicism. This leads either to the adoption of the entire collocation in its identical form as shown in example (1) above, or to the literal translations of the Anglicisms source language collocates as displayed in example (2) above. Either way, the influence of Anglicisms exceeds their mere word-forms found in German text.

The present study argues that replications of source language collocational structures exist within the target language, because according to Sinclair's 'idiom principle' and follow-up studies, words frequently appear to be chosen in pairs or groups. This is inherently different from strategic forms of lexical transfer such as intentional code-switching to fill a lexical gap, transfer of cognates, or the borrowing of words from another language for pragmatic purposes.

[1] SL stands for 'source language'.
[2] TL stands for 'target language'.

It is very likely that the frequencies of Anglicisms will continue to increase and that other linguistic levels outside of lexis will be increasingly affected (cf. Görlach 2002: 12). While the present study operates on the level of lexis, it attempts to illustrate the influence of Anglicisms beyond their isolated word-forms. Instead, their influence extends to the words in mediate and immediate adjacency of Anglicisms. The following hypotheses led to the conduction of the study.

1.1 Initial hypotheses

It is assumed that there exist word co-occurrences that are typical of certain types of text like business or news magazine articles in any given (source) language. It is expected that collocations are reproduced in target languages which have come under the increasing influence of Anglophone word-forms.

Hypothesis 1

The use of Anglicisms produces replications of source language collocational structures within the target language, i.e. Anglicisms feature identical or equivalent collocates in the source and target language corpora.

Hypothesis 2

Hypothesis 1 is true for specialized and general journalistic texts.

1.2 Methodological approach

The used methodology stresses the need for analyzing authentic language data and takes business and news magazine articles as a starting point of analysis. To investigate norms of use, and consequently test the hypotheses, computerized corpora of naturally occurring data are essential. Intuition, whilst being a valuable resource, is in no way sufficient to cope adequately with this task. Large collections of language data are needed, and this places the analysis beyond the level that any purely manual approach could attempt. This study therefore belongs methodologically to the area of computer-based corpus linguistics. It employs a number of additional research paradigms from the disciplines of contrastive linguistics, translation studies and communication or media studies. The goal is to bring

together perspectives and knowledge from these disciplines, to describe the collocational behavior of Anglicisms in business and news magazines.

The first stage of analysis is the creation of four research corpora. The corpora comprise a total of 9,324,491 million words. Figure 1 gives an overview of these corpora.

THE CORPORA		
LANGUAGE	**GERMAN**	**ENGLISH**
BUSINESS MAGAZINES	*WirtschaftsWoche 2008* 2,754,827 words	*BusinessWeek 2008* 1,413,686 words
NEWS MAGAZINES	*Der Spiegel 2008* 3,332,784 words	*Newsweek 2008* 1,823,194 words

Figure 1 The corpora

The corpora consist of German and American business and news magazine articles from *WirtschaftsWoche* and *BusinessWeek*, *Der Spiegel* and *Newsweek*. The 2008 volumes of these four magazines form the basis for the empirical research. The two types of magazines can be classified according to their 'origin', 'function', 'subject matter' and 'audience' (cf. section 4.1).

The corpora are representative samples of business language found in publications for knowledge dissemination and the language of general news reporting. Accepting the Firthian principle that language is varied and heterogeneous, the study differentiates the language of business magazines as 'specialized' from that of 'general' news reporting as found in news magazines.

The lexical analysis software *WordSmith Tools 5* (Scott 2008) is used to determine statistically which words significantly co-occur with previously identified key words (nodes) in the source language corpora. Because the majority of Anglicisms enter the German language without undergoing any changes in spelling, they can easily be identified afterwards in the target language corpora along with their collocates. Once a lexical base for analysis is established, more detailed analyses can be carried out.

Inherent in the methods described above are certain aspects that should be stated clearly at this stage. The present study focuses on the 'idiom principle' of Sinclair (1987, 1991) which sees language as being made up of prefabricated chunks of words. The concept of collocation is, therefore, essential for the research. Chomskyan notions of rationalist linguistic analysis that rely on intuition for the generation of data are rejected. It is not argued that intuitions have no place in such analysis, but that intuitions are often inaccurate or incomplete and that they should be firmly based on attested data (cf. Stubbs 1995: 249).

Intuition is still needed, but it is needed in the interpretation of quantitative data, not in the creation of them. In the context of this study, such interpretation is particularly useful for the analysis of equivalency relations of source and target language collocates (cf. section 3.2).

It is important to bear in mind that a corpus-based methodology also has its limitations. A rather small set of key words (cf. section 5.2) is selected to provide not only for a purely quantitative, but also a qualitative analysis. A limited in-depth examination is preferred, since it is expected to lead to a more valuable outcome and deliver more profound insights into the subject matter than an exclusively statistical approach.

In order to facilitate the reading of this work, an overview of what is found in each chapter is presented in section 1.3.

1.3 Structure of the study

This study is divided into seven chapters. Following this introductory chapter, collocation as a multifaceted phenomenon is introduced in chapter 2. The chapter provides a brief background on collocation within the broad field of syntagmatic relations. Previous descriptions of collocations and different approaches to define collocations are outlined. Subsequently, the key elements of collocation will be described. Chapter 2 concludes with an operational definition of collocation for the systematic, computer-aided identification of collocations in the corpora.

Chapter 3 introduces the linguistic disciplines and research paradigms relevant for the study and illustrates how these relate to its success. This chapter attends in particular to contrastive linguistics, translation studies and corpus linguistics and introduces terminology central to these disciplines and significant for this study.

Chapter 4 describes the four research corpora which were collected for the present study. It reports on the external criteria for corpus classification and the placement of the corpora along the general-specialized scale. Linguistic differences between business and news magazines are discussed as well. The corpora are mapped against descriptions of technical discourse and popular scientific writing. Moreover, the reasons for choosing to work with comparable corpora rather than parallel corpora are to be clarified. Finally, the procedure of compiling the corpora is illustrated and their matching criteria are presented. Chapter 4 fulfills a further important descriptive role, it describes the magazines as a mass communication medium and linguistic research object (corpus).

In chapter 5, the methods used for the empirical research are set out and further developed to meet the demands of a contrastive, corpus-based examination of authentic collocations in business and news magazines. The chapter defines ‘Anglicism’ as it is understood in this study. Furthermore the process of selecting adequate key words and the functioning of the lexical analysis software are explained.

Chapter 6 is perhaps the most important one in the study. It is dedicated exclusively to the quantitative and qualitative corpus analysis and the summary of the achieved results.

Chapter 7 concludes the study. The findings of the corpus analysis are summarized and the results of the study are mapped against the initial hypotheses. Chapter 7 ends with a proposal for further contrastive investigation of the representation of collocations in source and target languages.

2 Collocation

Collocations have captured the attention of different branches of linguistics for a long time. They hold a recognized linguistic and lexicographic status by now, but they still lack a systematic characterization and "there is no universally accepted formal definition of collocations" (Mel'čuk 1998: 23). The concept of collocation captures a range of similar phenomena. Although the term is used and understood in different ways (cf. Bahns 1993: 57), all definitions maintain a focus on the co-occurrence of words[3].

The different views on the exact structure of collocations aim at the question of how to subdivide the large group of co-occurring words in a language into smaller, ideally clear-cut categories. The present chapter illustrates collocations as defined for this study, within the large class of related structures. Despite all variations, a clear methodological grounding for the study of collocations can be offered by viewing them as an embodiment of the 'idiom principle' and an operational definition of 'collocation' can be reached. The initial working definition of collocation in this study, "words that keep company with one another" (cf. subsection 2.2.1.1), will be refined by adding more specific criteria, and the chapter ends with the definition of collocation used in this study.

2.1 Collocation as a multifaceted phenomenon

Collocations operate on the syntagmatic rather than on the paradigmatic level. The problem with syntagmatic phenomena is that they belong to several different disciplines in linguistics. Collocations may be attributed to the field of phraseology but they allow for considerable variability of the co-occurrences (cf. Burger 2007: 175), which occasionally possess idiomatic structures. Being a very complex and arbitrary phenomenon, phraseological units have not been thoroughly researched yet.

A large part of the vocabulary of a language is made up of phraseological units. Phraseological units include compound nouns (*balance sheet*), phrasal verbs (*to comply with*), idioms (*ball park figure*) and collocations (*liquid assets*). Among these, collocations are least fixed. Other phraseological units are usually of a more static nature. By assigning collocations the status of phraseological units, their collective characteristics are emphasized: all phraseological units have "idiosyncratic interpretations that cross word boundaries" (Sag et al. 2002: 2). Phraseological units are pervasive in texts of all genres

[3] Cf. Firth ([1951] 1957a: 181 and 1968: 182), Cowie (1978: 132), Hausmann (1985: 118), Cruse (1986: 40), Kjellmer (1987: 133), Sinclair (1991: 170), Sag et al. (2002: 7), Bartsch (2004: 76).

and domains (cf. Kjellmer 1987: 140) and “collocations make up the lion’s share of the phraseme inventory” (Mel’čuk 1998: 24). Some researchers even claim that most sentences contain at least one collocation (cf. Hoey 2005: 7). Seretan (2011: 2) summarizes the difficulty of defining collocations despite their high frequency as follows:

> The importance of collocations lies in their prevalence in language, whereas the difficulty in handling them comes, principally, from their ambiguous linguistic status, their equivocal position at the intersection of lexicon and grammar, and the lack of a precise and operational definition.

In general, the term ‘collocation’ refers to the linguistic phenomenon that some words occur preferably with certain others (rather than their “synonyms”). Oftentimes, the reader or listener expects the appearance of one word in the immediate vicinity of another. This is not owed to constraints on the level of syntax, but on that of usage (cf. van Roey 1990: 46). Unlike idioms, collocations have a rather transparent meaning and are easy to decode (cf. Fillmore et al. 1988). Yet they are difficult to encode since they are unpredictable for non-native speakers and, “in general, do not preserve the meaning of (all of) their components across languages” (Seretan 2011: 2).

Collocations are constrained by syntactic (grammatical), semantic and lexical properties of words. At each level, linguists have attempted to formulate rules and constraints for their co-occurrence (cf. Fellbaum 2007: 8). Within the numerous approaches three major theories can be identified (cf. also Bahns 1996, Herbst 1996 and Klotz 2000).

Firstly, one can speak of a collocation if the combined appearance of collocates is semantically inexplicable. For example, in English, you *brush your teeth*, but you do not **clean your teeth.* This notion of collocation, represented mainly by Hausmann, considers especially didactic and lexicographic aspects. During the 1980s Hausmann advocated the systematic treatment of the rediscovered concept of collocation in linguistics (cf. 1984, 1985, 1989). He defines collocations as combinations of two lexemes whose combining potential is limited by semantic rules and habitualness (cf. Hausmann 1984: 398). As constituents of collocations Hausmann admits only content words (nouns, verbs, adjectives) and disregards function words (prepositions, conjunctions, determiners etc.). In addition, the syntactic relationship between collocates is a central defining feature. Hausmann considers collocations as syntactically motivated combinations. The word-forms must be related syntactically and must be syntactically well-formed. This structural condition prevails over the proximity condition requiring them to appear within a short space of each other. In Hausmann’s view, the constituents of a collocation do not have equal status. He differentiates between ‘base’ as the dominant constituent and ‘collocate’ as the dominated

element of a collocation[4]. The relation between the two constituents of a collocation is said to be a directional rather than a mutual one (cf. Hausmann 1989: 1010). Hausmann's mainly lexicographic aim makes his notion of the directional nature of a collocation suitable for the establishment of dictionary reference structures. The 'base' is potentially listed in a collocations dictionary. A number of linguists, for example Cowie (cf. 1978: 132, also 1992, 2001) agree largely with Hausmann's notion. The inability to say why words collocate still represents a challenge today, although research has been done on this for example by Mel'čuk (1988).

A number of researchers added a statistical component to their definition of collocation as expressed for example by Kjellmer: a collocation is "a sequence of words that occurs more than once in identical form in a corpus, and which is grammatically well structured" (1987: 133). In his definition, Kjellmer includes corpora as research environments for collocations. This marks the transition to the second influential concept of collocation coined in British contextualism, the statistically and corpus oriented approach.

This concept of collocation was put forward mainly by the so-called 'Neo-Firthians' Halliday and Sinclair. It is stated that words can be regarded as collocations if they frequently co-occur within a text. This statistically oriented approach is closely linked to computer-based corpus linguistics, which has gained importance and whose potential has by far not been fully tapped.

Thirdly, collocations have also been addressed, though only to a limited extent, from the perspective of 'text cohesion'. Collocations in this approach are understood similarly as in contextualism as "the association of lexical items that regularly co-occur" (Halliday/Hasan 1976: 284). From the point of view of text cohesion, collocations contribute to the semantic unity of a text. The cohesive effect of collocations derives from the tendency of words to share the same lexical environment (ibid.: 286). This view of collocation explicitly states that collocation refers not only to pairs, but also to longer "chains of collocational cohesion" (ibid.: 287). Halliday and Hasan's concept of collocation is mentioned here for the sake of completeness, but is of secondary importance for linguistic theory (cf. Steinbügl 2005: 4) and this study.

The notion of collocation coined by Firth and his successors in the tradition of British contextualism is the most promising and useful approach for this study, but it has to be expanded by qualitative aspects in order to capture the many facets of the collocations in

[4] Hausmann introduced his definitions of 'Basis' and 'Kollokator' in 1979 and further specified their differentiation in 1989.

this study. The historical background of the developments and terminology of collocation research and the contextualist concept of collocation, leading to a corpus-oriented concept of collocation, are presented in the following section.

2.2 Brief historical background

According to some researchers, including Gitsaki (cf. 1996: 13), the concept of collocation, though not named as such, was known and described already by the ancient Greeks. Carter and McCarthy (1988: 32) point out that the term 'collocation' has been used since the eighteenth century and Bartsch (2004: 28) states that 'collocation' was first used in a clearly linguistic context in 1750.

During early linguistic research on syntagmatic relations, forerunners of the term 'collocation' were coined. In 1909, Charles Bally studied syntagmatic relations between word combinations. He used the term 'fixité variable' to describe the different degrees of fixedness of word combinations (cf. Hausmann 1979: 189). Bally differentiates between 'associations libres', 'groupements usuels' and 'unites phraséologiques'. Depending on the applied definition of collocation, their understanding today is closely related, if not equivalent, to what Bally labeled 'groupements usuels' and 'unites phraséologiques'.

During the early 1930s, Walter Porzig observed syntagmatic relations between words, which led to awareness of the phenomenon that in the use of one word another word is implicitly included (cf. Porzig [1934] 1973: 78). Porzig referred to these relationships as 'wesenhafte Bedeutungsbeziehungen' (cf. ibid.: 79) and regarded them as semantic relations. His concept of 'wesenhafte Bedeutungsbeziehungen' ultimately rests on the claim that the meaning of a word is established with reference to the syntagmatic relations it contracts with other words (similar to the later view held by Firth, Palmer and Coseriu). Whether Porzig assumed a direct syntactic relation underlying these syntagmatic relations is not clearly stated. Porzig did not explicitly set apart different types of such 'wesenhafte Bedeutungsbeziehungen', but his ideas contain some of the basic ideas of later research concerned with phraseology and collocations.

In an article in 1967, Eugenio Coseriu developed the concept of 'lexical solidarities'. Coseriu's 'lexical solidarities' can be regarded as a further development of Porzig's concept (cf. Lipka 1990: 164). Neither Porzig nor Coseriu were concerned with the nature and structural properties of the relations between the constituents of collocations (cf. Bartsch 2004: 35). As a fully formed concept, 'collocation' was established only in the twentieth century.

Harold Palmer was perhaps the first to pay attention to ‘collocation’ in the modern sense. He included over 6,000 frequent collocations (cf. Howatt 1984: 238, Nelson 2000: 159) in his teaching materials for students to memorize as one linguistic item. Palmer (1938/1968: x) defines collocation in this context as follows.

> ‘[C]ollocation’ (a succession of two or more words that may best be learnt as if it were a single word) […].

Clearly, Palmer did not limit the number of constituents of a collocation. Also, from his examples it can be seen that Palmer admitted both lexical and grammatical words as collocates (*a good many*, *make a fool of*). In an attempt to define collocations more closely, Palmer (1938/1968: xii) already differentiated ‘collocations’ from ‘phrases’, always maintaining his focus on learners of English:

> Phrases are distinguished from collocations. While collocations are comparable in meaning and function to ordinary single “words”, (and indeed are often translated by single words in the student’s mother-tongue), phrases are more in the nature of conversational formulas, sayings, proverbs, etc.

This didactic interest in collocations provided a strong motivation for their study, collection and analysis from the perspective of (foreign) language teaching. In the first edition of the Advanced Learner’s Dictionary, A.S. Hornby included collocational information (cf. Seretan 2011: 8) aimed at foreign learners of English. This pedagogical trend in the research of collocations was later continued, most notably, by Anthony Cowie and Peter Howarth. Thus, collocations unveiled largely from pedagogical observations on language acquisition that associated them with a high level of proficiency, which could only be achieved by speakers through memorization (cf. ibid.: 9). Until today, this interest remains. Hornby’s dictionaries continue to be a great commercial success (cf. preface OALD) and collocation dictionaries are compiled for many languages.

Bally, Porzig, Palmer and Hornby had already conducted research on habitual word co-occurrences and the term ‘collocation’ was already in use in linguistics. But it was John Rupert Firth who established collocation as a central concept of his theory. Firth popularized the term ‘collocation’, derived from the Latin word *collocare* - to place together, to assemble (cf. Seretan 2011: 9).

2.2.1 The contextualist concept

'Contextualism' developed as a theory of language in direct opposition to the decontextualized observation of language within American (Bloomfield) and European (Saussure) structuralism. Contextualists argue that language should be studied in authentic instances of use, not as intuitive, invented, isolated sentences. The unit of study must be whole texts and the study of 'context' should be central in linguistics (cf. Firth 1968: 174/179). Firth, the founder of British contextualism, worked in the structuralist tradition that was prevalent in his time. His ideas were strongly influenced by those of anthropologist Bronisław Malinowski (cf. Steiner 1983: 96). The influence of Malinowski's work shaped Firth's conviction that language should be studied as a social and cultural phenomenon by regarding its 'context of situation' beyond the purely linguistic facts (cf. Firth 1957c, Robins [1971] 2004: 33). This has been attempted variously in recent years, but has yet to be fully integrated into a comprehensive linguistic theory (cf. Bartsch 2004: 31). Firth's further developments of Malinowski's (1923) concepts of 'context of situation' and 'context of culture' form the basis of a significant part of his theory of language (cf. Robins [1971] 2004: 33).

Firth argued that the meaning of a word derives just as much from the particular situation in which it occurs as from the syntactic or syntagmatic relations it enters. This idea, which mixes language with the objects physically present during a conversation to ascertain the meaning involved, is known as Firth's 'contextual theory of meaning' or his theory of 'context of situation'. According to Violi (2000: 103), a syntagmatic and pragmatic view can be distinguished.

> [In contextualism] a syntactic or syntagmatic approach is adopted when considering the meaning of a linguistic (or other) sign to be a function of its relation to other linguistic (or other) signs in its context, and a pragmatic approach is adopted when meaning is defined as a function of its situational context.

Some of Firth's ideas on meaning were developed in his article "The Technique of Semantics" (1935). This article marks the beginning of contextualism, which reached its peak during the late 1930s and the first decade after World War II. Firth emphasizes both the relational and the situational context. He also recognizes contextual relations at all levels, phonology, grammar, or lexicography as manifestations of meaning (cf. also [1951] 1957a).

Contextualist descriptions of language regard linguistics essentially as a social science. The social context of the linguistic code is the culture - seen as a network of information systems, and the social context of language behavior is the situation in which socio-cultural meanings are exchanged by means of, amongst other things, the linguistic code (cf. Halliday 1984). Finally, this socio-cultural perception of meaning is connected to pragmatic approaches to semantics (especially those of Wittgenstein 1953: 80) and may be best summarized in Halliday's (1984: 22) words:

> Context is in this kind of model a construct of cultural meanings, realised functionally in the form of acts of meaning in the various semiotic modes, of which language is one. The ongoing processes of linguistic choice, whereby a speaker is selecting within the resources of the linguistic system, are effectively cultural choices, and acts of meaning are cultural acts.

2.2.1.1 Firth

'Collocation' and 'collocability' were introduced to the academic discussion by Firth ([1951] 1957a: 194), who mentioned these terms in his essay "Modes of Meaning" for the first time:

> I propose to bring forward as a technical term, meaning by 'collocation', and to apply the test of 'collocability'.

Firth developed his linguistic models based on the notion of 'meaning by collocation'. Contextualists assume that in characterizing a word, its context plays the most important role: "You shall know a word by the company it keeps!" (ibid.: 179). While this characterization provides a good understanding of the concept of collocation, it remains quite vague, as nothing is said about its linguistic status and properties.

In contextualism, the concept of collocation plays a central role; collocating words define each other. In particular, contextualists argue that the meaning of words is defined by their co-occurrence (or collocation) with other words (cf. Seretan 2011: 16). Firth discusses 'meaning by collocation', which he defines as an "abstraction on the syntagmatic level […] not directly concerned with the conceptual idea approach to the meaning of the words" (Firth [1951] 1957a: 196). Thus, part of a meaning of a word is the fact that it collocates with another word. The words with which it collocates, however, are often strictly limited.

Firth, like Palmer, illustrated collocations mainly by means of examples and did not develop a clearly outlined concept in his essay. This vagueness in definitions led to several

interpretations which vary depending on the basic assumptions of the respective branch of linguistics. The lack of specificity concerning his elaborations on collocation remains a major point of criticism (cf. Robins 1961: 198, Steiner 1983: 129, Kohn 1992: 370, Herbst 1996: 380, Lehr 1996: 21).

In his later article "A Synopsis of Linguistic Theory" Firth (1957b: 11 f.) attempts a more precise definition of collocation:

> [T]he collocation of a word […] is not to be regarded as mere juxtaposition, it is an order of *mutual expectancy*. The words are mutually expectant and mutually prehended. (emphasis added JS)

And

> [T]he habitual collocations in which words under study appear are quite simply the mere word accompaniment, the other word-material in which they are *most commonly* or *most characteristically* embedded. (emphasis added JS)

From these quotes it can be concluded that Firth's understanding of collocation puts emphasis on the criteria of mutual expectancy and habitual co-occurrence, which manifests itself in frequency and (statistical) probability (cf. also Lehr 1996: 24). These criteria for assigning the status of collocation will be discussed in detail in section 2.3.

The description of collocations within the framework of contextualism passed through several stages. The revolutionary part of Firth's thinking was to look at lexical relationships at a syntagmatic rather than paradigmatic level, whereas previous grammars had only considered structural relations on the paradigmatic level (cf. Gitsaki 1996: 1). Firthian linguistics was committed to the study of syntagmatic relations between linguistic items and the interdependency between grammar and lexicon. Firth's ideas were taken up and further developed by Halliday and Sinclair in articles that have been since regarded as landmark.

2.2.1.2 Halliday

Halliday (1966) reiterated Firth's idea that part of the meaning of a word is the fact that it collocates with others. Contrary to Firth, who was essentially concerned with syntagmatic relations, Halliday points out that those syntagmatic relations are in many cases reflected on the paradigmatic level. He introduces the terms 'lexical series' and 'lexical set' (cf. ibid.: 150-7, Halliday et al. 1964: 33). Lexical series (*oaktree, ashtree, plantree, beechtree)* consist of compound items having one constituent in common which is "morphologically

unmarked" (*tree)*. A 'lexical series', may or may not coincide with the grouping recognized as a 'lexical set' (Halliday 1966: 157).

Halliday defines a 'lexical set' as a group of lexical items which possess a similar 'range of collocation'. In contrast to the members of a 'lexical series', members of a 'lexical set' do not have compounds with a common constituent. The members of a lexical set display similar collocational behavior within a certain field, i.e. "they have a number of highly probable collocations in common […]" (Halliday et al. 1964: 33 ff.). Halliday (1966: 153) summarizes this in the following quotation:

> Collocational and lexical set are mutually defining as are structure and system: the set is the grouping of members with like privilege of occurrence in collocation.

For example *deal*, *agreement* and *contract* belong to a 'lexical set'; they collocate readily with *international* or *fixed* (cf. *BusinessWeek 2008*: APR). Halliday demonstrates this with *chair, seat* and *settee or bright, shine* and *light* which belong to one lexical set.

Yet, Halliday acknowledges that the paradigmatic relations between the elements of a lexical set depend on the syntagmatic relations they enter in particular cases. A lexical set is therefore without consistency (cf. 1966: 150 f.). He offers *strong* and *powerful* as examples. These adjectives belong to the same lexical set, because both readily collocate with *argument*. On the other hand, each adjective contracts syntagmatic relations in which only one of two is acceptable – *strong tea/*powerful tea*.

Halliday (ibid.) also illustrates that collocations cut across grammatical boundaries and occur in different syntactic constructions (*argued strongly/strength of argument*). He describes these as "instances of one and the same syntagmatic relation" (ibid.). Halliday argues further that the constituents of a collocation do not necessarily have to enter a grammatical (syntactic) relation (ibid.) – *I wasn't altogether convinced by his argument. He had some strong points but they could all be met.* This is in line with more recent research on collocations, which states that sometimes collocation stretches across sentence boundaries (cf. Schenk 1994: 6). As mentioned earlier in section 2.1, in his later works Halliday (1976) focuses on the cohesive effect of collocations, where this is of greater importance.

Halliday presents a far more detailed concept of collocation than Firth did. In essence, Halliday sustains Firth's ideas and affirms that "lexical choice […] is different from grammatical choice" (1964: 34). Regularities belong to the area of grammar, while co-occurrences of words are not predictable or explicable in most cases; therefore such co-occurrences escape the rules and cannot be described systematically. For Halliday 'collocation' illustrates syntagmatic relations and 'lexical set' paradigmatic ones. Firth and

Halliday agree that different branches of linguistics would particularly benefit from research and insights into collocations: lexicography, stylistics and didactics (cf. also Konecny 2010: 43 f.).

2.2.1.3 Sinclair

Collocation is also covered in depth by Sinclair in his article from 1966. Sinclair's approach to collocation continues the tradition of British contextualism. Grammar and lexis are still regarded as different aspects of language (cf. Sinclair 1966: 411, Jones/Sinclair 1974: 15 f.). Sinclair did not always see grammar and lexis as inseparable and it is interesting to note that his views have changed since the 1966 article, where he still kept grammar and lexis apart (cf. Nelson 2000: 160). As opposed to Halliday, who concentrated mainly on developing a detailed theory of grammar (cf. Halliday 1985, 1989), Sinclair focuses rather on lexis. Comparable to Halliday, the criterion of statistical probability is central to Sinclair's concept of collocation.

Sinclair defined such terms as 'node', 'collocate', 'cluster' and 'span' (cf. 1966: 415, 425, Jones/Sinclair 1974: 16, 21, 1991: 115) as they are used today and produced a frequency list of collocates. Sinclair refers to a word whose collocational behavior is explored as 'node'. A given number of words on each side of a node are described as 'span'. Words within this span are potential 'collocates' of the node. The term 'cluster' comprises collocates that show a high frequency in turning up for a particular node (cf. 1966: 417).

Sinclair found it difficult to decipher the maximum span (cf. 1966: 414) within which two words could still be considered node and collocate. The extent of the span was to be fixed at the optimum value at a later stage (cf. Sinclair 1991: 106). In his definitions of collocation, Sinclair occasionally opted for slightly vague paraphrases of span: "Collocation is the co-occurrence of two items in a text within a *specified environment* […]" (cf. Jones/Sinclair 1974: 19; emphasis added JS). In his earlier works (1966) he had set the span to ±3, later (1974) he researches collocates within a span of ±4. According to Lehr (1996: 38) the accurate size of the collocational span remains unsolved until today. In connection with his queries about the ideal collocational span, Sinclair (1966: 414) especially notes that some highly frequent collocates typically do not occur in direct adjacency to the node. He concludes that the proximity of node and collocate does not offer solid evidence of the collocational strength and that a larger span is not necessarily attended by a lower degree of mutual expectancy (cf. subsection 2.3.4).

Sinclair distinguishes between ‘upward’ and ‘downward’ collocations. If the frequency of a node in a text is higher than that of its collocate this is a matter of ‘downward collocation’. Conversely he speaks of ‘upward collocation’ when the node is less frequent than the collocate it co-occurs with (cf. Sinclair 1991: 116). Sinclair notes a systematic difference between upward and downward collocation (ibid.). Upward collocation is the weaker pattern in statistical terms, and collocates tend to be grammatical elements, for example prepositions, conjunctions and pronouns. In contrast, downward collocations consist of a large number of nouns and verbs.

Furthermore, Sinclair differentiates between ‘casual collocations’ and ‘significant collocations’ (cf. 1966: 418, Jones/Sinclair 1974: 19) on the basis of their frequency. Sinclair, unfortunately, does not make explicit where exactly the line between casual and significant collocations should be drawn. Nevertheless, Sinclair’s framework constitutes part of the foundation of corpus-based research on collocations (cf. subsection 2.2.2).

Closely related to his research into collocations, Sinclair suggests that language obeys two opposed principles. Krishnamurthy (2000) reminds us that Firth had already noted that collocation is an abstraction on the syntagmatic level. Sinclair proposes two models to account for syntagmatic relations “in order to explain the way in which meaning arises from language text” (1991: 109). The ‘open-choice principle’ refers to the regular choices in language and accounts for the utterances produced by the application of grammatical rules. The ‘idiom principle’ stipulates that these regular choices are further restricted by the presence of prefabricated phrases that are already available to speakers (cf. ibid.: 110 ff.). The introduction of the ‘idiom principle’ finally offers a methodological and theoretical grounding for collocation (cf. subsection 2.2.1.3.2 below). The principles are not compatible as “the switch from one model to the other will be sharp” (ibid.: 114). However, existing in parallel, both are applied when it comes to text production (cf. 1987: 324, 1991: 114). Sinclair mentions the open-choice and the idiom principle for the first time in his 1987 article “Collocation: a progress report” and states that collocation illustrates the idiom principle.

2.2.1.3.1 The open-choice principle

The open-choice principle is also referred to as a ‘slot and filler’ model, which defines text as a series of syntactic slots that may be filled with practically any word from a lexicon. Text is then seen as the result of a very large number of complex choices taken at the same time on all levels of language. At each point where a unit is completed (a word, phrase or

a clause) a large range of choices opens up; the only restraint being grammaticalness (cf. Sinclair 1987: 320, 1991: 109). The open-choice principle can be imagined as an analytical process which goes on all the time but whose results are only intermittently called for - if the idiom principle does not lead to satisfactory results. Lexical choices which are unexpected in their environment authenticate the application of the open-choice principle. Virtually all grammars are constructed on the open-choice principle (cf. Sinclair 1987: 320).

Sinclair regards this model of language, which distinguishes between grammar and lexis and uses grammar to provide a string of lexical choice points, as a secondary model (cf. 1991: 114). Normal text would not be produced simply by operating the open-choice principle. He states that text in general shows a potential for being analyzed as the result of open choices, but that the idiom principle dominates (ibid.). Words do not simply occur at random in a text and the open-choice principle does not account for sufficient restraints on consecutive choices. This is where the idiom principle applies.

2.2.1.3.2 The idiom principle

The idiom principle assumes that communicative performance leans on stored prefabricated sequences of words and that therefore the use of ‘rules’ of a ‘system’ is to a certain extent not very operative. Language is treated as many sets of units of meaning which may consist of more than one word.

Collocation, like any other linguistic phenomenon, depends on the developments of the extra-linguistic world and its conceptual categorization. Items that occur together physically, much like concepts from certain areas of expertise, are likely to be mentioned together. Nevertheless choices in language are not influenced merely by external factors or organizing features. The choice of register and the pursuit of particular purposes in determined social settings automatically narrow down slot-by-slot preferences (cf. Sinclair 1991: 110) as well. This opposes the open-choice principle since one cannot choose simply any word from a lexicon to fill a slot, although multi-word units allow for some internal lexical variation and/or changes in word order. In sum, many uses of words and phrases attract other words in strong collocation (cf. ibid.: 111).

Collocation illustrates the idiom principle since “the choice of one word affects the choice of others in its vicinity” (ibid.: 173). If words collocate significantly with each other, then this collocation is the result of a single choice (cf. Sinclair 1987: 323). Words appear to be

chosen in pairs or groups quite frequently (cf. Sinclair 1991: 115). Sinclair pays less attention to the distance between constituents of a collocation than others and he explains that "on some occasions, words appear to be chosen in pairs or groups and these are *not necessarily* adjacent" (ibid.: 115; emphasis added JS). As long as they regularly co-occur within the determined collocational span, they embody the idiom principle (cf. Sinclair 1987: 325 f.).

Erman and Warren (2000) attempt to quantify proportions between the two models in authentic texts. As much as 55% are said to belong to the idiom principle, whereas only 45% agreed to the open-choice principle. Partington (2004: 131) goes as far as to propose that there are probably no words that do not have specific lexical preferences. Treating collocation similarly to a single word or lexical item[5] also raises questions about the impact collocations have on the structure of language and especially that of vocabulary. Sinclair states that most text is made up of the occurrence of frequent words and of the frequent senses of less frequent words. Hence normal text is largely delexicalized and appears to be formed by exercise of the idiom principle, with occasional switching to the open-choice principle (cf. Sinclair 1987: 323). Grammar, as a linguistic level, is almost exclusively subject to application of the open-choice principle, whereas on the level of lexis, the idiom principle dominates (cf. Konecny 2010: 49). Thus, lexical choices which are unexpected in their environment will presumably occasion a switch. Choices which, if grammatically interpreted, would be unusual are an affirmation of the operation of the idiom principle (cf. Sinclair 1987: 324).

Sinclair's (1966) and Halliday's (1966) articles stress the need for computer-based corpus linguistics, and, in doing so, they were ahead of their time. Assumptions that had to be based on scattered observations of language in Firth's time can now be tested against large-scale corpora of different origin. In his late works, Firth himself estimates this development: "The use of machines in linguistic analysis is now established" (Firth 1957b: 31). A concept of collocation based on contextualist descriptions and adapted to computer-based analysis has been applied by numerous linguists (for an overview cf. Konecny 2010: 53) and is proposed for the purpose of the present study.

[5] The term 'lexical item' is used to refer to a unit of description made up of more than one word-form and is also available for units with an internal structure, like compounds (cf. Sinclair 1998: 23).

2.2.2 The corpus-oriented concept

The more recent developments in the study of collocations would have been largely impossible without the use of electronic corpora and computers. Once the use of large computerized corpora had become commonplace, collocations were modeled using the statistical notion of 'significance'. Collocations were defined as those combinations whose probability of co-occurrence, estimated on the basis of their co-occurrence frequency observed in a corpus, is much larger than chance (cf. Church/Hanks 1990: 23). Sinclair is the first to assign the status of collocation to words co-occurring in a text and differentiates between 'casual' and 'significant' collocation based on above-average frequency (cf. subsection 2.2.1.3). In many cases, research into collocations is based on this approach which was developed in contextualism. Concepts developed in the contextualist approach like 'collocate', 'node' and 'span' have been implemented in the corpus-based analysis of language.

In contextualist and corpus-oriented approaches, collocations consist of words highly probable to co-occur and presume statistical analysis of texts. A semantic predominance of one of the collocational elements is generally not assumed in contextualism. This opposes Hausmann's view (cf. section 2.1) of a directional relation between constituents of a collocation (cf. 1979, 1989: 1010). Sinclair distinguishes 'upward collocations' and 'downward collocations'. These, however, are frequency-based and not semantic concepts. In general, contextualists regard collocations as symmetrical relations, paying no attention to the relative importance of the words involved. To emphasize this, Sinclair (1991: 173) notes that "collocation is one of the patterns of mutual choice." There are no restrictions on word classes or syntactic structures. As Sinclair (ibid.: 179) puts it, collocation refers to "lexical co-occurrence, more or less independently of grammatical pattern or positional relationship." Collocates need not be adjacent to the node. Collocations are not based on semantics as a distinctive characteristic. This statistical view is predominant in the work of Halliday and Sinclair. Several linguists with an interest in corpus linguistics, computer linguistics, or corpus-based lexicography, have published in this tradition since (cf. for example Kjellmer[6] 1984, 1987, 1990, 1991, Church/Hanks 1990, Stubbs 1993, 1995, 1996, Esser 1999).

Since its introduction into linguistics by Firth, collocation has been used in many different ways, but in its original sense it plainly denotes the company a word keeps, ignoring any

[6] Kjellmer is mainly interested in analyzing different text types and characterizes them according to concentration of collocations.

semantic or syntactic reasons for this company (cf. also Handl 2005: 25). In other approaches (cf. section 2.1), the syntactic relationship between collocates is a central defining feature. These views are more restricted and emphasize that collocations are semantically motivated, syntactically well-formed constructions (cf. Cowie 1978: 132, Hausmann 1989: 1010, Bartsch 2004: 76, Mel'čuk 1998/2003). Hausmann remains the leading critic of the contextualist concept of collocation (cf. 2004: 320 ff.).

A corpus-based identification of collocation also has its drawbacks (cf. Krenn 2000: 3). At present, computers cannot tell the difference between types of phraseological units. This may cause compound nouns, phrasal verbs, or idioms to be identified as collocations. On the other hand, typical collocations which are not highly frequent or present at all in a given corpus cannot be captured based on a statistical definition. Recently, it has been proposed to use the term 'association' or 'co-occurrence' for the general statistical understanding of collocation and to reassert the term 'collocation' for the restricted understanding corresponding to the 'linguistically-motivated approach' (cf. Seretan 2011: 14). This view is rejected in this study.

The above is the framework within which the role of collocation is considered. Collocation, as has been mentioned, illustrates the idiom principle. Key elements which are decisive for assigning the status of collocation in this study are reviewed below.

2.3 Key elements of collocation

2.3.1 Co-occurrence and context

A collocation consists "[…] of two or more words […]" (Sinclair 1991: 115). This seems to be an obvious and unquestionable condition. Nevertheless, even at this point researchers disagree slightly. Despite the fact that the practical work is concerned almost exclusively with collocations made up of exactly two lexical items, in theory there is no length limitation for collocations (cf. Sinclair 1991: 170, Halliday/Hasan 1976: 287). Complex collocations like *major turning point* (*Newsweek 2008*: FEB) can be found often in language. In the literature, collocations consisting of three or more words are also often referred to as 'multi-word units' or 'lexical bundles'. They are extended collocations statistically co-occurring regardless of their idiomaticity (cf. Biber et al. 1998, 1999, 2004). Their frequency is seen as a reflection of the extent to which a multi-word combination is stored and used as a prefabricated chunk (cf. Biber et al. 2004: 376). Thus, they also exemplify the functioning of the idiom principle.

Even as the current study is concerned solely with combinations on the word level, Firth himself used ‘collocation’ to refer to the co-occurrence of items at all grammatical levels, not just the word level (cf. Partington 1998: 16). Some linguists study collocations below word level (cf. van der Wouden 1997: 20 f.). This is sometimes referred to as ‘morphological collocation’ and may refer to a particularly common use of a given prefix. However, co-occurrences below the word are not usually considered to be collocations (cf. Tognini-Bonelli 2001: 105).

Many word-forms show a tendency to co-occur with certain grammatical choices. Therefore, the distinction between grammatical and lexical words is relevant to the treatment of collocations and this study. There are various studies dealing with either lexical or grammatical collocates or both types. Firth ([1951] 1957a) already noticed that collocates can be both grammatical and lexical (cf. also Sinclair 1991: 116). Today, there is general agreement in the literature on the division of collocates into lexical or grammatical categories, though there is less agreement on their relative importance. Lexical collocation is defined by Hill (cf. 2000: 50 ff., also Lewis 2000, Lewis/Hill 1998) as having five main categories (*adjective/noun, verb/noun, noun/verb, adverb/adjective and verb/adverb*). Others have defined up to 37 categories of collocation, eight of which could be considered as lexical collocation (cf. Gitsaki 1996: 23), consisting of various combinations of nouns, adjectives, verbs and adverbs.

Firth ([1951] 1957a) called the syntactic (grammatical) constraints on a word’s selection of neighboring words ‘colligation’, a term adopted by Sinclair (1996, 1998, Sinclair et al. 2004). Hoey (2005) uses ‘colligation’ as a term for statistically significant co-occurrences of a lexical word with grammatical words. Examples are **MANAGER**, **SHARES** and **ASSETS**, which tend to be followed by OF. **BANK** in many cases is preceded by THE and followed also by OF (cf. *BusinessWeek 2008*: JAN-DEC). Kjellmer (1987, 1990), has looked at the collocation of grammatical classes, a phenomenon *also* known as ‘colligation’ such as, for example, the frequent co-occurrence of ‘determiner+noun+preposition’. Accordingly, the above collocation THE **BANK** OF, for example, could be subsumed under this more abstract notion of colligation. The present study only deals with lexical words as collocates (cf. section 2.4). Thus, it disregards grammatical words and grammatical classes as collocates in the sense of the definitions of colligation above. The corpus analysis will discuss which word-classes the identified collocates represent.

There is a further point in Sinclair’s definition that needs to be made explicit. Sinclair uses in his definition of collocation the informal concept of ‘word’, which in linguistics has many interpretations. First but foremost, collocations consist of ‘word-forms’, not of

'words'. A layperson typically means by 'word' an orthographic word, for example, "[…] the material between two blanks in a text" (Handl 2005: 29). Unfortunately, this definition fails sometimes. Consider for example compounds like *news story* or the adjective *long term*. For the latter, the English spelling conventions even lead to different realizations: "long-term"/ "long term" (*BusinessWeek 2008*: APR). Despite this separation by a blank or a hyphen in the text, the status of compounds and compound adjectives as one 'word' is not diminished. Accordingly, collocates can of course also be compounds.

Another closely related difficulty which arises during corpus analysis is the fact that constituents of compounds can erroneously be identified as collocates during the quantitative corpus analysis. Consider the following example.

(3)

NODE	
BANKING	
COLLOCATE SL	**COLLOCATE TL**
INVESTMENT	INVESTMENT

For the node **BANKING** the identical collocate INVESTMENT is frequently detected in the source and target language corpora. However, INVESTMENT cannot be counted as a lexical collocate of **BANKING** because INVESTMENT and **BANKING** are constituents of the compound INVESTMENT **BANKING** (cf. subsection 6.2.2) and hence form a single 'word' in the sense of Sinclair's definition. Moreover, to avoid ambiguities, the study uses the term 'word-form' instead of 'word'.

Another important condition for their treatment as collocation is that the items in question must co-occur together in context. 'Context' here purposely denotes both the direct textual (also referred to as 'co-text'), the wider semantic as well as the situational and cultural context of a word-form. Yet again, how far does the context of a collocation stretch (cf. subsection 2.2.1.3)? As opposed to the German-Romance tradition (cf. Klotz 2000: 74), Halliday (cf. 1966: 151) insists that collocations can cross sentence boundaries. Occasionally, sentence boundaries may separate words co-occurring in collocation. Greenbaum claims the status of collocation for two items "[…] if they belong to a single remembered set" (1988: 14). Accordingly, context must be seen as variable, which in case of ambiguities has to be extended to guarantee that the analysis does not ignore relevant collocations (cf. also analysis of the node **JOB** in subsection 6.3.2). This is ensured by evaluating findings of a corpus not purely by quantitative means, but also by judging them with the linguist's intuition. This of course directly concerns relatedness and neighborhood (the size of the collocational span), two criteria discussed in the next subsection.

2.3.2 Relatedness and neighborhood

Previously it was shown that collocational partners have to occur in a common co-text and context. The notion of 'co-text' has remained particularly vague up to this point. The 'co-text' of a selected word-form refers to its verbal environment (cf. Tognini-Bonelli 2001: 87). The size of the span defines the amount of analyzed co-text of a node. The word-forms in question may be adjacent, but the direct adjacency of collocational partners may be interrupted (by function words like prepositions, punctuation marks or else). Even if they are not neighboring in those cases, they are still accepted as collocates, because they are part of a logical relation. Collocations are processed dynamically, so that additional material can be inserted in between node and collocates. Ideally then, in an empirical study, the entire context of a potential collocation needs to be considered. Subsequently, it would have to be decided for each instance whether the words in question combine to form a collocation or not.

Quantitative analyses require impartial and measurable criteria. In corpus linguistics there is a general consensus that the maximum span for a collocation is normally preset to ±4 or ±5 word-forms around the node (cf. Sinclair 1991, Kennedy 1998, McEnery and Wilson 1996). A narrower span returns too few collocations and a wider one comes up with too many potential collocates, so that the results are no longer representative. Needless to say, a span of ±4 or ±5 will not include all and only relevant collocations[7] and there will always be a certain error rate in such statistical analyses.

The question of the extent of the span is relevant and it is also important in which grammatical form node and collocate appear within it. Just as sentence boundaries do not automatically prevent collocations, collocates can also change word-class or re-arrange their syntactic structure, like in "pricey oil"/ "high price of oil"/ "high oil prices"/ "high-priced oil" (*BusinessWeek 2008*: JUL). The collocational relation, or relatedness, stays the same, no matter in which grammatically determined form they appear (cf. also analysis of the node **TALK** and its collocate STRAIGHT in subsection 6.3.2).

Above, different types of relatedness have been accentuated. Of course there can also be semantic relations between collocational partners, but this is not a necessary condition and "[…] semantic bonds alone cannot account for the status of collocation" (Handl 2005: 35). It can also be stated that collocation rather influences semantics (meaning by collocation) than vice versa. Questions arise on how new collocations emerge in a language, whether

[7] Mason (1999) developed a method to fix the span individually for each collocation.

they are based on extra-linguistic co-occurrence or start out from a dominant part and therefore bind other words with a certain meaning. All of them relate to the study of semantics rather than to that of collocation. On the one hand, meaning can be seen as an integral part of every item in a language. On the other hand, meaning may be regarded as a bundle of characteristics a word-form acquires from its context. Thus, collocations would not result from the meaning of their constituents but were the starting point for their meaning. This is directly related to Firth's definition of collocation "you shall know a word by the company it keeps" (Firth [1951] 1957a: 179), and meaning by collocation (cf. ibid.: 194, subsection 2.2.1.1).

2.3.3 Recurrence and frequency

As mentioned previously, the vast majority of collocation definitions indicate the recurrent nature of collocations. It is their frequent recurrence that determines their institutionalization. It can be assumed that what is observed frequently and recurrently across a corpus of authentic text produced by native speakers conforms to the accepted usage of language. The recurrence of collocations enables their recognition and learning based on experience. Collocations must be identifiable on the basis of their frequency, i.e. they must recur often enough to be perceived as habitual combinations. Collocational knowledge does not derive from the awareness of their individual constituents or grammar rules of a language. On the contrary, collocations are acquired through experience, through encountering text in the course of our lives (cf. Hoey 1991: 219) and through language use.

Language and thus collocations are subject to language change. New collocations are often formed in analogy to existing ones. They begin as free lexical combinations and then gradually change their status through recurrence and become collocations. New collocations can come into existence more or less ad hoc and are then established in the language through constant re-use by language users. A single occurrence of a word combination will not enter the area of collocation, because it is considered to be a free lexical choice which is not governed by prefabricated patterns or lexical preferences in a language. Ad hoc word-formations function the same way. They can only leave their status as neologisms and become institutionalized and lexicalized in a language if they are used by more than one person in more than one situation. Sinclair (1966) defines 'significant collocations' as combinations that "co-occur more often than the respective frequencies [of the constituents] and the length of the text in which they appear would predict" (Jones/Sinclair 1974: 19).

Yet, definitions in the literature remain largely rather vague on what exactly counts as frequent or how frequency or recurrence are to be quantified (cf. Bartsch 2004: 59). Corpus-based research into collocations should aim to clarify this.

Although frequency is a widely accepted criterion, it is not without controversy. This is closely connected to the criticism on the statistically oriented definition of collocation in general (cf. Hausmann 2004: 320 ff.). The present study follows this criticism only partly. Frequency cannot serve as an absolute criterion for assigning the status of collocation, but at the same time it is a basis for collocational analysis. Ad hoc collocations may exist, but if they are not repeated, they will vanish again. Collocations which fulfill the frequency criterion at a certain time are not bound to stay in language forever. Language, including collocations, is subject to diachronic change.

At this point it has to be added to the criterion of recurrence that a collocation has to be recurrent in different types of texts and texts from different domains. Otherwise it could be regarded as an idiosyncratic word combination, which is of little relevance to the language as a whole. This is not to say that there do not exist genre specific or technical collocations (cf. Caro Cedillo 2004), but their influence is typically limited. The four corpora in this study are collected from business and news magazines and ultimately this results in limited generalizability. Nevertheless, meaningful conclusions about the language of business and news magazines can be drawn.

2.3.4 Mutual expectancy and predictability

Another criterion closely connected to recurrence, frequency and relatedness is that of mutual expectancy (cf. Firth 1957b: 12, subsection 2.2.1.1). Cruse (1986: 40) also considers that in a collocation, “the constituent elements are, to varying degrees, mutually selective.” Sinclair (1991: 173) later notes also that “collocation is one of the patterns of mutual choice.” He thus indicates that one word automatically induces its partner, and the other way round. Mutual expectancy is of course preceded by a frequent common use. The notion of mutual expectancy has been re-interpreted as ‘predictability’ by other linguists, for example Greenbaum (1970), Bublitz (1996) and Herbst (1996). Schmid (2003: 243) clarifies the relationship between predictability and mutual expectancy as follows:

> In a way, predictability is […] the pragmatic counterpart to mutual expectancy: The former looks at word combinations from the language users’ perspective and the latter from the language-immanent perspective of the words themselves.

Mutual expectancy and predictability seem to capture the psychological essence of the phenomenon of collocation, the associative relation between words (cf. Lipka 2002: 181-82).

The assumption that occurrence of one word-form seems to almost automatically trigger the selection of other words in its context is reflected by intuitions which native speakers have of collocations. It is its psychological character which turns mutual expectancy into a promising criterion to decide on the status of co-occurrences, since their associative relations in the mind can be felt intuitively (cf. Handl 2005: 38). Herbst (1996: 389) considers this criterion to be "one of the most prominent features of a collocation", which is observable from results of association tests with native speakers. This element of collocation is difficult to operationalize for the identification of collocations in corpora, which is why mutual expectancy can only be employed by the human analyst as a criterion.

It is disputed in the literature whether constituents of a collocation are mutually expectant, i.e. whether all constituents exert the same influence in the selection process. Hausmann (1985, 1989) and Kjellmer (1991: 112) suggest that a dominant constituent exerts a stronger influence, i.e. that the selection process is directional. Moreover, it is problematic that mutual expectancy applies similarly to idioms and collocations, which means that it is not suited to tell the two apart. Schmid points to another difficulty by noting that the notions of mutual expectancy and predictability "[…] are highly subjective and of little reliability" (2003: 243). He also admits that empirical testing may lead to larger objectivity but requires "[…] a lot of effort while being restricted to relatively small sections of the lexicon" (2003: 244). The testing of psychological connections or associative bonds between word-forms is outside the scope of the current study. However, the psycholinguistic concept of 'priming' deals in depth with such issues (cf. Hoey 2005).

2.3.5 Semantic transparency and idiomaticity

As has been stated earlier, idioms are excluded from the present study. In order to distinguish collocations from idioms, a final criterion is introduced: semantic transparency. Collocations are usually distinguished from idioms based on the degree of semantic transparency of their constituents. While co-occurrence, neighborhood, recurrence and frequency of collocations are empirically testable in corpora, it is more difficult to establish their degree of semantic transparency.

Collocations need to be semantically transparent (cf. Cruse 1986: 40), i.e. the meaning of the collocation is obtained through the composition of the meanings of node and collocate.

Idioms, whose meaning is considered much more opaque than that of collocations, "do not reflect the meanings of their component parts" (cf. Benson et al. 1986: 252 ff.). Strictly speaking, this is, of course, a matter of degree. The meaning of most collocations can be composed of the meanings of their constituents. The problem with the notion of semantic transparency is that it is difficult to precisely establish the semantic contribution of the individual constituents of a collocation. It is generally assumed that collocations populate the gray area between one extreme of entirely transparent combinations, and the other extreme of completely opaque combinations. McKeown and Radev support this and see collocations as expressions that "fall somewhere along a continuum between free word combinations and idioms" (2000: 509). For illustration see Figure 2 below.

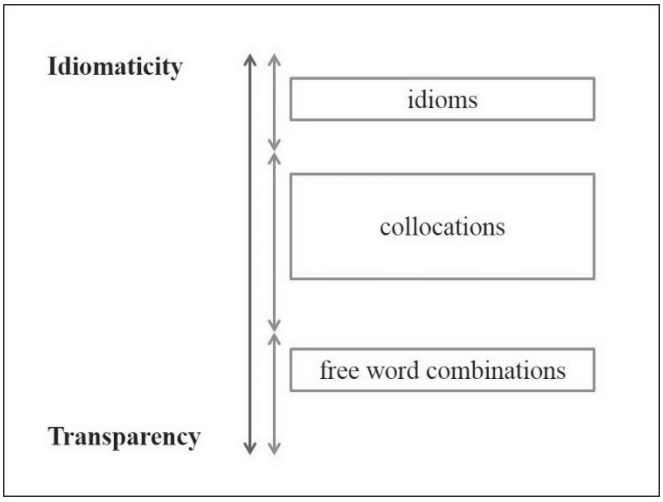

Figure 2 Idiomaticity/Transparency scale

Not all collocations are completely transparent. Some collocations consist of a transparent and an opaque constituent, which nevertheless lead to an intelligible combination. In most collocations all constituents contribute either lexical or functional meaning (cf. Bartsch 2004: 61). Some collocations may be said to have acquired additional meaning over and above the meaning contributed by the individual constituents. Manning and Schütze go even further by indicating that "there is always an element of meaning added to the combination" (cf. 1999: 151, 184). However, most researchers agree that collocations are fairly transparent and their meaning is deductible from the meaning of its constituents.

Expressions where none of the single meanings is conceivable in the understanding of the whole item, idioms, represent the extreme case of semantic opacity[8] (cf. Handl 2005: 39), as for example "cream of the crop" (*BusinessWeek 2008*: APR) or "ball park figure" (*BusinessWeek 2008*: MAY). While the first idiom refers to the "best people or things"

[8] Semantic opacity is considered to be the opposite term of semantic transparency.

(COBUILD 2009: 360) in a particular group and the second one denotes “an approximate estimate” (COBUILD 2009: 105). Idioms can be said to be wholly non-compositional (cf. Wood 1986: 31). Dobrovok’skij (1997: 45) criticizes this view by stating that some idioms are perceived as conceptual metaphors and therefore at least partially transparent based on previously acquired culture specific knowledge of native speakers. On the contrary, the non-native speaker stores and retrieves idioms as autonomous units of vocabulary. Still, from both perspectives, the translation of the single constituents of idioms does not reveal its proper meaning. Thus, the analysis of idioms is not beneficial for investigation of the initial hypotheses (cf. section 1.1) because the analysis focuses only on such co-occurrences which are semantically transparent.

Although studies have attempted to grade the compositionality of phraseological units, (cf. McCarthy et al. 2003, Baldwin et al. 2003) a clear distinction between collocations and idioms cannot be drawn (cf. Moon 1998, McKeown and Radev 2000). Nevertheless, “idiomaticity applies to encoding for collocations, but not to decoding” (Fillmore et al. 1988). Collocations then are easily interpretable on the basis of their individual constituents (but are difficult to generate because they are unpredictable for the non-native speaker). To sum this up, semantic transparency can be seen as a scalar criterion which separates collocations from idioms. Idiomaticity, on the other hand, includes anything from semantic transparency to semantic opacity and applies to both collocations and idioms.

2.4 The concept of collocation used in this study

It has been shown that there is no universally valid response to the query on what can pass as a collocation and that all descriptions are, to a great extent, interrelated.

The present study is mainly governed by the question whether Anglicisms, utilized by non-native speakers, display similar collocational behavior in the target language compared to their source language. The concept of collocation applied in this study has to allow for valid conclusions about the implementation process of Anglicisms in German business and news magazines. In this study, collocation is thus used as a notational term whose specific meaning is narrowed down according to the requirements of the particular research questions.

With the intention of arranging a representative corpus analysis, the study needs to rely on objective and manageable criteria. Rooted in the tradition of British contextualism and supported by corpus linguistic methods of analysis, collocation in this study is defined

statistically as regular and significant co-occurrences. Collocation is described with terminology according to Sinclair, who advanced the description of the concept of collocation beyond Firth's initial remarks. The study uses the term 'node' for the key word that is being studied and the term 'collocate' for any word that occurs frequently in the specified environment of a node (collocates need not be directly adjacent to the node). The span will be fixed to ±4. Early evidence (cf. Jones/Sinclair 1974: 21-23) suggested that this was a reasonable limit.

The basic assumption is that a collocation holds between two lexical items that co-occur more than five times in the corpora. While Kjellmer (1987: 133) states that collocation is "a sequence of words that occurs more than once in identical form in a corpus" five instantiations are regarded a reasonable minimum frequency in bear relation to the corpus size in the present study.

Collocations may at times function beyond sentence boundaries or syntactic structures, but the constituents have to display a logical relatedness at all times. This is to be judged as the case arises. Schenk (cf. 1994: 6) also states that sometimes collocation stretches across sentence boundaries.

Generally, statistical analyses and therefore corpora have at all times shown a preference for form-based research (cf. Muller 1963). In the present study, the key word lists consist of word-forms mainly because "[…] ultimately it is the word-forms that are related to meaning and patterned relations […] and not words or lexemes" (Esser 1999: 157). Besides, different word-forms can have different collocates (cf. Stubbs 1996: 172). Thus, the corpus analysis is word-form and not lemma based. Because mutual expectancy resembles psychological associations between the constituents of collocation, and is best judged by native speakers' intuition, it remains unaccounted for in the corpus analysis. Genuine idioms which are semantically opaque are excluded. Conversely, a word combination has to be semantically transparent to be considered a collocation.

The present study focuses on lexical collocates. Lexical words (collocates) are nouns, adjectives, adverbs, and main verbs. Grammatical collocates, on the other hand, are auxiliary or modal verbs, pronouns, prepositions, determiners, and conjunctions (cf. Biber et al. 1999). Concerning frequency, grammatical collocates recur more often because they come from a closed class where recurrence is much more prominent. The open class of lexical words (collocates) on the other hand leads to much more variation in syntagmatic relations and is thus more rewarding but also more difficult for collocational analysis.

It should have become clear that there are many different views on collocation and suggestions for criteria of the phenomenon. It is believed that the classification of collocation

reached is eligible to the specific aims of the present study and considers all important aspects in order to pursue an adequate analysis.

3 Related fields of research

The present study is in the tradition of British contextualism. In addition, further research paradigms and methods are essential. The study draws particularly on the concepts and methodology of ‘contrastive linguistics’, ‘translation studies’ and ‘corpus linguistics’.

3.1 Contrastive linguistics

3.1.1 Short historical background

Contrastive linguistics is founded on the assumption that languages can be compared and contrasted (cf. James 1980: 3). Contrastive linguistic analysis is concerned with the comparison of two or more languages in order to determine both the differences and the similarities between them (cf. Fisiak 1981: 1). ‘Typology’, an older and more established branch of comparative linguistics, already provides general statements about differences between languages and fundamental principles of language organization (cf. König 1991: 136). The methods and findings of typology may inspire and guide contrastive analyses of two languages in the search for further differences and similarities (cf. ibid.: 138). The linguist contrasts languages to gather confirmatory evidence for hypotheses suggested by the analysis of a single language (cf. James 1980: 7). Thus, contrastive studies of two languages are especially valuable because they reveal linguistic patterns that would not be apparent from a study of each of these languages in isolation (cf. König 1991: 152). Contrastive studies mostly deal with the comparison of languages that are ‘socio-culturally linked’, i.e. languages whose speech communities overlap in some way, typically through (natural or instructed) bilingualism (cf. Gast forthcoming: 1).

The idea to compare languages to obtain more information about their functioning is very old (cf. Rein 1983: 7) and there is a large body of literature on contrastive linguistics that cannot be exhaustively summarized here. What follows is a brief overview of the development of contrastive linguistics as a discipline. General overviews include Alatis (1968), Fisiak (1980, 1981), and James (1980).

Contrastive linguistics in the modern sense arose from pedagogical considerations (James 1980: 3). It can be considered an offspring of structuralist linguistics applied to the problems of foreign language teaching (cf. Dirven 1974: 2). The research field of ‘contrastive

analysis'[9] was outlined with the primary objective of making foreign language teaching more efficient. This is voiced by Charles C. Fries (1945: 9):

> [T]he most efficient materials [for foreign language teaching] are those based on a scientific description of the language to be learned, carefully compared with a parallel description of the native language.

Contrastive linguistics has been promoted as a means of predicting the difficulties in learning a particular target language for students with particular source language[10] backgrounds (cf. Lado 1957). Vice versa, early contrastive studies were motivated by the belief that making similarities explicit for the learner may facilitate the process of foreign language learning (cf. Nickel 1971). Interlingual lexical studies, stemming from the needs of foreign language teaching, translation and bilingual lexicography, resulted in works devoted to lexical pitfalls such as 'false friends'. Besides many results within the structuralist framework, the contrastive program was never researched on a wide empirical basis and lacked a solid foundation in the psychology of language learning and language production.

König (1991: 137) states that thirty years of contrastive linguistics had shown that work of this kind should not be tied too closely or exclusively to the goals of language pedagogy. This view had been expressed before (cf. Fisiak 1980). The link between contrastive analyses of two languages and the pedagogical goals initially associated with it were severed gradually over time. Corpus linguistic methods have given new impulses to descriptive studies in contrastive linguistics and the discipline has witnessed revitalization in recent years (cf. König/ Gast 2009: 7). The emergence of large bilingual corpora[11] allows contrastive linguists to test and quantify intuition-based contrastive statements within a body of empirical data that is advantageous – qualitatively and quantitatively – compared to the type of contrastive data that had been available before (cf. section 3.3).

The globalization of society leads to an increased awareness of the importance of interlingual and intercultural communication and played a major role in the further development

---

[9] An earlier term used to refer to contrastive linguistics and may be considered the more general term (cf. Dirven 1974: 2).

[10] The terms 'source language' and 'target language' originate in translation studies. In contrastive linguistics it is usually referred to 'L1' and 'L2'. There exists a close symbiosis between contrastive linguistics and translation studies. Because the present study is concerned with a directional comparison of languages it was considered adequate to use the terms 'target language' and 'source language' throughout.

[11] For example, the English-French 'Hansard Corpus' which was released in 1995 by Salim Roukos et al. or the 'English-Norwegian Parallel Corpus' (ENPC) compiled by Stig Johansson and Knut Hofland and released in 1997.

of contrastive linguistics. It seems that multilingual studies in general have captured researchers' attention lately. Contrastive linguistics has also moved on towards analyzing English in cross-cultural and cross-linguistic communication, especially in international contexts (cf. Crystal 2004). Research in this area is essential for effective cross-cultural exchange of ideas. In line with state-of-the-art research interests in contrastive linguistics, the present study analyzes the collocational behavior of Anglicisms. Additionally, the present study maintains that cross-cultural studies of English as used by various communities are important in order to understand further cultural concepts and influences involved in the use of English.

3.1.2 'Theoretical' versus 'applied'

Early contrastive analysts were not systematically concerned with methodological problems, and emphasis fell on the actual practice and applications rather than the building of a theory and methodology of contrastive linguistics (cf. Krzeszowski 1990: 23 ff.). Many linguists, occasionally or systematically, practiced what, in general, falls into the domain of contrastive linguistics by evoking cross-language evidence to support theoretical claims (cf. Nickel 1971, Lakoff 1972, Comrie 1976). Few works, such as James 1980, Di Petro 1971 and van Buren 1976 were explicitly devoted to the theory and methodology of contrastive studies. Comprehensive publications concerned with English-German contrasts include Hawkins 1986, Mair/Markus 1992 and König/Gast 2009.

The literature (cf. James 1980: 127) distinguishes between 'theoretical' (or 'pure') and 'applied' contrastive linguistics. Theoretical contrastive studies have been defined as giving "an exhaustive account of the differences and similarities between a given pair of languages" (Fisiak 1981: 2). Applied contrastive studies are assumed to be a part of applied linguistics and should direct the comparison of two languages towards a specific non-linguistic purpose, such as the explanation of interference errors (cf. Fisiak 1981: 3).

The methodological dichotomy of theoretical and applied contrastive linguistics does not remain without criticism and mediating views have been expressed. In Gast's (forthcoming) understanding, contrastive linguistics, although not a branch of applied linguistics, aims to arrive at results that carry the potential of being used for practical purposes, e.g. in foreign language teaching and translation. As contrastive linguistics provides the descriptive basis for such applications, its research program can also be summarized as 'comparison with a purpose'. Gast (forthcoming) articulates on the same issue:

> With its largely descriptive focus contrastive linguistics provides an interface between theory and application. It makes use of theoretical findings and models of language description but is driven by the objective of applicability.

Systematic comparison is only possible within the framework of an explicitly stated theory. This is true of any kind of contrastive linguistic analysis, because reliable exploration of facts cannot be conducted without a theoretical background which provides concepts, hypotheses and theories.

The present study does not deal with all the problems connected with the theory and methodology of contrastive linguistics. While the pedagogic interest in contrastive analyses has declined, the value of contrastive linguistics extends far beyond explaining and hypothesizing about language teaching. Contrasting can be an excellent way of highlighting structures of the languages compared.[12] Vachek (1964: 307) points out that "the relative importance of a linguistic fact […] of a given language […] may be set off by a well-considered use of foreign comparative material." Similarly, Schröder (1991) assumes that to study texts comparatively across cultures may yield a variety of valuable findings for language learners and practitioners. Contrastive linguistics has proved to be of considerable importance both from a practical and a theoretical point of view and is important for the understanding of language in general as well as the study of the individual languages compared (cf. Johansson/Hofland 1994: 25).

All in all, the present study is a contribution to contrastive linguistics, to be exact, to contrastive collocation research. Whether this study should be classified as exclusively or primarily 'applied' or 'theoretical' ultimately cannot be determined and is of little importance to the efforts of this study. What can be said is that the study exceeds the mere description of errors and can consequently be theory grounding. More precisely, this study documents the results of a first analysis which bear the potential for various practical applications. The results of the present study may suggest the development of new theoretical concepts.

[12] This is an approach to language comparison practiced within the *Prague School* ('analytic comparison').

3.1.3 Tertium comparationis

'Tertium comparationis' characterizes a common platform of reference (cf. Krzeszowski 1990: 15). It is an instrument of language comparison and contrast. The problem of establishing comparability and of finding the 'third of comparison' (*tertium comparationis*) is a major issue in any kind of contrastive work (cf. König/Gast 2009: 5). Unfortunately, the definition of this central concept of contrastive linguistics is highly variable. All contrastive studies involve the basic assumption that the linguistic items to be compared have something in common against which differences can be stated. This may refer to their common structural levels of description regarding phonology, morphology, lexicology, syntax and text-linguistics. Depending on the platform of reference (or *tertium comparationis*) adopted, and the kind of equivalence involved, linguistic items turn out to be either similar or different.

In the present study, several levels of such a common platform of reference are of importance. At the text-linguistic level a first *tertium comparationis* for English and German are the text types: business and news magazine articles. A second *tertium comparationis* at the level of lexis is the identical manifestation of the nodes in both languages, i.e. the sets of analyzed key words (Anglicisms) which are identical in both languages and whose collocates are being contrasted. Thirdly, pragmatic equivalence, i.e. the overlap of semantic content of equivalent collocates in both languages provides a further *tertium comparationis*.

Closely related to pragmatic equivalence is the following discussion which is dedicated to 'formal correspondence' and 'semantic equivalence' as *tertia comparationis*. They are meaningful for the corpus analysis and their description leads up to section 3.2 which is concerned with translating collocates, i.e. equivalence relations. Formal correspondence and semantic equivalence can serve as *tertia comparationis* for lexical contrastive studies (cf. Lado 1957: 52 f.). In contrastive linguistics, 'correspondence' is used to describe the resemblance and difference between words in terms of their linguistic form (cf. Munday 2009: 177, Catford 1965). Consider the following example in Figure 3 below.

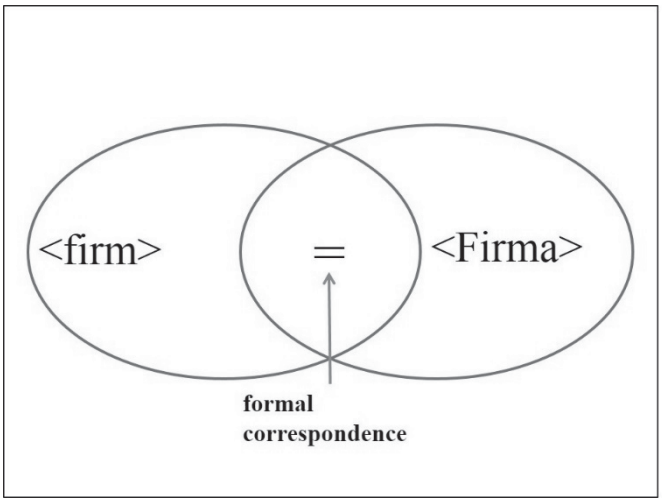

Figure 3 Formal correspondence

Formal correspondence refers to the features of form which are reproduced in the target language, i.e. there is a clear relation in terms of shared elements or features between word-forms as to their sounding and writing such as *firm* and *Firma*. In Figure 3 <firm> and <Firma> are not only formally correspondent, but can also be considered 'pragmatically equivalent' within the present study (the concept of 'pragmatic equivalence' will be further explained on pages 52 ff.). Many words that are formally correspondent in two languages are also similar in meaning and at times even semantically equivalent, for example English *sock* and German *Socke*. Sometimes such word-forms are pointed out as 'cognates' to the learner. In linguistics, cognates across languages are often defined as words that have a common etymological origin (cf. Sunderman/Schwartz 2008: 19). Here, a synchronic definition of cognates as used within psycholinguistics is employed instead, namely any two words with shared aspects of writing, sounding, and meaning across two languages (cf. Carroll 1992).

On the other hand, formal correspondence occasionally leads to the establishment of so-called 'false friends'. "This characterizes the situation where formal […] correspondence is accompanied by completely different reference" (Esser 1980: 186). Examples are English *gift* ('present') and German *Gift* ('poison'). It has been pointed out that formal likeness alone cannot serve as a *tertium comparationis* without support from semantic equivalence (cf. Liston 1970: 44). On the contrary, the contrasting of linguistic items which are semantically equivalent (but formally different) is generally considered meaningful. For example German *Bilderrahmen* and English *picture frame* can be considered semantically equivalent, but formally different.

'Semantic equivalence' means that the content of an expression remains identical when transferred from source to target language. Equivalence at word level means that a source

language word-form and target language word-form have the same meaning and refer to the same non-linguistic entity. Strictly speaking, two entities are equivalent if neither has features that the other lacks. Semantic equivalence is most easily demonstrated with entities which only have one meaning, thus one sense. This is rarely the case, but can be illustrated with the example in Figure 4 below.

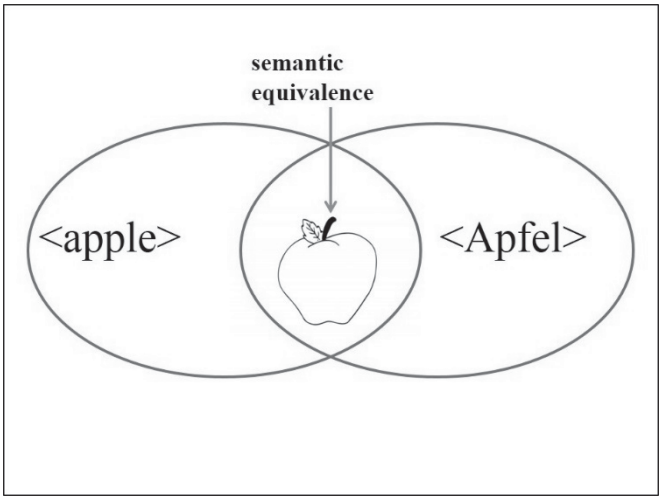

Figure 4 Semantic equivalence

It is certainly debatable whether *apple* and *Apfel* display limited formal correspondence in addition to their semantic equivalence. Establishing full semantic equivalence is difficult, since "the notion of equivalence is normally relative owing to language-dependent characteristics and cultural influences" (Munday 2009: 185). This view had also been expressed earlier by Esser (1980: 185) who points out that the relativity of semantic equivalence is to some extent owed to 'language particular concept formation'. This relates directly to Coseriu's (1970: 11) distinction between language particular concepts and those concepts which are not language particular.

The difficulty of determining semantic equivalence can be illustrated with the example of English *tree* and German *Baum*. *Baum* and *tree* are formally different, but they have the same referent (cf. Esser 1980: 185). Still, the two words are not semantically equivalent. This is partly owed to language particular concept formation and partly due to the different meanings of *tree* and *Baum*. For illustration see Figure 5 below.

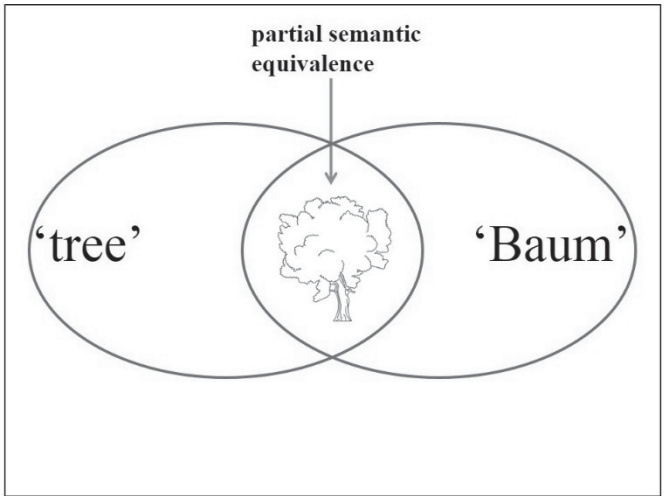

Figure 5 Partial semantic equivalence

Figure 5 shows a partial semantic equivalence relation. This is often the case and relates to the fact that most entities like ‘tree’ and ‘Baum’, for example, encompass more than one sense and that they are not in all senses semantically equivalent. Again, it should be pointed out that cases of ‘mono-sense’ like illustrated with ‘apple’ and ‘Apfel’ in Figure 4 are rather the exception. The different senses of ‘tree’ and ‘Baum’ are as follows:

> *tree*$_1$ – a woody perennial plant, typically having a single stem or trunk growing to a considerable height and bearing lateral branches at some distance from the ground. (www.oxforddictionaries.com)
>
> *tree*$_2$ – a drawing that connects things with lines to show how they are related to each other (also *tree diagram*). (www.oxforddictionaries.com)
>
> *Baum*$_1$ – ein Holzgewächs mit festem Stamm, aus dem Äste wachsen, die sich in Laub oder Nadeln tragende Zweige teilen. (www.duden.de)
>
> B*aum*$_2$ – Graph mit mehreren Knoten, deren Verbindungslinien kein geschlossenes Netz bilden, sodass je zwei Knoten durch genau einen Weg miteinander verbunden sind. (www.duden.de)
>
> *Baum*$_3$ - Der Baum ist ein Teil eines Segelbootes oder einer Segelyacht und dient zum Aufspannen und zum Einstellen des daran befestigten Segels. (www.duden.de)

It can be said that ‘Baum’ comprises senses that ‘tree’ lacks and the other way around. This is problematic, because two entities are semantically equivalent only if neither has features that the other lacks. Consider the following Figure:

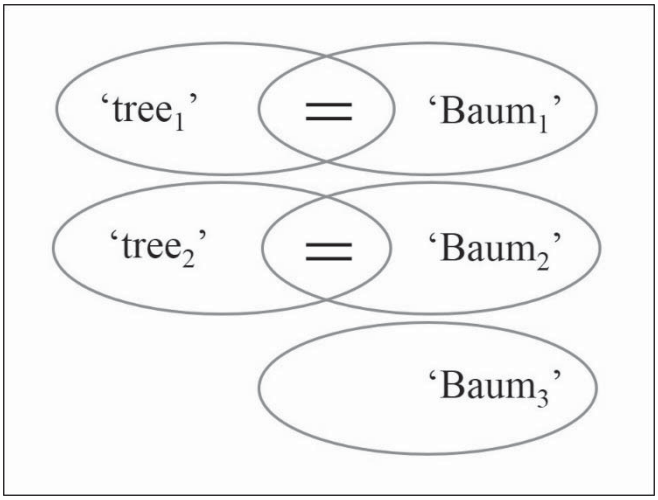

Figure 6 Senses of *tree* and *Baum*

Figure 6 illustrates the different senses of *tree*, namely 'tree$_1$' and 'tree$_2$', and those of *Baum*, namely 'Baum$_1$', 'Baum$_2$' and 'Baum$_3$'. 'Tree$_1$' and 'tree$_2$' are polysemous, as these two senses of *tree* are related. The same holds for 'Baum$_1$'and 'Baum$_2$'. 'Baum$_3$' in contrast is a homonym, as its sense is not related to the senses of 'Baum$_1$' and 'Baum$_2$' and is hence not a case of polysemy.

The overlapping circles demonstrate that *tree* and *Baum* show partial semantic equivalence because the senses of 'tree$_1$' and 'Baum$_1$' as well as the senses of 'tree$_2$' and 'Baum$_2$' are semantically equivalent. 'Baum$_3$' does not show any semantic equivalence with any of the senses of *tree*. The correct English translation for *Baum$_3$* is *boom*, not *tree*. In the following, partial semantic equivalence as demonstrated in Figure 6 is referred to as 'pragmatic equivalence'. Hence, pragmatic equivalence is established when at least one of the senses of two entities is semantically equivalent. This is decisive for the corpus analysis in order to determine equivalence relations between collocates.

Among scholars there is controversy around the concept of 'equivalence', particularly its definition, relevance, validity and applicability. The assessment of equivalence ranges from rejection (cf. Snell-Hornby 1988) to assuming that it holds helpful implications (cf. Baker 1992) to indispensability (cf. Nida/Taber 1969). Holmes (1988) and Toury (1980) argue that equivalence should be replaced by a more relative term. Much of this controversy results from the assumption that full equivalence is an achievable goal. But in reality not everything can be translated exactly into a different language. Holmes expresses this when he assesses the use of the term equivalence to be "perverse", since to ask for sameness is to ask too much (cf. 1991: 28). Accepting that equivalence can only be achieved to a certain degree, the present study analyzes collocates in their source and target language which show partial semantic equivalence, i.e. pragmatic equivalence. Pragmatic

equivalence will also be dealt with in detail in section 3.2. Collocates can be formally correspondent and pragmatically equivalent at the same time. Referring back to Figure 3 and the formal correspondence of <firm> and <Firma>, it can also be stated that *firm* and *Firma* are pragmatically equivalent as one of their senses is semantically equivalent. This could be demonstrated in the same manner as for *tree* and *Baum* above.

In contrastive practice, semantic equivalence is often identified with translation equivalence. Halliday et al. (1964: 115) state that comparability can be established by reference to translation. Vinay and Darbelnet (1995 [1958]: 255) claim that expressions between language pairs are acceptable as 'full equivalents' as long as they are listed in a bilingual dictionary as such. This understanding veers toward the concept of pragmatic equivalence. Consequently, the concept of equivalence is closely related to another branch of linguistics, translation studies.

Translation can serve as a tool of contrastive analysis and the findings of contrastive linguistics may have useful implications for translation studies. Recent publications and conferences have brought together specialists from the disciplines of contrastive linguistics and translation studies. After all, they do cover partly common ground, and with the emergence of corpora they have started to converge (cf. Granger 2003: 17). Today, corpora provide a common resource for academics of contrastive linguistics and translation studies.

3.2 Translating collocates

As has been explained above, the notion of 'equivalence', holding between two items in different languages which are being compared, is vital to any contrastive analysis. Similarly, 'equivalence' is a fundamental concept in translation studies and central to the practice of translation. The notion of 'equivalence' is also relevant to the present study. Consider Hypothesis 1 (cf. section 1.1) which is repeated for convenience below:

Hypothesis 1

The use of Anglicisms produces replications of source language collocational structures within the target language, i.e. Anglicisms feature identical or equivalent collocates in the source and target language corpora.

The present study differentiates between identical and equivalent collocates. It is necessary to understand the difference between these types of collocates. It is not claimed that the value of the following definitions extends beyond the purposes of the present study.

In the present study 'identical collocate' refers to collocates of a given node with identical written word-forms in the source and target language corpora. Example (4) below illustrates one of the quantitatively identified[13] identical collocates of the node **BUSINESS**.

(4)

NODE	
BUSINESS	
COLLOCATE SL	**COLLOCATE TL**
DEVELOPMENT	DEVELOPMENT

<DEVELOMPENT> (source language) and <DEVELOPMENT> (target language) are considered identical collocates of **BUSINESS** because they are identical in form and meaning. For a better understanding, sample sentences from the source and target language corpora follow.

(4a) […] a 39-year-old woman who is in charge of business development at the carrier […]. (*BusinessWeek 2008*: SEP)

(4b) In den meisten Unternehmen kümmern sich Stabsstellen für business development darum […]. (*WirtschaftsWoche 2008*: MAY)

A definition of 'equivalent collocate' is much more complex. The more general the meaning of a word is, the more polysemous that word will be, and the more translational equivalents it will have in other languages (König/Gast 2009: 217). Most words are polysemous, but words usually occur in a specific situation and linguistic context. Coseriu (1981: 191) elaborates on this issue as follows:

> In der Übersetzung geht es nicht um unmittelbare Entsprechungen zwischen den Bedeutungen verschiedener Sprachen, […] sondern um Entsprechungen zwischen Bedeutungsverwendungen, d.h. Bedeutungsvarianten. Und eine Bedeutungsvariante weist in jedem Fall mehr Züge als die ihr entsprechende funktionelle Einheit auf, da zu den einzelsprachlich gegebenen Zügen bei jeder Variante noch die kontextuell und situationell gegebenen Züge hinzukommen.

[13] The qualitative corpus analysis (cf. subsection 6.1.2) will show that DEVELOPMENT is not an identical lexical collocate of **BUSINESS**, but that they are constituents of the compound **BUSINESS** DEVELOPMENT.

This usually means that only one sense (of polysemous or homonymous words) is involved in a given co- and context and thus only one translation (which is considered its pragmatic equivalent). It is ultimately the task of the researcher to dispel ambiguities. In such instances, this study relies thus on qualitative, manual, introspective analysis. Pragmatic equivalence, i.e. the closest approximations to word-for-word translations, provides a *tertium comparationis* (cf. subsection 3.1.3) in this study.

The present study refers to bilingual dictionaries as collections of translation equivalents on the basis of partial semantic equivalence, i.e. pragmatic equivalence. Although "a dictionary seems the obvious place to find a record of the meanings of a word" (Halliday 2004: 23), there is a paradox hidden in this. The bilingual dictionary is meant to be the primary tool for translation and although its content, the translation equivalents, are based on and derived from translation acts, it is often claimed that the translation equivalents offered in the bilingual dictionary are only of limited use to the translator. This paradox comes about through a process of abstraction or idealization that lexicographers are prone to as much as contrastive linguists, or any linguist. Halliday (2004: 24) comments on this:

> To do the job of more or less individually presenting words in an accessible list, the dictionary takes words away from their common use in their customary settings. In fact a dictionary is a highly abstract construct.

Dictionaries observe language in use and abstract from all the various contextual uses of a word its distinct core sense or senses. Lexicographers concentrate on those meanings that produce a neat lexical equation. The judgment as to the optimal translation and its correctness is ultimately left to the authority of a competent bilingual speaker and is an offer of plausibility which can be accepted or argued about. Again, the present study is bound to rely on the researcher's intuition to determine the equivalence of collocates. This process is guided by the notion of pragmatic equivalence. The target language word-form should bear the relevant features of the meaning of a word-form in the source language in a given context (cf. Kussmaul 1995: 86).

The aim is to apply the notion of 'equivalence' as it is understood in contrastive linguistics and translation studies to the corpus analysis of the present study. However, even experts in the field admit that equivalence "is influenced by a variety of linguistic and cultural factors and is therefore always relative" (Baker 1992: 6). This study differentiates between identical and equivalent collocates. Equivalent here refers to 'pragmatic equivalence' which is claimed for collocates of English and German. The fact that English and German are both members of the Western branch of the Germanic language family and thus closely

related genetically (cf. König/Gast 2009: 4) possibly alleviates the establishment of contrast and equivalence, respectively.

'Equivalent collocate' denotes collocates in the target language which are a translation of a source language collocate. Consider the following example of equivalent collocates of **BUSINESS**. (The collocations in examples (4) and (5) are taken from the business magazine corpora).

(5)

NODE	
BUSINESS	
COLLOCATE SL	**COLLOCATE TL**
LEADING	FÜHREND

BUSINESS collocates significantly with LEADING in the source language corpus and FÜHREND[14] in the target language corpus. Examples are:

(5a) […] the technical offerings of Thomson's leading legal research business, Westlaw. (*BusinessWeek 2008*: MAY)

(5b) […] will die Inge Watertechnologies wieder zum weltweit führenden business aufbauen. (*WirtschaftsWoche 2008*: JUL)

Every time possible equivalent collocates occur during corpus analysis, comparability needs to be established between German and English collocates of identical key words (nodes). Whether a source language collocate and a target language collocate are equivalent is connected to the specification of equivalence. The question of defining equivalence is concerned with semantics, particularly with transfer of semantic content from source language to target language (cf. Bassnett 2002: 34). Translation as a method of contrasting must be regarded with caution though, and it is important to distinguish between 'semantic equivalence' and 'translation equivalence' (as introduced by Catford 1965: 27-34).

As has already been pointed out in subsection 3.1.3 above, semantic equivalence is an ideal and is theoretically untenable for each instance in this study. Facing the issue of partial semantic equivalence when searching for translation equivalents of collocates, the present study reverts to the concept of 'pragmatic equivalence'. Pragmatic equivalence in this

[14] Although the corpus analysis is word-form based, search for FÜHREND has to include the search for its inflectional variants FÜHRENDE, FÜHRENDES, FÜHRENDEN because they are all translation equivalents of LEADING.

study describes the assessment of partial semantic equivalence (of collocates) by the researcher. ‘Pragmatic equivalence’ is understood as the closest approximation possible to the meaning of source language word-forms in a given co- and context.

For example, one of the translations of the English noun *firm* is German *Firma*. *Firm* has several other translations in German, for example *Unternehmen* and *Anwaltsbüro* or *Anwaltskanzlei* (www.pons.eu). Therefore *firm* and *Firma* do not meet the requirements of ‘semantic equivalence’, but only that of ‘partial semantic equivalence’, i.e. ‘pragmatic equivalence’ because at least one of their senses can be considered as semantically equivalent. Whether translation equivalents, such as *Firma*, *Unternehmen* or *Anwaltskanzlei/-büro* really prove to be pragmatic equivalents depends on their context. *Firm* and *Firma* are considered equivalent collocates in the corpus analysis because they form translation equivalents in a given context. <Firm> and <Firma> also display formal correspondence. It has been mentioned before that this is not necessary to assign the status of equivalent collocates. Below, Figure 7 summarizes correspondence and equivalence relations of the equivalent collocates *firm* and *Firma*.

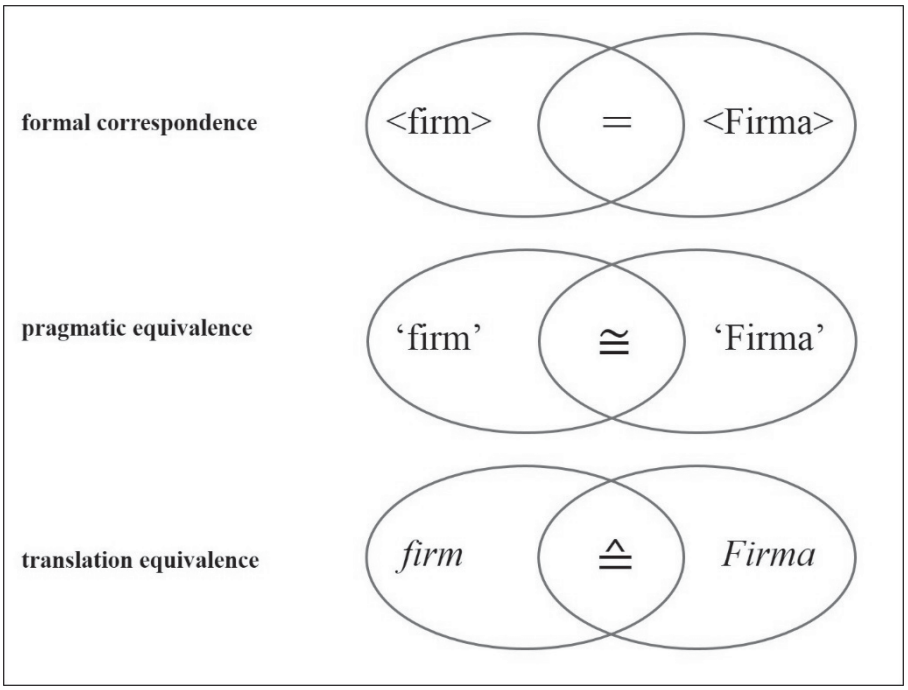

Figure 7 Correspondence and equivalence relations of *firm* and *Firma*

It may be argued that a study which deals with collocations should also be concerned with equivalence above word level. Equivalence at word level is concerned with the meaning of single words. Equivalence above word level explores combinations of words and phrases, and collocations as well. As stated by Cruse (1986: 40), collocations are “[…] fully transparent in the sense that each lexical constituent is also a semantic constituent.” Therefore, node and collocate each contribute with their individual meaning to the overall

meaning of the collocation. There are, however, cases in which one of the constituent elements of a collocation acquires a slightly different meaning from the one it normally carries outside that particular collocational environment. One example would be *white coffee* (coffee with milk). Although the adjective *white* is used in a figurative sense, the other constituent, *coffee*, maintains its literal meaning. Thus, the collocation *white coffee* is semantically transparent and clearly distinct from idiomatic expressions like *white elephant* (a possession that is useless, especially one that is expensive to maintain or difficult to dispose of (www.oxforddictionaries.com)) in which both constituents are used figuratively and in which the overall meaning of the expression is completely idiomatic (cf. Martelli 2007: 27).

While the present study excludes idioms, semantic transparency remains a scalar criterion (cf. subsection 2.3.5). The present study employs a largely statistical definition of collocation and is dedicated to the analysis of transparent collocations. It is assumed that the specific meaning of a given node or collocate contributes to the overall meaning of a collocation. Equivalent collocates are translated as autonomous word-forms which form part of a transparent collocation. Therefore, a discussion of equivalence above word level can be dispensed with. Nevertheless, consideration of the entire collocation (node + collocate) may limit the number of pragmatically equivalent collocates. Thus, contextual influence is also reflected in the restrictions upon the translations deemed adequate in the analysis. It should have become clear that the choice of the "right word" to use in the target language is often a subtle process.

In the context of this study it is important to note one additional fact. The first step in the corpus analysis is to identify highly frequent collocates in the source and target language corpora of a given node which is identical in the English and German texts. This process is carried out automatically with the help of lexical analysis software (cf. section 5.3). In search for equivalent collocates, the researcher compares these so compiled lists of highly frequent identified collocates. The lists may hold identical collocates or (pragmatically) equivalent collocates. Thus, the actual performance of translation for the identification of equivalent collocates is strongly limited. The task of the researcher really is to compare those highly frequent collocates in source language and target language corpora and determine identity or (pragmatic) equivalency between these word-forms retrieved from the corpora.

3.3 Corpus linguistics

Traditionally, linguistic analyses have emphasized structures and were of a rather abstract nature. A different perspective, which is the focus of corpus linguistics and this study, is to emphasize frequency in language use (cf. Biber et al. 1998: 1). It is now common to address theoretical issues through the examination of bodies of naturally occurring language use (cf. Bybee 2006: 712). The present study investigates how journalists utilize the resources of their language and it studies collocations in actual and genuine instances of use.

The rise of corpus linguistics is one of the most recent and influential developments in linguistics. The possibility of compiling and analyzing electronically stored corpora has led to this new academic discipline (cf. Kennedy 1998: 3). It has to be pointed out though that there have been corpus-based studies before the invention of computers. Quirk, for example, began his *Survey of English Usage* in 1959 (cf. Quirk et al. 1985: 33) and in some traditions, like American structuralism, linguistic analyses were based on authentic language data even earlier (cf. Fries 1940: 24, 1952: 3). However, the first comprehensive machine-readable corpus was compiled in the USA as the *Brown Corpus* in 1963/1964. Corpus work has by no means been restricted to the English language, but the development here has been most remarkable (cf. Svartvik 1992: 8).

Research possibilities within corpora are innumerable (cf. Kennedy 1998: 88 ff.). The most obvious aim, however, is to examine the preferences of a language, since the computer-based corpus enables the researcher to analyze words in their context in the form of so-called ‘concordances’. It is thus easier to observe regularities that would be difficult to catch when scanning through texts manually. The focus of corpus linguistics is on performance rather than competence and on observation of language in use leading to theory than vice versa (cf. Leech 1992: 107).

One of the major issues in corpus design is their ‘representativeness’ or ‘typicalness’ (cf. Kennedy 1998: 63). This is important because corpus linguists are usually not only concerned with what is possible, but mainly with what is likely and frequently to occur in language use. While it is sometimes argued that corpora should be as large as possible (cf. Sinclair 1991: 18), considerations of representativeness should prevail over efforts to merely increase the size of a corpus. A corpus aims to be representative in the sense that findings based on its analysis can be generalized for the language as a whole or a specified subset of it. The outcome of a corpus linguistic study can then, to a certain extent, be seen as a basis for broader conclusions. Results cannot be regarded as definite facts about a language, but as strong and highly probable tendencies all the same.

Corpora can supply the empirical basis for linguistic research, as they are very useful to objectify intuitions about linguistic phenomena. While there is a general tendency to supplement qualitative analyses with quantitative methods to a major or minor degree, the extent to which corpora should play a role in linguistic analysis is disputed. Corpus-based analyses must go beyond simple counts of linguistic features. McEnery/Wilson emphasize this by expressing that it is essential for corpus data not to be used solely for quantitative research, and that, in fact, many researchers have used corpora as a source of qualitative data (cf. McEnery/Wilson 1996: 61). It is important to include qualitative, functional interpretations of quantitative patterns. Biber et al. (1998: 5) summarize this as follows:

> The goal of corpus-based investigations is not simply to report quantitative findings, but to explore the importance of these findings for learning about the patterns of language use.

The present study is corpus-based rather than corpus-driven since it puts the concept of collocation, which has been formulated before large corpora were available, to the test (cf. Tognini-Bonelli 2001: 65). The basic aim is to retrieve collocations from four corpora (cf. chapter 4) on the basis of frequency. The corpora supply language data, the distribution of which then serves as a basis for testing the hypotheses. Consequently, corpus linguistic methodology here serves to reveal specific preferences in word combinations and a certain patterning which can be observed across languages. The corpus-based approach takes advantage of computers' capacity for fast, accurate and complex analyses of large amounts of data (cf. Biber et al. 1998: 233).

It is beyond doubt that linguistic studies gain from using corpora, and that since the 1990s the utility of corpora for language analysis is no longer seriously disputed. Yet, corpus-based methodology also has its limitations (cf. Granger 2003: 22), such as restricted availability and the preference for form-based research. Corpora can suffer from problems related to size and representativeness (cf. Sinclair 1991: 17 ff.). Still, even if the use of large computerized corpora is not useful for all kinds of linguistic studies, the use of solid empirical data, whether electronic or not, should be encouraged (cf. Granger 2003: 23). The positive effects of corpus analysis far outweigh its potential disadvantages. Nevertheless, this study is also aware of the criticism of corpora and does not rely purely on automatically processed data.

4 The corpora

The present study deals with a contrastive analysis of journalistic texts taken from American business and news magazines and their German counterparts. From the beginning, the importance of empirical and corpus-based research for this project has been stressed. Four corpora were collected to form the basis of analysis. Chapter 4 addresses issues related to the compilation of the corpora and thus clarifies the rationale of this study, the significance of the research and its results.

4.1 External criteria for corpus classification

4.1.1 Overview

There is a wide array of non-linguistic and linguistic criteria to choose from for the compilation of corpora. The text itself, its production and reception need to be considered (cf. Stubbs 1996: 3). Such linguistic characteristics and extra-textual criteria are the basis for the classification of the corpora in this study. The external criteria for text classification in this study are ‘origin’, ‘function’, ‘subject matter’ and ‘audience’. Magazines can be categorized into types according to these criteria. ‘Origin’ classifies a text with regard to its medium, i.e. whether it is spoken or written. In the context of the present study, ‘origin’ also designates the specific magazine from which the texts were collected. The criterion ‘function’ describes the magazine’s prototypical uses. Magazines can be classified as primarily informative, entertaining, persuasive or social, for example. ‘Subject matter’ refers to the articles’ themes and a magazine’s editorial scope. All mass communication media are intended for audiences that are large and their exact size and membership are unknown. The criterion ‘audience’ looks at the target audience of each magazine type, the audiences’ particularities and varieties, and at the relationship between journalist and readership.

‘Origin’, ‘function’, ‘subject matter’ and ‘audience’ are central to some of the approaches that are taken to describe and systematize the correlation between types of text and conventionalized linguistic and non-linguistic features. These approaches have been structured, among others, by ‘genre’ and ‘register’ analysis. Corpus classification is related to these concepts and ‘genre’ and ‘register’ have been long established to serve as guidelines for the compilation of corpora such as the prominent *Survey of English Usage,* the *Brown Corpus,* or the *British National Corpus* for instance.

A number of definitions of ‘genre’ have been influential, notably those of Biber (1988) and Swales (1990). Biber describes ‘genre’ as “[…] refer[ring] to categorizations assigned on

the basis of external criteria" (1988: 70). His list of external criteria includes "[…] subject-matter, purpose, […] the relation between the communicative participants, the relation of the participants to the external context, and the relation of the participants to the text itself" (1988: 70). Further examination of Biber's work clearly shows that he views the notion of genre from a similar perspective to that expressed by Swales (1990: 58).

> A genre defines a class of communicative events, the members of which share some set of communicative purposes. […] [E]xemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience.

These definitions by Biber and Swales each hold three out of four parameters taken as external criteria for corpus classification in this study. 'Purpose' (Biber) and 'communicative purposes' (Swales) match the current study's criterion 'function'. 'Subject-matter' (Biber) and 'content' (Swales) are in line with 'subject matter' used in this study. The relations of the communicative participants to one another, the external context and the text itself (Biber) relate very well to the field covered by the criterion 'audience' in this study. In comparison, Swales restricts the complexity of this criterion by using the term 'intended audience'.

The approaches within different branches of literary and linguistic theory which define situational or subject-specific language have resulted in the terms 'genre' *and* 'register'. Their meanings are interrelated and many definitions of these concepts exist (cf. Esser 2009: 75). Halliday (1978: 111) defines 'register' as "the configuration of semantic resources associated with a situation type: it is the meaning potential that is accessible in a given social context." Registers, according to Halliday et al. (cf. 1964: 90), may be distinguished according to 'field of discourse', 'mode of discourse' and 'style of discourse'. In the context of this study, Halliday's 'field of discourse' conforms to the external criterion 'subject matter'. Esser (cf. 2009: 78) explains that 'mode of discourse' corresponds to the text-external parameter 'medium', which is consistent with 'origin' in the current study. Finally, 'style of discourse' "refers to the social relations between the speaker/writer and the listener/reader" (ibid.). 'Style of discourse', then, is closely related to 'audience' as it is used in this study, because not only the target audience of a given magazine is examined, but also the relationship between the journalist and the reader. In summary, "[t]he linguistic features which are typically associated with a configuration of situational features-with particular values of the field, mode and tenor – constitute a register" (Halliday/Hasan 1976: 22).

Closely related to these efforts to categorize text is the concept of 'text types'. In the 1960s, Halliday et al. (1964: 89) already formulated the idea to distinguish "groupings of text that

are similar with respect to their linguistic form" (cf. Esser 2008: 77). Later, in the 1980s, Biber introduces 'text types' which represent such groupings of texts that are linguistically well-defined, distinguished by the occurrence of linguistic features on the levels of grammar, syntax and lexis, irrespective of genre categories (cf. 1988: 70, 1995: 10). Thus, linguistically similar texts from different registers or genres may represent a single text type (cf. Biber 1985, 1986). Because text types are the result of preceding linguistic analysis, this concept is not functional for external corpus classification.

To summarize the above, the corpora in this study are defined primarily on the basis of text-external criteria, tailored to journalistic writing. Business and news magazine articles are viewed as language varieties that are embedded within socio-cultural, institutional and discipline-specific histories of practice. Criteria that are similar to the ones used in this study can be found in some of the most influential approaches taken to define 'genre' and 'register'.

4.1.2 Origin

The corpora consist of journalese magazine texts intended for mass communication. They are written, monologic texts. It may be argued that the possibility of a dialogue between the journalists and readers of magazines cannot be ultimately dismissed. A conceivable form of dialogue might be letters to the editor. All the same, such written and asymmetrical communication will not typically be considered a dialogue. A letter to the editor is a one-time reader response to an article and undergoes the magazine's editing process in all cases. A reader cannot be sure that his letter to the editor will be printed and may not expect a response.

4.1.3 Function

The nature of language use is closely related to the demands that are made on it - the functions it has to serve. The functions of business and news magazines differ and these differences are relevant to the analysis of the corpora. Assigning communicative purposes to a text is neither easy nor straightforward, but journalistic texts perform certain conventional functions in society. All magazine articles are texts with the primary communicative purpose to inform, report and persuade. These texts have various functions which serve the interests of writers and readers: to reflect, shape and lead public opinion (cf. Crystal/Davy 1969). Business or specialist magazines serve additional functions. They are written by a

specialist journalist for a reader in the field. It is a periodical which provides a business, industry, or specialized profession with information. They "accumulate and preserve specialist knowledge, […] and reveal the personal attitude of the text author to his/her subject and to the addressee" (Gläser 1998: 469). This emphasizes the significance of 'function' as one distinguishing factor between the corpora or different magazine types.

4.1.4 Subject matter

The general or news magazine is a periodical that usually contains a miscellaneous collection of articles, stories, and pictures from various fields of interest. The majority relate to arts and entertainment, politics, general news, sports, lifestyle and health topics. Business magazines provide information on a comparatively narrow spectrum of recurrent topics. The key areas of coverage in business magazines include national and international business and finance. Business magazines provide comprehensive information on the economy, economic policies, trade, banks and corporations, management, politics, regulation, capital, commodity markets, the stock market and private investment, innovation, and technology. A news magazine's editorial scope is much broader than that of a business magazine and they target different reading publics. This introduces 'audience' as a primary influence on the subject matter of a journalistic text.

4.1.5 Audience

Each mass communication medium has a target audience which has a bearing on the subject matter, functions and language of a business or news magazine. The journalist assumes an 'ideal audience' or an 'ideal reader', including their attitudes towards the subject matter and often their social class, age, gender and ethnicity (cf. Kress 1988: 107, 183). The concept of audience is critical in the choice of language used in a journalistic text. Most of the studies on journalistic texts even argue that the audience is the primary influence on the linguistic features used in the text. This suggests the importance of attracting and maintaining the audience's attention in getting the message across (cf. Mohamad Ali 2005: 54). Consequently, awareness of the reader's strengths and weaknesses always informs the writing (cf. Michaels 1981: 311).

This study involves two types of contrasting: first, contrasting collocations in American English versus German, and secondly, contrasting the results across magazine types, that

is, business magazines versus news magazines. The review of the factors ‘origin’, ‘function’, ‘subject matter’ and ‘audience’ above has underlined their significance for classifying the texts collected for the corpora in this study. Although there is always certain variability within a set magazine type, ‘origin’, ‘function’, ‘subject matter’ and ‘audience’ serve as distinguishing factors for business and news magazines.

The following section reports on the non-linguistic differences between the two types of magazines and include a discussion on magazine publishing as a process and profession closely related to the focus of the research.

4.2 The general-specialized dichotomy of magazines

4.2.1 Overview

Magazines can be classified as special interest or general interest items. The specialized magazine exists because the general magazine cannot possibly find space to gratify all readers’ varied and intense interests (cf. Wolseley 1985: 315). Naturally, the magazine type determines much of the shape of its articles. For a more general circulation publication, there is a greater need for explanations and items to attract the reader’s attention. For a more sophisticated audience, the emphasis shifts to greater depth and quantity of information.

The general-specialized dichotomy of magazines frequently refers to the nature of their audience. The scope of the editorial content is also very important to distinguish specialized from general magazines. Another distinguishing factor is how media language accommodates readership or differs according to editorial content (cf. Bell 1991). These factors account for the complexity of variation of journalistic texts. They are highly interrelated, since editorial constraints reflect perceptions of readers’ preferences and their ability to understand complex subjects. Such perceptions are based on extensive research concerning readership conducted by various agencies.

A general interest magazine is typically one with a demographically and/or geographically diverse audience. On the contrary, specialized magazines have more narrowly focused audiences. Assigning the status of ‘specialized’ is based on audiences’ shared interest in a specific subject matter or audiences’ shared demographic and related characteristics. Magazines can also be divided into ‘general’ and ‘specialized’ publications based on the sheer size of their audience. In this regard, the term ‘general’ is sometimes used interchangeably

with ‘mass’. A very good indicator of whether a magazine is a general or specialized interest publication is its circulation. The terms ‘general’ and ‘specialized’ are also used to refer to a magazine’s editorial scope, whereas ‘general’ and ‘specialized’ refer to the range or diversity of topics included in the magazine.

The general-specialized dichotomy is applicable to the present study insofar as news magazines have a larger, more demographically diverse audience than business magazines. This is owed to their broad editorial scope, which is more appealing to the general public than that of a business magazine. The printed editions of all magazines that are analyzed in this study, namely *BusinessWeek*, *WirtschaftsWoche*, *Newsweek* and *Der Spiegel*, are available nationwide in their countries of publication. Beyond this, they all maintain a website that offers most of the magazine’s content online, worldwide. Geographic diversity of their audiences is thus relevant only to a limited extent. However, printed editions of *Newsweek* are published in English-speaking countries around the world other than the US. *Der Spiegel* publishes English summaries of international news on their website, which allows for the conclusion that there is an interest in the magazine’s content outside the country and outside the community of native German speakers. *BusinessWeek* and *WirtschaftsWoche* also publish nationwide editions in the US and Germany. While most of their content is available on the internet free of charge, printed editions in other countries than their country of origin do not exist. *WirtschaftsWoche* does not publish any content in English. This suggests that the selected news magazines may indeed also have a geographically broader audience than the business magazines in this study.

Categorization necessarily leads to some kind of abstraction. The general-specialized dichotomy subsumes all periodicals into one of two groups; all magazines that are not general are specialized. Although this approach may be criticized (cf. also Abrahamson 1995: 13), for the present study, it is useful to classify the corpora by grouping the magazines into general and specialized interest categories. This helps to organize the language data under investigation and to provide information about magazine dimensions.

4.2.2 News magazines

A news magazine as a written form of mass communication is intended for a large audience and has a multiplicity of purposes. It is a general interest magazine, addressing a broad and general reading audience. Journalistic news stories in news magazines are published frequently and distributed widely in order to spread news of current events. Hence, the stories and issues must be presented in a simple form for easy comprehension.

The news magazine is supposed to supplement the newspaper, not replace it. News magazines are important because of the perspectives they offer, and at times they also attend to news omitted by newspapers or broadcasters due to reasons of lack of space or time, policy, presumed lack of reader interest and the like (cf. Wolseley 1985: 273). News magazines summarize, explain and evaluate news. The news magazine journalist cannot and does not need to hold the same level of expertise in a specified area of reporting as a business magazine journalist. This is determined on the basis of the extensive scope of topics covered and the wide-ranging audience to which the news magazine journalist must appeal. The topics in newsweeklies typically range from politics, business, science, the arts, the courts, personal achievements, sport, travel, entertainment, fashion and social affairs to descriptions of the latest in a variety of topics.

News magazines also report on current events in the business world, but the business sections are generally kept rather short and the style of writing, the amount of technical vocabulary, and the depth of coverage in a news magazine and a business magazine are different from one another. The articles found in business magazines mostly aim at the actual business community to both give information and perhaps influence the economic development of the country. These need to be distinguished from the business section articles in news magazines which aim at the general public. Their main purpose is to give information on business events of note. This was also found in work carried out by Posteguillo and Palmer (1997), who discussed the use of language in business articles found in business and news magazines. They found significant differences, which they attributed to the different intended audiences of the articles; specific readership of business people on the one hand, and a wider, less business aware audience on the other. Consequently, there is a gradation of difficulty in the articles – from simpler language for the general public, moving to more complex for the business people themselves.

4.2.3 Business magazines

Business magazines are specialized publications. They are also called ‘economic journals’. Business magazines reflect the entire spectrum of commerce, and their market changes as new industries or businesses emerge. Their influence is great among the decision-makers of companies and government.

Out of all categories, the magazines of business are the largest single group in the magazine world (cf. Wolseley 1985: 315). A business magazine is devoted to the technical, industrial,

and business interests of its readers. Business magazines deal with manufacturing and marketing operations. They provide news on each industry and all events which affect it directly or indirectly. This includes employees and governments, civil or military happenings. They report on developments in science and technology, the research lab and the working day, and they bring ideas and information for the attentive reader (cf. Ford 1969: 114/116).

As discussed in section 4.1, the editorial scope and nature of the audience are significant to determine any given magazine type. Schenk and Rössler (1997) come to the conclusion that the range of topics treated in various business magazines is the same or extremely similar, targeting a restricted audience. Business English in journalistic texts is tailored to specific readers who understand the meaning of technical expressions and whose interests are being served (cf. Cameron/ Low 1999: 338). Laymen rarely find comprehensible and relevant information for their interests in business publications. Periodicals dealing with matters of general current interest, such as news magazines, serve this purpose much better. However, even within news magazines it is obvious that the business sections are more or less written for experts, specifically for those who are employers, producers, property-owners or stockholders (cf. Schenk/Rössler 1997: 17).

There have been many studies on the language of business in both written and spoken communication. Studies in Business English have looked into linguistic and non-linguistic issues. They have investigated the specialized words and terminologies that differentiate it from everyday language (cf. Pickett 1986, 1989). This work cuts across many fields like register, discourse and genre analysis, corpus linguistics, lexical, collocational and multi-word item studies. However, limited attention has been paid to written business texts for knowledge dissemination such as journals, magazines and newspapers (cf. Mohamad Ali 2005: 27).

Business journalism ideally embodies a balanced mix of technical knowledge, linguistic proficiency, logical reasoning and emotionality. This has been earlier referred to as ‘laymen-related technical language’ (cf. Möhn/Pelka 1984). Gunnarsson (cf. 1987: 78) classified business journalism as popular scientific literature. By contrast, Näslund (cf. 1987: 126) regards specialized journalism as a hybrid format, placed on the periphery of technical communication. ‘Audience’, while being an extra-linguistic criterion for text classification, also determines lexical choice. The criteria ‘origin’, ‘function’ and ‘audience’ do not give rise to separate, special languages as such, but they restrict language choice and allow for a certain amount of specialist lexis to be found in business magazines.

This is of central importance to the following section, which deals with internal, language-immanent, characteristics of the corpora, specifically lexis. Because collocations operate on the lexical level, corpus classification according to external criteria is not sufficient.

4.3 Internal characteristics of the corpora

An easily definable body of lexis and grammar prototypical of the texts in business or news magazines does not exist. The language of general news reporting is not homogenous, and neither is that of business and trade. 'Business', however, is not only a profession, but also a matter of public debate in politics and media, more so than other areas of expertise (cf. Burger 2007: 173). 'Business English' can be seen as distinct from general English though at the same time is still quite clearly attached to it (cf. Nelson 2000).

This study argues that business journalistic texts are different from other journalistic texts and that business and news magazine articles represent general language usage versus specialized language. Moreover, specialized discourses like business journalism have their own linguistic preferences and communicative functions in disseminating information. Variation in specialization also leads to a distinct lexis. In order to assess what contributes to this division and to characterize business journalism more closely, it is mapped against descriptions of technical discourse and popular scientific writing in the following subsections.

4.3.1 Features of technical discourse

Generally speaking, technical text is meant to serve as a resource of clear, effective and suitable communication on technical matters (cf. Gläser 1990: 6). Technical text is a cohesive body of text written for a specific purpose with at least one primary communication goal in a particular field or discipline (cf. Haramundanis 1998, Miller 1991, Walker, 1996). Technical writing is most strongly associated with the various fields of science (cf. Martin 1993a: 206 ff.). At first, this exclusively referred to writing in the fields of technology, engineering and the natural sciences. In recent times, it has come to include writing done in any profession (cf. Pfeiffer 1998: 8). Technical text tries to teach or to inform the reader. The writing of technical text requires sophisticated knowledge of both technical and non-technical language and a sense of how to tailor language to the needs and capabilities of specific audiences (cf. Huckin/Olsen 1991: 83). Technical text is generally written for, read by and published in literature supported by practitioners of its own disciplines (cf. Vatavuk

1992: 1 f.). This study argues that business journalism provides the authenticity of a specialized or technical discourse.[15]

Journalists in general and the business journalist in particular communicate in a specialized field. They communicate their messages along with particular communicative intentions. Besides the overall structure of news writing, other persuasive textual and rhetorical styles are used (cf. Biber 1995). The content of business journalistic articles may reflect the communicative intentions of business writers, which are persuasive and purposive in nature, as there is the intention of either guiding readers in the same field or building the image of the business and management experts. Lexis also plays a crucial role in the level of specialization and persuasion in journalistic writing. The concept of technical discourse as "languages that possess different lexical, grammatical and other linguistic features than ordinary language, and which are used for particular and restricted types of communication" (LDAL 1985: 159) thus, is well suited to describe business journalism.

In contrast to the articles from business magazines, the discourse of news magazines is not a technical one. Relatively few technical terms are used, and where they are used they tend to be borrowed from other disciplines rather than established by the news discourse itself. Technical terms have the function of accumulating a number of less specialized meanings in a single lexical item, although meaning accumulation is just one aspect of their function (cf. Martin 1993a: 229). The amount of technical vocabulary and how it is introduced is one of the decisive criteria when distinguishing between general and specialized discourse. As mentioned earlier, discourse purposes are primarily characterized by 'who communicates to whom' (cf. subsections 4.1.3 and 4.1.5, also Schröder 1991: 6). In news magazine articles, difficulties in terms of understanding do not arise for the reader, because language is kept at a level at which knowledge of the vocabulary is assumed. This accounts for the typically non-technical nature of news magazines. It is important to stress that the interpretation of technicality being developed here shows that technical terms are fundamental to differentiating general from specialized discourse.

[15] In the 1980s, the term 'Language for Specific Purposes' (LSP) was introduced to describe the kind of language use linked with specialized communication in a wide variety of subject areas (cf. Swales 2000, Gläser 1998). At times, 'Business English' is also referred to as LSP. The present study does not describe Business English with terminology coined in LSP research, because it is mainly used in the environment of teaching and pedagogy and thus not very suitable for the purpose of the present study.

Gläser (cf. 1990: 46-47) describes the concept of technical discourse in more detail and develops criteria to distinguish between internal and external technical communication. Whereas internal technical communication requires appropriate professional competence of the communication partner, external technical communication is hallmarked by a decrease in technicality mainly due to popularization strategies (cf. ibid.: 71). Such strategies may be reflected by degree of abstraction and formalization, the consequent use of technical vocabulary or the abandonment of formulae. Articles taken from *WirtschaftsWoche* and *BusinessWeek* qualify most likely as external technical communication, more specifically, they are best classified as popular scientific magazine articles (cf. ibid.: 183 ff.).

4.3.2 Features of popular scientific writing

Business journalism qualifies as technical discourse, but this does not necessarily equal 'scientific' or 'popular scientific' discourse. First of all, the scientific status of economics is debated among experts and theorists of different disciplines. Yet, as an essential prerequisite, the present study assumes economics to be a factual and applied discipline which relies on descriptive theories and models constructed by economists. Hence, articles from economic journals may qualify as popular scientific discourse, put forward by a journalist exercising economic reporting, covering areas such as finance, technology and management and possessing expert knowledge in those areas. Business journalism thus qualifies as external technical communication and popular scientific writing.

Popular scientific writing appears in a variety of sources, ranging from news-rack publications to university presses. Science popularizations are anything but intellectually jejune (cf. McRae 1997: 10). The label 'popular scientific' unfortunately connotes what Richard Whitley rejects in the usual attitudes about science popularizations, namely, a demarcation between "a structured intellectual elite of knowledge producers" and a "diffuse mass of ignorant knowledge consumers" (1985: 6). Such a distinction trivializes the readers of science popularizations as much as it degrades their authors (cf. McRae 1997: 11).

The business journalist, when writing an article, acts on the assumption of different levels of understanding of 'the economy'. The interaction between these different frameworks of knowledge – that of the journalist (who either is an expert himself or might rely on an expert in the field) and that of the audience - creates understanding of and for the reader, to put it to use within the informational and evaluative frameworks at their hands. When creating such texts, the intended audience dominates the author's intention (cf. Kelley

1997: 148). Business magazines are primarily concerned with economic and financial reporting, but its narrative practices profoundly shape the articles. Although the information provided remains scientific, its format changes as it moves from professional to popular expression. Scientific research is translated for the popular audience and modified accordingly. While their emphasis remains dominant, facts do not speak for themselves. The narrative around the raw data is increasingly cast in the form of an argument (cf. Russell 2010: 10).

Business journalism can be characterized as projecting a particular image: that it is an objective, efficient conveyor of cognitive complexity. Under the rubric of objectivity, those aspects of style that focus attention away from people and toward things are placed. Examples are the increased use of passive forms and verbs which relate to the activities of things rather than people and the relative absence of personal pronouns. Typically, the results of scientific research rather than the data on which they are based are foregrounded. Technical verbs are scarce, and those that exist are rarely used. Nominalization is used to facilitate classification and the vast majority of technical terms are nouns (cf. Martin 1993b: 213). At clause rank, the meaning of technical terms is constructed through use of the relational identifying clause. This is the favored clause type in scientific discourse for what are commonly referred to as ‘definitions’ (cf. Halliday 1985: 112 ff.). Gross et al. (2002: 215) identified another set of features that make syntax more efficient:

> Increasing cognitive complexity has forced the shift in scientific [texts] to more complex noun phrases in the subject position; as compensation, there has been a corresponding decrease in the length of sentences and in the number of clauses per sentence.

In response to the high volume of publications in the field of economics, a mode of presentation most likely to get the point across efficiently is adopted. This includes presentational and organizational features of popular scientific articles that also promote reading efficiency. Among these are introductions, headings, the numbering of figures, the format and placement of citations, and finally, conclusions. These features promote efficiency by permitting readers to discover as quickly as possible whether they need to pay attention to an article, and if so, on which constituents of the article their attention should dwell. The majority of these observations apply regardless of whether their source is the social sciences, the humanities, or the natural sciences.

All in all, the business magazine corpora in this study represent economics literature. They display characteristics of technical language embedded in journalistic writing and qualify

as popular scientific writing because of the communication and popularization strategies used by the business journalist.

This study researches and contrasts the peculiarities of the collocational behavior of Anglicisms in news versus business magazines. *Newsweek* and *Der Spiegel, BusinessWeek* and *WirtschaftsWoche* were selected as their representatives. The size of the corpora collected from each magazine allows for a detailed analysis and a high level of explanatory and predictive capability. Each magazine can be taken as an independent variable, where the dependent variable or focus of investigation is a linguistic feature, collocation, narrowly defined. Details on the concept of 'comparable corpora' and the requirements to successfully work with them are presented in the following section.

4.4 Working with comparable corpora

Comparable corpora are sets of texts from pairs or multiples of languages which can be contrasted and compared because of their common features. They have been defined as collections of "texts which, though composed independently in the respective language communities, have the same communicative function" (Laffling 1992: 20). Thus, working with comparable corpora means that the texts chosen are independently created and typically unrelated, except by the analyst's recognition that the original circumstances that led to their creation have produced similarities. In the present study, the corpora achieve their comparability through similar extra-linguistic criteria (cf. section 4.1). The construction of suitable comparable corpora is a fruitful way of enabling contrastive language studies**.**

Contrastive linguistics typically differentiates between two main types of corpora:

1. Corpora consisting of original texts in one language and their translations into one or more languages – translation corpora.
2. Corpora consisting of original texts in two or more languages, matched by criteria such as the time of composition, text category, intended audience, etc. – comparable corpora.
   (cf. Johansson/Hasselgård 1999: 148)

Besides being fairly easily accessible, comparable corpora have the major advantage of representing original texts. The main drawback lies in the difficulty of establishing comparability. This is precisely why the comparison between corpora of original texts in different languages is regarded as particularly difficult (cf. Granger 2003: 22).

Comparable corpora provide data to study how the same concept can be rendered independently in different languages. Structures to be contrasted have to be described by means

of the same theoretical model (Sajavaara 1981: 54). Contrasting in the present study is carried out by investigating collocations in all corpora according to the same concept of collocation (cf. section 2.4). Comparable corpora always concern a restricted sub-language (cf. Peters et al. 2000: 74). They are processed to obtain information on cross-linguistic natural language lexical equivalents within a given domain.

Due to the nature of this study, the characteristics of the texts that were collected for the corpora are, to some extent, predetermined. They are all journalistic texts, with the primary purpose of being informational. Despite differences which may exist between the respective cultures and languages the texts were produced in, it is claimed that there are types of texts in each language which perform approximately the same function (cf. James 1980: 117 ff.). Magazine articles perform certain conventional functions and can be labeled as more or less institutionalized. At the same time, the texts selected from news magazines vary in their origin, subject matter, communicative intentions or functions and audience when compared to those collected from the business magazines (cf. section 4.1).

Naturally, there are a vast number of business and news magazines from which to choose for analysis. Generally, choice is based on the comparability of the magazines' overall content, informational intent and readership. *Newsweek* and *Der Spiegel*, *BusinessWeek* and *WirtschaftsWoche* were selected as representatives for the present study. To qualify as comparable corpora for each magazine type, they must share several matching criteria. The magazines chosen to represent each publication type here are similar in several major respects. These are their self-perception and how they are perceived by readers, their communicative intentions, their subject matters and the space, location and priorities they assign to their news categories, the demographics of their audience, their appearance and layout, their circulation, price and points of sale. All corpora are synchronic and consist of complete texts as opposed to text samples.

In an effort to present the minimized variables, facts about the data collection process and the magazines and their matching criteria are provided in sections 4.5 and 4.6.

4.5 Data collection for the business magazines

In addition to print editions, the content of most newspapers and magazines is available on the internet. Data for the business magazine corpora were gathered online from www.businessweek.com and www.wiwo.de. In this study, lexical analysis software is applied. Language data, therefore, need to be formatted accordingly. The magazine articles were accessed

according to publication date. In order to apply automatic retrieval software, text (.txt) files without formatting information were created, one file per magazine and month.

Text acquisition was slightly complicated by a restrictive copyright situation. Staff of both business magazines confirmed that the digitally published content is not entirely identical with the content of a corresponding printed issue. This results from individually reached licensing agreements with all journalists and correspondents. *BusinessWeek*'s online archive offers full tables of contents for each print issue dating back as far as 1995 and hyperlinked headlines are available online. Registering is not required and no fees are charged. Access to the digital archives of *WirtschaftsWoche* is granted only in return for registering and the payment of a monthly fee. Full tables of contents according to print editions are not provided. Articles are grouped according to publication date and archiving dates back as far as 1989.

In addition to making articles available, both archives hold advertisements and diagrams. Digital formats like 'blogs' and 'slide-shows', composed by editorial staff, as well as letters to the editor from print editions and feedback posted online by website visitors are stored as well. However, the analysis in this study is designated to popular scientific business magazine articles. Therefore, all the above additional formats and interviews are excluded from the corpora. Moreover, the initial hypotheses of this study (cf. section 1.1) are pinned partly on the assumption that similar training and collaboration within editorial departments blurs idiolects, paving the way for augmented probability of discovering analogies. For this reason, the compilation of the corpora excluded anything but the text portion of magazine articles composed by editorial employees.

The business magazine corpora compiled for the current study consist of all articles which were published in 2008 in *BusinessWeek* and *WirtschaftsWoche*. The time frame was set to one year for two main reasons. First, it is important to avoid misrepresentation, due to the regular review of certain occasions at a given time of the year in the magazines and the vocabulary associated with it. Nevertheless, predominance of a certain topic (and the vocabulary associated with it) is also observable across a year, considering, for example the fact that 2008 was an election year. Overall, such prevalence is of less significant consequence when using larger corpora. Secondly, the size of a corpus primarily depends on the type of research (cf. section 3.3), and lexical studies require larger corpora than are needed for studies of syntactic features (cf. Esser 1999: 157).

The business magazine corpora in this study hold 4,168,513 million words, collected from 50 issues of each magazine. Considering the restrictions of the business magazines' domain coverage and the specialized language, this represents a reasonably large portion of the

subset of language under study. The corpora are considered representative of leading business journalism and popular scientific magazine articles in Germany and the US with respect to aspects such as content and language use (particularly lexis). The word[16] count of the business magazine corpora, according to month and magazine, is summarized in Figure 8 below.

MONTH	*BusinessWeek*	*WirtschaftsWoche*
JANUARY	80,203	218,047
FEBRUARY	108,482	224,055
MARCH	145,498	274,234
APRIL	146,544	223,953
MAY	110,716	215,731
JUNE	117,068	273,359
JULY	124,262	190,045
AUGUST	78,465	210,330
SEPTEMBER	110,875	280,066
OCTOBER	122,222	215,217
NOVEMBER	123,916	216,868
DECEMBER	145,435	212,922
TOTAL	**1,413,686**	**2,754,827**

Figure 8 Word count for *BusinessWeek* and *WirtschaftsWoche* corpora

The fact that the target language corpus is almost double the size of the source language corpus is a great advantage for the present study, i.e. more instantiations of Anglicisms occur. Because the frequencies of a certain node (key word) will not be contrasted proportionally to running text, problems for statistical analysis do not arise. That is to say, whether a given collocation occurs in the source or target language corpus more or less often in relation to the size of that corpus is not important.

In order to complete the assessment as comparable corpora, separate descriptions of both business magazines follow in subsections 4.5.1 and 4.5.2. This leads to the definition of the business magazine corporas' matching criteria in subsection 4.5.3.

[16] Here referring to a chain of letters between two blanks.

4.5.1 BusinessWeek

In September 1929, just before the Great Depression, the major industrial publishing house McGraw-Hill[17] decided to launch a new business weekly – *BusinessWeek*. The magazine was founded to provide "essential information and insight that help individuals, markets and societies perform to their potential" (www.businessweek.com). *BusinessWeek* is "[…] published for management. Each week it reports the news, the ideas and the trends that have an impact on the economy or industry – or that can provide new insights for the business executive in the operation of their own business" (Ford 1969: 137).

Like many print publications, *BusinessWeek* suffered considerable economic strain as a result of the recent financial crisis. In October 2009, *BusinessWeek*'s publishing house McGraw-Hill announced that it was selling the magazine to Bloomberg after a months-long sales process. Bloomberg L.P. is a privately held company started by New York Mayor Michael Bloomberg. Bloomberg runs the largest private network in the world which is a source of knowledge for businesses, governments and professionals – known for its financial focus. *BusinessWeek* was renamed *Bloomberg BusinessWeek* during relaunch in 2010. The magazine's website remains available at www.businessweek.com. Josh Tyrangiel was recently appointed editor of *Bloomberg BusinessWeek*. He succeeds editor Stephen J. Adler, who announced his resignation as editor-in-chief in October 2009. The change in ownership and editorial scope of the magazine does not affect the data collected for the study and their classification. The texts for the corpora were all written and published in 2008, when *BusinessWeek* still belonged to McGraw-Hill and Adler was still in charge.

In 2008, the average weekly circulation amounted to 933,566 copies within the US. *BusinessWeek*'s editorial staff consisted of 155 employees at its New York headquarters and 19 correspondents distributed among nine bureaus in the USA. Each issue of *BusinessWeek* published under Adler and his staff features in-depth perspectives on business, corporate leadership, finance, innovation, media, marketing, news, regulation, small business, sustainability and technology. Since 1994, the website has provided the majority of articles from print editions electronically; stored in a searchable archive. Figure 9 below summarizes information on online readership.

[17] The Corporation is a leading global information services provider. Additional information is available at www.mcgraw-hill.com.

READERSHIP OF www.businessweek.com	
DEMOGRAPHICS	
female	32%
male	68%
average age	49
EDUCATION	
College+	67%
HOUSEHOLD INCOME($)	
average	95,256
100,000+	39%
OCCUPATION	
C*-Level; EVP/SVP/VP**; Senior Management; Business Decision Makers	63%
* Chief-Officer **Executive Vice President/ Senior Vice President/Vice President	

Figure 9 Readership of www.businessweek.com

Because the corpus texts were collected from the *BusinessWeek* website, it was deemed more meaningful to give information on website visitors, rather than on the readership of print editions; admittedly they coincide significantly. *BusinessWeek* is written to the big business beat. World trade and statistics occupy foreground and background. Typically, the writing moves in short sentences.

4.5.2 WirtschaftsWoche

Originally, the German economic journal was founded in 1926 as *Der deutsche Volkswirt* and was renamed *WirtschaftsWoche* in 1970. A gradual transformation of the magazine from a highly theoretical specialized business publication to a modern business magazine began. *Der Volkswirt* was intended to become a bridge between scholarship and practice. This bridging function was apparently well received by the readers, as its circulation tripled between 1967 and 1971. It was designed to combine all the advantages of two American business magazines: *The Economist* and *The Nation.* Consequently, it intended to be the

> […] premier source for the analysis of world business and current affairs, providing authoritative insight and opinion on international news, world politics, business, finance, science and technology, as well as overviews of cultural trends and regular industry. (www.economist.com)

At the same time it would have to be a

> weekly journal of opinion, featuring analysis on politics and culture and making an earnest effort to bring to the discussion of political and social questions a really critical spirit, and to wage war upon […] exaggeration and misrepresentation by which so much of the political writing of the day is marred. (*The Nation's* founding prospectus at www.thenation.com)

Today *WirtschaftsWoche* is certainly a major business magazine with the longest history in this field in Germany. Currently, it is Germany's most quoted[18] business magazine and it is actually sold and read more than competing magazines in the market. *Wirtschafts-Woche* competes with numerous influential business magazines like *Capital* or *Manager Magazin*. *WirtschaftsWoche* prides itself on being the most up-to-date business magazine, because it is published weekly as opposed to its major competitors, which are published either bi-weekly or monthly.

WirtschaftsWoche is exclusively geared to leaders in business and politics and to employers. Its publisher, Handelsblatt Verlagsgruppe GmbH, states that "[i]t combines topical with analytical coverage for executives and decision-makers in Germany and German speaking Europe" (www.wiwo.de). *WirtschaftsWoche* provides comprehensive background information on the economy, industry, trade, banks and corporations, management, politics, the stock market and private investment, innovation, and technology. Detailed special reports cover innovative areas such as multimedia, telecommunication and biotechnology.

Editor-in-chief Roland Tichy stands by the fact that *WirtschaftsWoche* has a clear political orientation and that this is one way in which the magazine distinguishes itself from its competitors. *WirtschaftsWoche* has an editorial team of 120 journalists with correspondents all over the world. In 2008, its weekly circulation added up to an average 186,000 copies. The *WirtschaftsWoche* website was launched in August 1996. A look at the website offers initial insight into the extensive use of Anglicisms within the magazine. Editorial

[18] *WirtschaftsWoche* was the business magazine to be quoted most frequently in 2008 and 2009 (http://de.statista.com/statistik/daten/studie/30671/-umfrage/die-meistzitierten-printmedien-deutschlands).

policies which would influence the use of Anglicism are unknown. Figure 10 below provides some facts about the readers of the *WirtschaftsWoche* website (www.iqm.de/medien).

READERSHIP OF www.wiwo.de	
DEMOGRAPHICS	
female	21%
male	79%
average age	47
EDUCATION	
College+	82%
HOUSEHOLD INCOME(€)	
average	84,000
100,000+	31%
OCCUPATION	
C*-Level; EVP/SVP/VP**; Senior Management; Business Decision Makers	55%
* Chief-Officer **Executive Vice President/Senior Vice President/Vice President	

Figure 10 Readership of www.wiwo.de[19]

4.5.3 Matching criteria

It was explained already in section 4.4 why the definition of matching criteria is essential when working with comparable corpora. Some of the information provided in the previous subsections on *WirtschaftsWoche* and *BusinessWeek* was taken from their own websites; therefore, it may not be truly unbiased. Still, the numerous comparisons reveal significant similarities.

BusinessWeek and *WirtschaftsWoche* are both weekly business magazines with a long tradition. Moreover, *WirtschaftsWoche* was modeled on the standards of two American business magazines: *The Economist* and *The Nation* (cf. also previous subsection 4.5.2).

[19] The readership of the business and news magazines in this study is evaluated according to the same criteria. For some magazines more than the displayed characteristics are available. The study focuses on such criteria which are equally accessible for all magazines for the benefit of facilitating comparison.

BusinessWeek and *WirtschaftsWoche* are specialized magazines. The external corpus classification criteria 'origin', 'function', 'subject matter' and 'audience' coincide. Both magazines wish to influence markets while being a solutions provider. *BusinessWeek* and *WirtschaftsWoche* have similar audiences. Their readership is typically assumed to have enjoyed the benefits of higher education, holding advanced general knowledge and voluntarily engaging in intelligible technical literature. Consulting Figures 9 and 10, www.businessweek.com and www.wiwo.de obviously have an almost identical readership. On average, readers are male, middle-aged, well-educated and affluent. On their websites, both magazines claim to address the global leaders of today and tomorrow, thought leaders, policy influencers and senior executives.

Both corpora display characteristics essential to business journalism, technical discourse and popular scientific articles. Thus, they can be assumed to make comparable use of specialist lexis or technical language and are similar with respect to style, argumentation and presentational, textual and syntactic features. The two magazines feature nearly identical areas of reporting, i.e. sections. These structural categories that the magazines provide sometimes differ in name slightly, but an overall evaluation of the magazines allows for the conclusion that *WirtschaftsWoche* and *BusinessWeek* focus on the business environment and economic and political conditions. Both magazines draw on large resources of editorial employees and international correspondents who are experts in the field of economic and financial reporting and possess extensive technical knowledge on these matters. The noticeable difference in the magazines' circulation can be explained by the significantly different size of their distribution areas. Circulation in 2008, in relation to the population of the countries of distribution, is comparable:

COUNTRY	MAGAZINE	CIRCULATION	POPULATION
USA	*BusinessWeek*	933,566	304.72 million
Germany	*WirtschaftsWoche*	186,000	82.12 million

Figure 11 Circulation of *BusinessWeek* and *WirtschaftsWoche* in 2008

4.6 Data collection for the news magazines

Data for the news magazine corpora were gathered from www.newsweek.com and www.spiegel.de. Not all articles from *Newsweek*'s print issues are available on www.newsweek.com, but missing data could be retrieved from www.factiva.com. *Factiva* aggregates content from both licensed and free sources. It provides access to more than 28,500 sources such as newspapers, journals, magazines, television and radio transcripts etc. from 200 countries in 25 languages,

including more than 600 continuously updated newswires. *Factiva* offers all articles of the US *Newsweek* edition from 1994 onwards. *Der Spiegel'*s online archive offers full tables of contents for each print issue published, dating back as far as 1947. The content has been digitalized, including the covers. All articles are available free of charge.

In the USA, *Newsweek* competes with *Time* and *U.S. News & World Report* in the newsweekly category to serve the reader who wants a weekly news roundup**.** During 2008 and 2009 *Newsweek* undertook a dramatic restructuring of its business (cf. Pérez-Peña 2009 NYT, Deveny 2009 TDB). The magazine repositioned its content towards opinion and commentary beginning with its May 24, 2009 issue, because it was citing difficulties in competing with online news sources to provide unique news in a weekly publication. Since the changes in ownership and editorial content took place after 2008, they did not affect data collection or corpus classification, and *Newsweek* was assessed a suitable comparable corpus. This will be stated again in more detail in subsections 4.6.1 and 4.6.3. Figure 12 below shows the word count for the *Newsweek* and *Der Spiegel* corpora.

MONTH	*Newsweek*	*Der Spiegel*
JANUARY	275,835	268,908
FEBRUARY	131,632	256,967
MARCH	174,178	322,722
APRIL	140,016	254,297
MAY	130,858	263,141
JUNE	177,557	334,609
JULY	98,970	257,361
AUGUST	92,231	257,186
SEPTEMBER	182,176	260,573
OCTOBER	147,394	259,296
NOVEMBER	133,384	277,332
DECEMBER	138,963	320,392
TOTAL	**1,823,194**	**3,332,784**

Figure 12 Word count for *Newsweek* and *Der Spiegel* corpora

The editing and formatting process of the news magazine articles is identical to that of the business magazine corpora (cf. section 4.5) and resulted in 24 .txt files; one file per magazine and month.

A total of 5,155,978 words were collected to form a representative corpus of German and American leading news magazine journalism with respect to content and language use. Again, the analysis in this study is limited to business and news magazine articles. Thus, analogous to the process for the business magazine corpora, letters to the editor and all additional editorial content designed and published exclusively on the internet were excluded from the news magazine corpora. Along the lines of resembling the contrastive schematic structure of business and news magazines and to create an ideal foundation for contrastive linguistic research of specialized versus general discourse, the business sections of *Newsweek* and *Der Spiegel* were excluded and do not form part of the corpora.

4.6.1 Newsweek

News-Week was first published in February 1933, by former *Time* editor Thomas J.C. Martyn. In 1937, *News-Week* merged with the weekly journal *Today*, which had been founded in 1932. Malcolm Muir took over as president and editor-in-chief. Muir was the one who changed the name to *Newsweek* and launched international editions (www.newsweek.com). Over time, *Newsweek* developed a broad spectrum, from breaking stories and analysis to reviews and commentary. In 1961, *Newsweek* was bought by *The Washington Post Company.*

The magazine is published in New York City. It is distributed throughout the United States and internationally. For the current study, only articles of the printed US edition were collected. In 2008, about 135 journalists worked for the US edition. Even though traditional staffing had been cut, online staff kept growing (http://stateofthe media.org/2011/magazines-essay/data-page-4/). After 80 years in print, the news magazine adopts an all-digital format. The last print edition in the United States was the December 31 issue of 2012.

Newsweek's global network of correspondents, editors and reporters cover domestic politics and international affairs, society, the arts and entertainment, personal finance, living and lifestyle, celebrities, general news and health related topics with some lighter stories included each week. In 2003, worldwide circulation was more than 4 million, including 2.7 million in the US. By 2010 it was down to 1.5 million (with newsstand sales declining to just over 40,000 copies per week) and remains at this level.

In 2010, *Newsweek* had been losing money for the past two years and was put up for sale. In August 2010 the magazine was sold to Sidney Harman for $1 (Peters 2010 NYT). Harman kept the vast majority of *Newsweek*'s 325 employees in their jobs, including the magazine's top management and editors (www.huffingtonpost.com). Since he had bought

Newsweek, Sidney Harman was looking for a life-injecting top editor (www.observer.com). In November 2010 *Newsweek* merged with the news and opinion website *The Daily Beast*. Tina Brown, *The Daily Beast*'s editor-in-chief, became editor of both publications.

As early as 1993, *Newsweek* launched an internet edition of its weekly magazine. Figure 13 provides some facts on the readership of www.newsweek.com, which is very similar to that of the printed edition.

READERSHIP OF www.newsweek.com	
DEMOGRAPHICS	
female	44%
male	56%
average age	48
EDUCATION	
College	73%
College+	46%
HOUSEHOLD INCOME ($)	
average	79,032
100,000+	31%
OCCUPATION	
Business decision makers	27%
Senior Management	10%

Figure 13 Readership of www.newsweek.com

4.6.2 Der Spiegel

The magazine was first published in 1946 under the name *Die Woche*. In January 1949 the magazine was renamed *Der Spiegel* (cf. Stave 1960: 228). In the beginning, John Chaloner, a British World War II officer who wanted to establish a free press in occupied Germany, discovered and promoted *Der Spiegel* founder Rudolf Augstein (cf. Bölsche 2007 DS). Augstein took over sections from the *Time* magazine founding policy to characterize the aim of his publication. He wanted to publish a magazine that provided reportage instead of commentary (cf. Just 1967: 14). The influence of the British occupying power, which had issued the license for *Der Spiegel*, was distinctive. With the news magazine, a new type of magazine could be introduced in Germany (cf. Stave 1960: 228).

Since 1960, *Der Spiegel* has differentiated between the sections Germany, International Affairs, Sports and Culture. Besides this, topics are grouped under very generic terms

(cf. Just 1967: 67) such as Politics, Science, Health, Career and Technology. From 1951 on, the magazine deliberately refrained from introducing a business section. This was done in order to avoid the aversion of numerous readers to the business sections offered in dailies, because the business sections were presented as if there were only experts (Just 1967: 68). After the gradual growth during the first years, the circulation of the magazine jumped around 1961. The economic boom within the magazine market can only partly explain this. Circulation in 2008 was around 900,000. In Germany, *Der Spiegel,* unchallenged in the newsweekly category for almost 50 years, now competes with *Focus* and even *Stern* to some degree.

Der Spiegel does not only deal with the most important news of the past week or deliver a summary of it. The magazine also reports on latently topical issues. The magazine assumes knowledge of the general news. It deals with issues and events that have already been reported elsewhere. But the magazine points out new aspects or connections, a task the daily newspaper cannot fulfill (cf. Just 1967: 52). *Der Spiegel* has a clear political orientation, although this is true for most German news publications.

Der Spiegel, which is supposed to report more personally and provide more background information than the daily newspaper, selects current issues most likely to appeal to and concern a broad audience of averagely interested laymen. Carstensen (1971: 28) describes its typical readership:

> Was im Spiegel steht sollte geeignet sein, den bildungsfähigen Durchschnittsmenschen zu interessieren.

Carstensen (ibid.: 21) also notes:

> Der Spiegel schreibt nicht fuer Liebhaber und Spezialisten, sondern fuer interessierte Laien, seien es nun Arbeiter, Angestellte oder Direktoren.

Sackarndt (1961: 70) characterizes the magazine's readership as follows.

> Kein weiter Kreis Gleichgesinnter. Sie [die Leser-schaft] ist in der Sichtweite zwischen Hochintelligenz und Querulantentum, zwischen demagogisch Geführten und Opponenten, affektuösem und rationalem Typ, Sensations-publikum und politisch-wirtschaftlichen Sachinteressen-ten nicht einmal psychologisch genauer bestimmbar.

Der Spiegel uses puns and plays with words incessantly, although these tools have to be kept at a comprehensible level for the reader for the intended effect (cf. Carstensen 1971: 45). Figure 14 below provides some facts on the readership of the magazine's website.

READERSHIP OF www.spiegel.de	
DEMOGRAPHICS	
female	36%
male	64%
average age	48
EDUCATION	
College	47%
HOUSEHOLD INCOME (€)	
average	60,000+
OCCUPATION	
Managerial position	47%
Senior Management	14%

Figure 14 Readership of www.spiegel.de[20]

Because *Der Spiegel* is a popular medium of information whose editors are using the language so freely and boldly as to suggest conscious experiment, it has been subjected to linguistic research carried out in particular by Carstensen (1965, 1971) and Lück (1963). Because of its popularity, any word that *Der Spiegel* introduces and uses consistently may be considered at least potentially a part of the language. For this and other reasons, many linguistic studies on *Der Spiegel* have been conducted, many of them focusing on its lexis, particularly concerning Anglicisms. Research by Carstensen shows that the editors of *Der Spiegel* use Anglicisms intentionally to be more precise, and more importantly to produce effect (cf. Just 1967: 149). Carstensen even referred to the magazine as "Haupteinfallstor von Amerikanismen in die deutsche Sprache" (1965: 22). This is an additional reason why *Der Spiegel* is particularly suited as a comparable corpus for Anglicism research.

4.6.3 Matching criteria

The previous subsections on the two news magazines *Newsweek* and *Der Spiegel* reveal significant similarities.

Newsweek and *Der Spiegel* are both weekly news magazines with a long tradition. Moreover, *Der Spiegel* was modeled on the standards of an American news magazine.

[20] Quelle www.spiegel-qc.de/deutsch/partner__preise/der_spiegel/ index.php.

Newsweek and *Der Spiegel* are general interest magazines. The external corpus classification criteria ‘origin’, ‘function’, ‘subject matter’ and ‘audience’ coincide. Both news magazines wish to inform their readers beyond the capacities of a daily newspaper. *Newsweek* and *Der Spiegel* have similar readership. Figures 13 and 14 show that the audience of both magazines is male and female in equal shares. About half of the magazines’ audience is college educated and their average household income is comparable.

Both corpora display characteristics essential to news magazine journalism. The two magazines feature nearly identical areas of reporting, i.e. sections. An overall evaluation of the magazines allows for the conclusion that *Newsweek* and *Der Spiegel* focus on similar areas of reporting. Both magazines draw on large resources of editorial employees and international correspondents. The circulation of *Der Spiegel*, when put at an appropriate level to population of distribution area, is slightly higher than that of *Newsweek* but the magazines’ circulation may still be considered comparable. Figure 15 below illustrates this. It should be pointed out that the circulation of *Newsweek* has gone down rapidly during the last few years. This is owed to a general crisis in the magazine market in the US. In 2003, the circulation of *Newsweek* was still at about 4 million.

COUNTRY	MAGAZINE	CIRCULATION	POPULATION
USA	*Newsweek*	2,750,000	304.72 million
Germany	*Der Spiegel*	1,071,256	82.12 million

Figure 15 Circulation of *Newsweek* and *Der Spiegel* in 2008

5 Methodology

This chapter initiates the transition to the corpus analysis. It concerns the issues related to the empirical research. The approach of this study is empirical in that it is based on the analysis of authentic language data. This chapter outlines the methodology to contrastively analyze the collocational behavior of Anglicisms in a total of four corpora of written journalistic texts. Details on processing the corpora and different steps of analysis are presented. This chapter shows how the theoretical concepts described previously are applied to the analysis of the corpora. Section 5.1 defines 'Anglicism' as it is understood in this study. In section 5.2 the lists of key words for the analysis are provided and section 5.3 describes the lexical analysis software. The aim of the corpus analysis is to determine whether an Anglicism is borrowed by itself or if its collocational surroundings are borrowed as well. The analysis should lead to new insight into the collocational behavior of Anglicisms in business and news magazines and highlight some tendencies in coping with the two phenomena of collocation and borrowing.

5.1 Anglicisms

5.1.1 Previous studies of Anglicisms and the language of the press

The language of the press has long been a preferred environment to conduct research on Anglicisms. This is also evidenced by the large number of works published in this field. The earliest prominent example is Zindler's 1959 dissertation "Anglizismen in der deutschen Pressesprache nach 1945". Zindler's study was the starting point for research by Carstensen, who is still considered the scholar who contributed most to this field. In 1965, Carstensen published his monograph "Englische Einflüsse auf die deutsche Sprache nach 1945". His analysis was based on West German newspapers and magazines published between 1961 and 1964, among them *Der Spiegel*. Another important study focusing on English vocabulary in the German press was carried out by Fink. In 1968, he analyzed American loanwords in the German daily press (*Süddeutsche Zeitung*, *Frankfurter Allgemeine Zeitung*, *Die Welt*). This was the first comprehensive study based on a particular corpus. The first computer-aided quantitative lexical study was carried out by Engels in 1976. She examined the occurrence of Anglicisms in the German daily newspaper *Die Welt*. More recent studies include dissertations based on *Der Spiegel* (Yang 1990; Kovtun 1996) and work by Langer who researched the business magazines *Capital* and *Deutsche Mark* (1996). Another smaller study is Hedderich's (2003) investigation of recent language change in Business German, which is based on the examination of three successive issues

of *WirtschaftsWoche* from 2002. Thus, the present study is in line with a certain tradition of the use of print media sources for linguistic research on Anglicisms.

As already stated in the introduction of this study, the arrival of English language material in German has been a controversial issue for more than a century. Onysko (2004: 59) states on the recent influence of the English language the following:

> In recent times, increasing globalization, tools of mass communication such as the internet, and, specifically in Europe, the expansion of the European Union have all strengthened the position of English as an international language.

Crystal (2004: 79, 109) adds that English is the language of science and technology, the official working language of 85% of supranational organizations, and dominates the internet at about 80% of its language output. As a result of the great impact of English 'loanwords', the basic question of how to define the term Anglicism as a concept has gained fresh momentum (cf. Onysko 2004: 59). Basically, we can distinguish between direct and indirect loan influences. The aim of this section is to reach a workable definition of 'Anglicism' for the corpus analysis.

5.1.2 Direct and indirect loan influences

A central issue in the field of language contact is how to classify the linguistic influence that one language exerts on another language. As far as the impact of English is concerned, 'Anglicism' is often used as a generic term to describe the occurrence of English language elements in other languages.

As early as 1936, Werner Betz proposed a conceptual structure for the field of lexical borrowing based on an analysis of Latin influences on German. The terminology he coined for the categorization of loanwords proved to have significant impact on the field of language contact. Some influential researchers like Carstensen (cf. 1965: 214 f.) and Weinreich (cf. 1979: 51), who focus on the phenomenon of Anglicisms in German, still apply the categories as put forward by Betz. Betz draws a general distinction between direct and indirect loan influences.

Indirect loan influences are difficult to prove and generally not discernible on the level of word-form. For example the German 'Wolkenkratzer' is said to have been coined following the English model of 'skyscraper' (cf. Carstensen/ Busse 2001). Since the corpus analysis will be based on the retrieval of identical word-forms, no further details on the

complex system and terminology of indirect loan influences are given at this point. In order to illustrate direct loan influence, Betz proposes the following diagram (cf. 1959: 128).

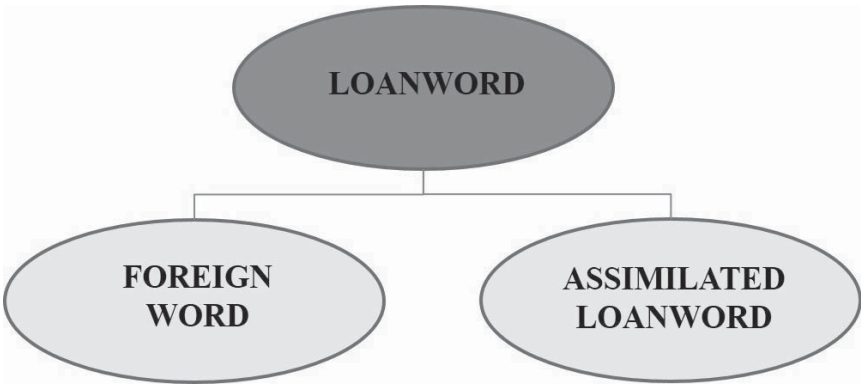

Figure 16 Classification of direct loan influences

The version above is slightly adapted since it uses the English terminology which was introduced by Duckworth (cf. 1977: 40).

Figure 16 illustrates that Betz divides loanwords into 'foreign words' and 'assimilated loanwords'. Thus, Anglicisms can fit one of the latter two categories. According to this traditional view, the basic difference between foreign words and assimilated loanwords is that the former are considered foreign objects in the language whereas the latter have adapted to the receiving language to such an extent that they are no longer considered foreign or cannot be recognized by laymen to have come from another language. The distinction between foreign words, assimilated loanwords and native words is made on the basis of formal criteria like spelling, morphology, word formation etc., and the perspective is synchronic.

Following the diagram in Figure 16 above, foreign words represent the most direct form of transfer. The spelling of the English word is kept in its original form, e.g. *crash*, *meeting* or *management*[21], and it may be pronounced either as in British or American English or it may be assimilated to the German phonological system (cf. Hedderich 2003: 48). At the beginning of the twentieth century there was a tendency to assimilate loanwords; examples include *Konzern* (← concern) and *Streik* (← strike). Thus, older loanwords are often not identified as such by the average language user, since they are fully integrated regarding spelling and pronunciation (cf. Busse/Görlach 2002: 13 f.). Once a word-form is borrowed, accommodation takes place. Görlach (cf. 2002: 10) suggests that in many cases Anglicisms develop distinctive semantic features through contextual restrictions, metaphorical and metonymic applications, euphemistic or facetious uses and other developments also visible

[21] In some cases they are capitalized, cf. subsection 5.1.5.

in native words. This also means that an Anglicism may eventually develop distinctive collocations in the target language.

The distinction between foreign words and assimilated loanwords cannot be taken into consideration in the present study, because the chosen method of corpus analysis does not provide evidence to make statements about the degree of integration of Anglicisms. Besides, the current study is only concerned with word-forms, specifically Anglicisms, which are spelled identically in the source and target language. Because the study deals with written word-forms, their assumed pronunciation is not relevant. In Betz' terminology, the Anglicisms investigated in this study are best classified as 'foreign words'.

5.1.3 Determining origin

Unfortunately, language internal criteria do not act as reliable indicators to distinguish between Anglicisms derived from British or from American English (cf. Carstensen 1975: 12). This difficulty is partly due to the fact that many Anglo-American word-forms were initially inserted to British English and only later transferred to German (cf. Busse 2008: 41). Most researchers, including Carstensen, Pfitzner and Zindler use 'Anglicism' as a generic term denoting lexical adoptions from British, American and other varieties of English. Görlach (cf. 2002: 3) mentions another reason why the distinction between American English and British English is impossible to make with regard to loanwords. Loanwords tend to be adapted in the borrowing process which deletes the distinctive features in those words which have different forms (in spelling and pronunciation) in the two varieties of English. With regard to spelling, this may refer to <s> versus <z> as for example in British English *industrialisation* versus American English *industrialization*. The pronunciation of English loans in German is a complex problem to unravel, but this does not concern the present study, because it deals with Anglicisms exclusively in the written medium. In summary, a lot of Americanisms have either been transmitted through British English or they are automatically adapted to the model dominating English language in Europe – British English. This variety is still dominant because language instruction in Europe (which is a major source of language contact) still favors British English over American English. In some cases, the adoption of words together with new inventions, concepts, or commodities imported from the US can be a clue but is not conclusive (cf. Drews 1999: 46, Görlach 2002: 3). The influence of renowned American media and companies should also not be disregarded.

Overall, linguists agree that the formerly prevailing influence of British English has been constantly declining ever since 1945 (cf. Engels 1976: 1, Busse 2008: 40) and that today American English loans exceed those of British English origin by far (cf. Carstensen 1975). This is usually explained by the economic, technological and cultural dominance of the United States (cf. Muhr 2004: 9). In summary, American influence is increasing worldwide, but British transmission is still effective, and there is not enough evidence to provide more than intuitions confirming generally held opinions (cf. Görlach 2002: 3).

5.1.4 Motives for borrowing

It has been stated several times before that English loanwords are particularly popular in German and that their influence is significant. This section is intended to shed some light on the motives for borrowing so frequently from English.

When Fink (1970) examined the motivation behind adoptions, he came to the conclusion that there were only two major reasons for the usage of Anglicisms.

1. Lack of German expressions in certain semantic fields.
2. The use of English words as an expression of closeness to and admiration for the US.

Although borrowing enriches language in that it helps to arrange lexical and semantic fields more accurately, in many cases an adoption is not mandatory (cf. Buck 1974: 131). According to Carstensen (cf. 1965: 266), Anglicisms therefore may be differentiated into ‘Bedürfnislehnwörter’ and ‘Luxuslehnwörter’. Others argue that every Anglicism enters the German language out of a necessity, even if this does not open up at first glance (cf. Donalies 1992: 103). With reference to Fink’s second motive for borrowing above, it can be stated that borrowing is not an exclusively linguistic phenomenon. Loanwords are further evidence of cultural contact and assimilation (cf. Hertzler 1965: 105). Historical and political tendencies of the past and present contribute to it. In the range of German and American English, the European Recovery Program may be seen as a starting point for eminent lexical transfer. Later on, Germany’s mostly westward-oriented policy of alliances and the political, economic and military dominance of the US played a major role. Over time, English was assigned the status of the language of negotiation in those fields. This exceptional position is strengthened via conventions of communication within large international organizations like UNO, NATO or the European Commission. The governance of the English language today is commonly connected to the globalization process; Haarmann even labels it to be ‘the engine of globalization’ (cf. 2002: 153).

Another reason for the frequent use of Anglicisms is the fact that English is taught as the first foreign language in many European countries, including Germany. The advent of the Internet in the early 1990s is regarded as another driving force for the adoption of Anglicisms (cf. Hedderich 2003: 47). Especially on the European continent, English expressions are by now credited with a certain social prestige, which also fosters their frequent use (cf. Wetzler 2006: 28, Bohmann 1996: 199).

It should have become clear that when discussing the motives for borrowing words from English, it is necessary to distinguish between linguistic causes, which lead to the formation of so-called 'Bedürfnislehnwörter', and extra-linguistic causes which lead to the coining of 'Luxuslehnwörter' (cf. discussion previous page). One rather linguistic aspect is that a borrowed English expression is often shorter than the corresponding German one. Many English word-forms are monosyllabic and consequently facilitate easier memorization and handling processes. Due to the present study's focus on corpora originating in journalistic writing, it should be pointed out that it is journalists who particularly strive for clear and brief enunciation thus achieving variation of style (cf. Carstensen 1975: 30).

5.1.5 Definition of 'Anglicism' in this study

Modern linguistics is usually said to attempt to describe the actual use of language while traditional grammar is condemned as prescriptive. Although the topic and use of Anglicisms in German is a highly controversial issue, the present study does not intend to judge whether borrowing from English may be regarded as enrichment or lead to a disfigurement of German. The present study is descriptive and aims to investigate to what extent Anglicisms display similar collocational behavior within source language and target language. In the research literature, the term 'Anglicism' is used for the most part to refer to a loan that various languages adopt from English. For the current study it is sufficient to prove a high frequency of an English word-form in German to speak of an 'Anglicism'.

The phenomenon of hybrid Anglicisms such as *Börsencrash, Rating-Agenturen or Cash-Polster* (*WirtschaftsWoche 2008*: NOV) for example, is also addressed in this study, particularly with regard to hybrid compounds and hybrid collocations. The concept of 'pseudo-Anglicisms' is also of some importance to the corpus analysis. A pseudo-Anglicism is an English word-form used in German and referring to a word existing in English with a different meaning (cf. Hedderich 2003: 48). The most prominent example is probably *H/handy*. In English something is *handy* if it is convenient to handle or use (www.oxforddictionaries.com). In German the noun *Handy* is used to refer to a cell phone. If a

key word displays perfectly divergent collocational behavior within the source and target language corpora, additional qualitative analysis will help to determine if the loanword might prove to be a 'pseudo-Anglicism'.

Since German uses the same alphabetical system as English and there are no modified letters (such as <ä>, <ö>, <ü>, <ß>) in the English alphabet, the adoption of the English spelling is possible in principle for all Anglicisms. Recent loans tend to retain the English spelling. Some Anglicisms in the target language corpora may have been capitalized to indicate nouns (cf. ibid.: 7) or italicized to indicate their foreignness. This is regarded as a measure of integration and can indicate the age of the loanword (cf. Görlach 2002: 22). It may also depend on the given magazine's editing conventions. In theory, formal aspects like capitalization might be useful to tell apart 'foreign words' and 'assimilated loanwords'. However, it has been pointed out above in subsection 5.1.2 that this distinction is not relevant for the corpus analysis. Apart from this, lexical items denoted as Anglicisms in this study will not have undergone any orthographic changes.

The present study relies largely on the dictionary by Carstensen/Busse to verify that the chosen key words are genuine Anglicisms. Carstensen's research was united in the three volume dictionary of Anglicisms that was first published in 1993 and was continued by Busse after Carstensen's death. The dictionary lists Anglicisms of different thematic areas and discusses the spelling, pronunciation, types of loan processes, and grammatical aspects. It serves well not only as a comprehensive information source for all questions regarding certain Anglicisms but also as a directory of words of English origin in German that can be used to search various corpora. Unfortunately, its latest edition dates back to 2001, and for some of the selected nodes an entry does not yet exist. These nodes were then defined as English lexical items by accordant entries in *The Oxford English Dictionary* at its website www.oxforddictionaries.com.

5.2 Lists of key words

It is impossible to analyze all Anglicisms detected in the target language corpora in the present study. Therefore, the number of Anglicisms for analysis was narrowed down by choosing smaller sets of Anglicisms, sets of 'key words' (nodes). All key words are highly frequent Anglicisms in the business and news magazine target language corpora. The corpus analysis is limited to those key words for the benefit of monitoring many more instances of each node and to enable additional qualitative analysis. The key words were

identified manually via close reading and their overall frequency within the target language corpora was calculated with the help of *WordSmith Tools*.

More precisely, potential key words (highly frequent Anglicisms) were encountered during the collection of data for the target language corpora. The overall frequency of these potential key words within the target language corpora was then calculated by using *WordSmith Tools.* In addition, these calculations allowed for a first estimate of the number of lexical collocates to be returned for such a potential key word. This is crucial for the choice of key words, because a given key word needs to feature ideally at least 5 lexical collocates to carry out according analyses. For the business magazine corpora, no minimum overall frequency of key words was set. This was different for the news magazine corpora. Here, it was determined that key words need to occur at least 100 times in the corpora in order to feature sufficient analyzable lexical collocates (cf. also subsection 6.3.1). Furthermore, it was aimed at selecting the key words evenly from all areas of reporting within the magazines. In so doing it can be assured that nodes belonging to certain sections (e.g. 'science and technology' or 'sports') of the magazines are not overly represented.

The study focuses on nouns as key words. This is because in domain-specific corpora the nouns primarily bear the weight of topic specificity, i.e. the technical messages. Nominalization is used especially frequently in popular scientific articles (cf. subsection 4.3.2). The vast majority of technical terms are nouns (cf. Martin 1993b: 213). The verbs tend to have a more general meaning (cf. Peters et al. 2000: 80). Compounds are not admitted as key words.

Given a particular key word, the aim is to identify its co-text in the source language and target language corpora. This is done by displaying the key word in the so-called 'KWIC-format' (cf. section 5.3 below for a more detailed discussion). Once *WordSmith Tools* displays the key word in context, the next step is to isolate its collocates from this co-text in the source language corpus, hypothesizing that the key word will be surrounded by identical or equivalent collocates in the target language corpora. For each occurrence of the key word, the software determines its collocates, containing the key word ±4 word-forms appearing to the right and left of the key word. Strong punctuation marks (full stops and semicolons) do not generally act as break points (cf. section 2.4) in the determination of collocates. It should be mentioned though that the present study considers the syntagmatic cohesion, i.e. the strength of the relationship (cf. subsection 2.3.2) between node (key word) and collocate. Such cohesion is present stronger in the phrase than the cohesion found in a larger section of text (cf. Peters et al. 2000: 81). But this does not mean that

collocations cannot stretch across sentence boundaries and has to be judged specifically in each case.

The analysis is dedicated to lexical collocates, because lexical items are more likely borrowed than grammatical ones (cf. Onysko 2007: 45). Since "[…] it is commonplace to distinguish roughly between grammatical and lexical items" (Sinclair 1991: 35) nouns, verbs, adjectives and adverbs are defined for this study as belonging to the class of lexical collocates.

The corpus analysis also employs so called 'stop lists'. Stop lists are manually produced collections of words used in automatic indexing to filter out words that would make poor index terms. Traditionally, stop lists are supposed to include the most frequently occurring words. In this study, stop lists are composed for each corpus to filter out the most frequently occurring grammatical words. The stop lists include functional words such as articles, pronouns, prepositions. These lists can be modified so that certain frequent terms can be eliminated if necessary. In certain cases this can also apply to lexical words. The node **CASH,** for example, frequently co-occurs with the quantifiers BILLION and MILLIARDE. Such identified equivalent collocates are not conclusive to determine similarities or differences in collocational behavior. Although they are lexical words, they come from a closed class.

To arrive at a manageable amount of material that is also sufficiently large to yield meaningful results, 36 key words (Anglicisms) were selected from the business magazine target language corpus *WirtschaftsWoche 2008*. The list of key words for the business magazine corpora is presented in Figure 17 below.

LIST OF KEY WORDS FOR THE BUSINESS MAGAZINE CORPORA		
1) ANALYST	13) ECONOMY	25) MARKETING
2) ASSET	14) EQUITY	26) OFFICER
3) BANKING	15) FINANCE	27) PORTFOLIO
4) BROKER	16) GROUP	28) PROVIDER
5) BUSINESS	17) IMAGE	29) PUBLIC
6) BUYOUT	18) INSIDER	30) RATING
7) CASH	19) INSIGHT	31) RESEARCH
8) CORPORATE	20) INVESTMENT	32) SERVICE
9) CRASH	21) INVESTOR	33) SOFTWARE
10) DEAL	22) MANAGEMENT	34) SUPPLY
11) DESIGN	23) MANAGER	35) VALUE
12) DEVELOPMENT	24) MARKET	36) WORK

Figure 17 List of key words for the business magazine corpora

Analogous to the key word selection for the business magazine corpora, 36 key words were selected from the news magazine target language corpus *Der Spiegel 2008*. Some of the key words chosen for the analysis of the news magazine corpora have been subject to previous linguistic research. However, none of them has been analyzed with regard to their collocational behavior. Thus, their previous linguistic analysis is no reason why such Anglicisms should not be used as key words in the present study. The list of key words for the news magazine corpora is presented in Figure 18 below.

LIST OF KEY WORDS FOR THE NEWS MAGAZINE CORPORA		
1) AIRPORT	13) FAN	25) MODEL
2) BESTSELLER	14) FARM	26) NETWORK
3) BOOM	15) FASHION	27) NEWS
4) CAMP	16) FREAK	28) PERFORMANCE
5) CLAN	17) GANGSTER	29) REALITY
6) COMEBACK	18) INTERVIEW	30) SOCIETY
7) COMEDY	19) JOB	31) SOUND
8) COMMUNITY	20) KILLER	32) STAR
9) DOTCOM	21) KNOW-HOW	33) STORY
10) E-MAIL	22) LIFE	34) TALK
11) ENTERTAINMENT	23) LUNCH	35) TEST
12) EVENT	24) MEETING	36) TREND

Figure 18 List of key words for the news magazine corpora

5.3 Lexical analysis software

Corpus linguistics requires the development of tools that permit the automatic processing of data. As a rule, a corpus does not contain new information about language, but the software offers us a new perspective on the familiar (cf. Hunston 2002: 5). In this study, a computational analysis of the corpora is performed with the program *WordSmith Tools*[22]. Figure 19 below shows a screenshot of the *WordSmith Tools* controller menu. The software offers three main functions: *Concord*, *KeyWords* and *WordList*. Their position is highlighted with ovals in Figure 19. The current study mainly utilizes the *Concord* function.

[22] A demo version can be downloaded from Mike Scott's website www.lexically.net/wordsmith/version5/index.html.

Figure 19 Screenshot of the *WordSmith Tools* controller menu

The size of the collocational span, which *WordSmith Tools* refers to as 'collocate horizons', the minimum frequency of a collocate and other *Concord* settings can be adjusted. This is shown in Figure 20 below and highlighted with ovals.

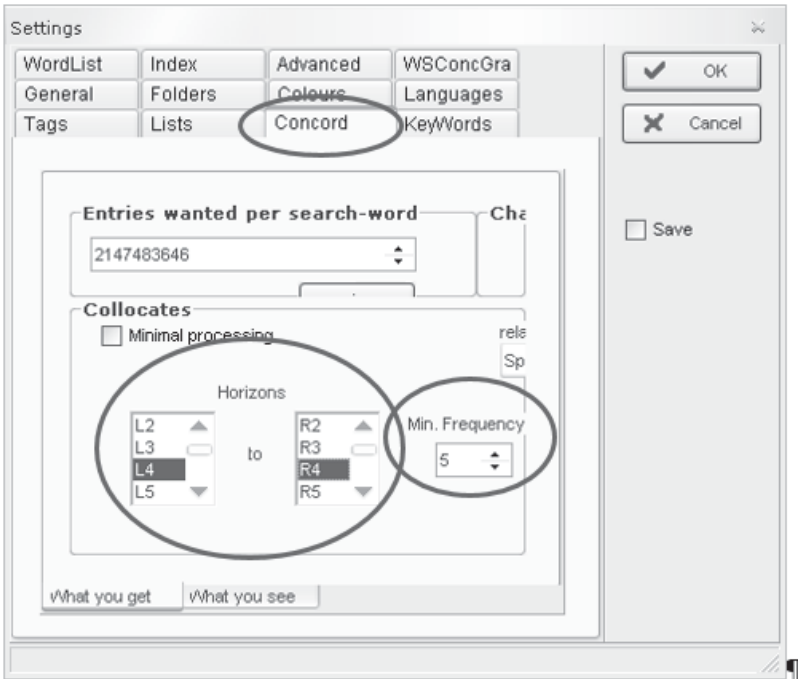

Figure 20 Screenshot of *WordSmith Tools Concord* settings

As a result of the *Concord* function, *WordSmith Tools* gives information on the number of occurrences of the key word in the corpus, on that part of the corpus in which they were found, and so-called concordance-lines, which show the key word in question in context (with the key word highlighted). The concordancer processes .txt-files and presents any word or phrase in context to determine its collocates. For many years the KWIC (Key Word in Context) format has been widely used in data processing (cf. Sinclair 1991: 32). This format makes it possible to extract all the instantiations of a given form quickly and reliably and to visualize them in context. The screenshot below shows concordance lines 1 to 15 for the key word **BANKING** in the source language business magazine corpus.

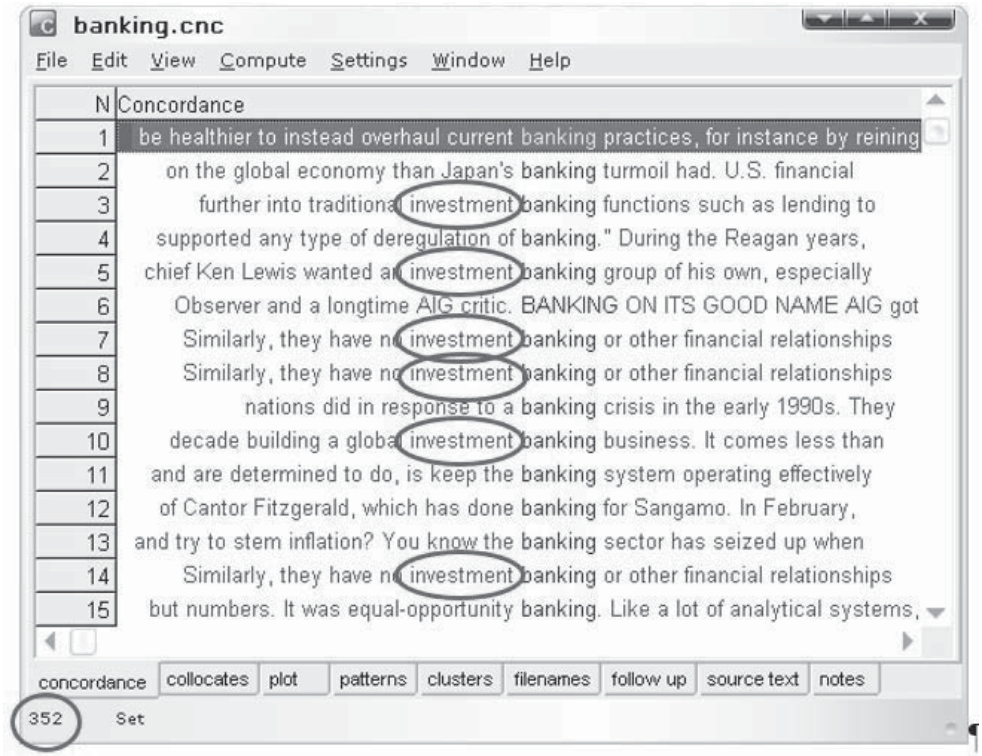

Figure 21 Concordance lines for BANKING

As indicated in the left hand bottom corner of Figure 20, **BANKING** appears 352 times in the source language corpus.

Although from Figure 21 it is already observable that INVESTMENT frequently precedes **BANKING** in L1 position, sometimes it can be difficult to see overall trends in concordance lines. Therefore they can be sorted or collocates of the node can be viewed.

WordSmith Tools lists collocates for every query according to customized settings. The statistically most significant collocation, i.e. the collocation which occurs most frequently, will head the list.

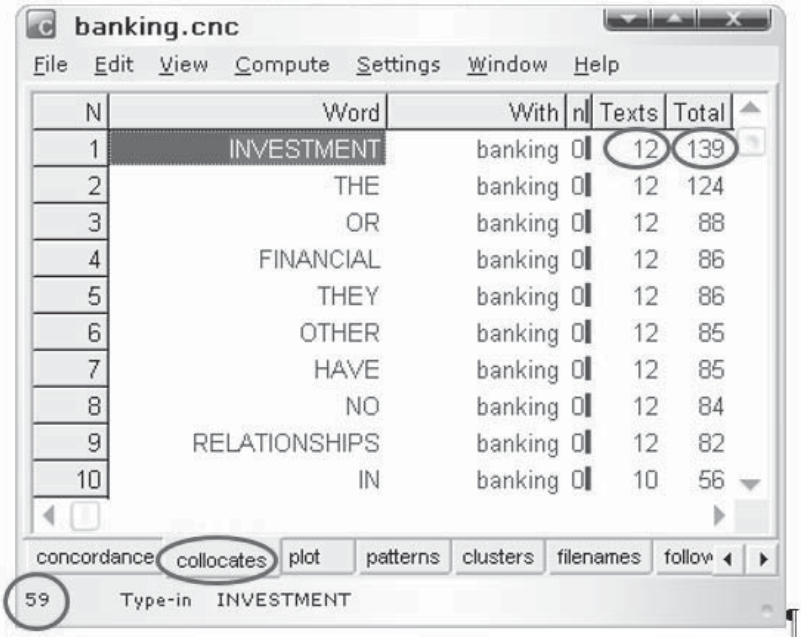

Figure 22 Collocates of BANKING

In Figure 22, the ten most frequent collocates of **BANKING** within the *BusinessWeek 2008* corpus are listed. In order to obtain exclusively grammatical or lexical collocates, stop lists are created (cf. section 5.2). When aiming at observing lexical collocates, grammatical words are put on the stop list and vice versa. In Figure 22 a stop list is not applied yet, otherwise the definite article THE for example, would have been excluded from the list of collocates. The application of a stop list reduces the number of identified collocates significantly. The number in the left hand bottom corner of Figure 22 indicates that 59 collocates of **BANKING** were identified in the *BusinessWeek 2008* Corpus. This number includes lexical and grammatical collocates. However, **BANKING** features only 27 lexical collocates (cf. subsection 6.2.2).

The numbers to the right of each identified collocate indicate in how many texts (one text per month was compiled) the collocation was found. Next to this, the total number of times the collocation was retrieved is listed.

The settings for automatic processing of the corpora were adjusted according to the concept of collocation as defined for the present study in section 2.4. Collocates have to appear within a span of ±4 left or right of a node. Their positions are referred to as L1-L4 and R1-R4. Analysis is not case-sensitive. Within these specifications, node (key word) and collocate have to co-occur at least five times within a corpus to qualify as a collocation (cf. section 2.4).

6 Corpus analysis

6.1 Introduction

This chapter analyzes the collocational behavior of the Anglicisms that were identified as key words in section 5.2. The major strength of a corpus-based methodology lies in its suitability for capturing variation and/or consistency in a large number of instantiations. This is owed to the possibilities of partially automated corpus analysis. Depending on the research questions and objectives of a study, a corpus can be annotated or unannotated. The corpora in the present study are unannotated; they consist exclusively of raw text data without offering any additional information about their lexical components.

The hypotheses (cf. section 1.1) are tested initially on the business magazine corpora; i.e. economic and financial reporting. Afterwards, the hypotheses will be assessed for the news magazine corpora; i.e. a representative sample of general news reporting covering broader fields from outside of economics. The hypotheses are tested on sets of comparable corpora in two different languages, i.e. corpora from the same domains in English and German. The distinction between the two types of discourse and their representation is made to support the main hypothesis that identical and equivalent collocates of Anglicisms are detectable in general and specialized journalistic text alike.

The corpus analysis pursues two main goals. The first one is to assess the extent to which identical and equivalent collocates are used. The second goal is to analyze the contextual factors which influence variability, i.e. when are identical or equivalent collocates used. In view of the fact that such an analysis features aspects which are hardly calculable with quantitative methods, qualitative investigation is required. Manual investigation is particularly important to distinguish collocates of a given node from constituents of proper nouns or compounds which are incorrectly identified as collocates according to the quantitative analysis. It is the contention of this research that quantitative and qualitative analyses promote and support each other to provide valid and reliable findings. Descriptive frequency information is combined with qualitative interpretation.

For the quantitative analysis, the study uses lexical analysis software which facilitates the identification of key words in context and the calculation of their collocates. It is possible to obtain an accurate account of the frequency and distribution of a given key word. The results enable direct comparisons between the different corpora (cf. McEnery/Wilson 1996: 61). The corpus analysis is qualitative because it interprets the collocational patterns, the relatedness of nodes and collocates and the equivalency relation between source language and target language collocates. This can only be achieved if all instantiations in

which a given node and collocate were identified as such by means of the quantitative corpus analysis are manually accessed. It is then determined in which positions within the collocational span the identified collocate co-occurs with a particular node. The quantitatively identified collocations are analyzed qualitatively with regard to their actual status as collocation.

Again, it should be emphasized that the corpus analysis is word-form based (cf. section 2.4). However, there is one exception. Inflectional variants visible in adjective endings in German need to be considered as translation equivalents (cf. section 3.2). For example the target language collocates FÜHREND, FÜHRENDE, FÜHRENDES are all considered (translation) equivalents of the source language collocate LEADING.

In the case of identical singular and plural forms of a given node or collocate, only those instantiations are counted which are equivalent to the previously identified source language form. For example the source language collocate COMPANY is equivalent to the target language collocate UNTERNEHMEN. The difficulty is that UNTERNEHMEN is spelled identically in its singular and plural forms, but only in its singular meaning can it be considered an equivalent collocate. The lexical analysis software cannot differentiate between the singular and plural meanings of a word-form, thus manual analysis is required in such cases.

A significant challenge is to retrieve all occurrences of a key word in the target language corpora where there might be variations in spelling due to a preference of the British or American English variety (cf. subsection 5.1.3). All potential variations are considered.

While in section 5.3 screenshots are used to exemplify procedure, the results of the corpus analysis are presented in manually produced diagrams and tables based on data generated by *WordSmith Tools*. This is preferred since it allows for restriction to data essential to the study. ‘Key words’ will from now on be exclusively referred to as ‘nodes’. Nodes and collocates are always given in capital letters and nodes are printed in bold in addition.

6.2 Business magazines

6.2.1 Quantitative analysis

When calculating the frequencies of the nodes within the business magazine source and target language corpora, a few opening observations were made. To begin with, the data show that highly frequent key words in the source language do not necessarily turn into

highly frequent Anglicisms in the target language and vice versa. For a better overview, Figure 23 below displays the frequency of the nodes, descending from highest to lowest frequency.

FREQUENCY OF NODES			
SOURCE LANGUAGE		TARGET LANGUAGE	
BUSINESS	(2,641)	MANAGER	(1,330)
MARKET	(2,441)	MANAGEMENT	(523)
ECONOMY	(1,250)	SOFTWARE	(382)
INVESTMENT	(1,240)	BUSINESS	(343)
WORK	(1,118)	SERVICE	(268)
GROUP	(996)	DESIGN	(210)
RESEARCH	(848)	MARKETING	(207)
MANAGEMENT	(840)	IMAGE	(201)
CASH	(823)	INVESTOR	(199)
DEAL	(780)	INSIDER	(195)
SERVICE	(710)	EQUITY	(192)
CORPORATE	(667)	INVESTMENT	(177)
VALUE	(662)	CASH	(172)
SOFTWARE	(625)	ANALYST	(156)
PUBLIC	(544)	GROUP	(156)
EQUITY	(539)	DEAL	(150)
MANAGER	(488)	RATING	(115)
ANALYST	(427)	CRASH	(111)
FINANCE	(414)	PORTFOLIO	(101)
MARKETING	(413)	RESEARCH	(96)
DESIGN	(385)	VALUE	(71)
BANKING	(352)	ECONOMY	(68)
PORTFOLIO	(323)	WORK	(62)
DEVELOPMENT	(318)	CORPORATE	(56)
OFFICER	(262)	MARKET	(55)
INVESTOR	(247)	ASSET	(50)
ASSET	(235)	BUYOUT	(49)
SUPPLY	(222)	BANKING	(44)
BUYOUT	(143)	FINANCE	(43)
RATING	(132)	SUPPLY	(42)
IMAGE	(105)	OFFICER	(41)
BROKER	(101)	BROKER	(35)
PROVIDER	(73)	INSIGHT	(35)
CRASH	(55)	DEVELOPMENT	(29)
INSIGHT	(39)	PROVIDER	(28)
INSIDER	(30)	PUBLIC	(26)

Figure 23 Frequency of nodes in the business magazine corpora

Figure 24 below shows the nodes and their frequencies in alphabetical order. Frequencies range from 30 to 2,641 within the source language and from 26 to 1,330 in the target language corpus. As a reminder, the total word count of the *BusinessWeek* corpus is 1,413,686 million words, whereas the *WirtschaftsWoche* corpus comprises a total of 2,754,827 million words (cf. section 4.5). Because of the difference in corpus size, the frequencies of all

nodes are also given per 1 million words. This again facilitates comparison of frequencies of the identical nodes in differently sized corpora.

FREQUENCY IN ALPHABETICAL ORDER				
	SOURCE LANGUAGE	FREQUENCY/ 1 MIO WORDS	TARGET LANGUAGE	FREQUENCY/ 1 MIO WORDS
ANALYST	(427)	302	(156)	57
ASSET	(235)	166	(50)	18
BANKING	(352)	250	(44)	16
BROKER	(101)	71	(35)	12
BUSINESS	(2,641)	1,868	(343)	125
BUYOUT	(143)	101	(49)	18
CASH	(823)	582	(172)	62
CORPORATE	(667)	472	(56)	20
CRASH	(55)	39	(111)	40
DEAL	(780)	552	(150)	54
DESIGN	(385)	272	(210)	76
DEVELOPMENT	(318)	225	(29)	11
ECONOMY	(1,250)	884	(68)	25
EQUITY	(539)	381	(192)	70
FINANCE	(414)	293	(43)	16
GROUP	(996)	705	(156)	57
IMAGE	(105)	74	(201)	73
INSIDER	(30)	21	(195)	71
INSIGHT	(39)	28	(35)	13
INVESTMENT	(1,240)	877	(177)	64
INVESTOR	(247)	175	(199)	72
MANAGEMENT	(840)	594	(523)	190
MANAGER	(488)	345	(1,330)	483
MARKET	(2,441)	1,727	(55)	20
MARKETING	(413)	292	(207)	75
OFFICER	(262)	185	(41)	15
PORTFOLIO	(323)	228	(101)	37
PROVIDER	(73)	52	(28)	10
PUBLIC	(544)	385	(26)	9
RATING	(132)	93	(115)	42
RESEARCH	(848)	600	(96)	35
SERVICE	(710)	502	(268)	97
SOFTWARE	(625)	442	(382)	139
SUPPLY	(222)	157	(42)	15
VALUE	(662)	468	(71)	26
WORK	(1,118)	791	(62)	23

Figure 24 Alphabetical order of nodes in the business magazine corpora

Although the target language corpus is about double the size of the source language corpus, it encompasses just enough occurrences to reliably determine collocates of all selected nodes. As has been pointed out before, a comparatively large target language corpus holds thus advantages for the corpus analysis. Furthermore, it can be observed from Figures 23 and 24 that the relation between frequency in the source and target language corpus is more or less arbitrary. **MANAGER** for example is the most frequent node in the target language corpus, but only ranks 17th in the source language corpus; whereas **ECONOMY** ranks third in the source language and 22nd among the target language nodes. Figures 23 and 24 reveal Anglicisms which are, judging strictly from absolute numbers (cf. Figure 24), more frequent in the target language than the same expression is in the source language corpus; they are **CRASH**, **IMAGE**, **INSIDER** and **MANAGER**. This leads to speculation whether frequency is an indicator for degree of integration and whether this has an effect on their collocational behavior within the target language.

Having looked up the above-mentioned nodes in Carstensen's/ Busse's *Anglizismen-wörterbuch* (2001), it turns out that those nodes which are more frequent in the target language corpus than in the source language corpus, are among the comparatively early introduced Anglicisms on the key word list. They were first recorded between 1928 and 1966. Another cause for their high frequency may be that the German translations offered are rather long and bulky. As has been mentioned earlier, journalists strive to keep their language brief and to the point (cf. subsection 5.1.4). Using 'Zusammenbruch' instead of **CRASH**, 'öffentliches Erscheinungsbild' instead of **IMAGE**, or 'Eingeweihter' instead of **INSIDER** would not serve this purpose well. Due to the fuzzy underlying semantic concept, matters stand more complicated with **MANAGER**. The word has many German translations, characteristically describing a 'Führungskraft' in one or another area. **MANAGER** is a very convenient circumscription since it does not call for further specification, but encompasses many German translations like 'Abteilungsleiter', 'Betriebsleiter' or 'Geschäftsführer'. However, it is impossible to pin down the exact causes of the excessive use of **CRASH**, **IMAGE**, **INSIDER** or **MANAGER** in German, most likely both factors add to it.

According to the definition of collocation in this study (cf. section 2.4), a total of 2,580 lexical collocates of the 36 nodes was detected in the business magazine source language corpus. The number of lexical collocates per node differs heavily. Within the source language corpus, the number of lexical collocates per node ranges from 3 to 314.

A total of 515 lexical collocates of the 36 nodes was identified in the business magazine target language corpus. Within the target language corpus, the number of lexical collocates

per node ranges from 2 to 135. These numbers show that the chosen 36 nodes feature about five times as many collocates in the source language business magazine corpus when compared to the target language business magazine corpus. As has been mentioned before, the target language corpus is roughly double the size of the source language corpus. It is not very surprising that Anglicisms feature more lexical collocates in their source language than in their target language. These numbers are also only somewhat conclusive, as they do not take into account the different frequencies of the nodes in the two corpora. However, they allow for an overview of the amount of researched data.

With such large amounts of data, limitation is necessary. Consequently, the corpus analysis is dedicated to the 'Top 5' collocates, i.e. the five collocates which co-occur most frequently with a node in the source and target language corpus. Frequency is the basis on which collocations are analyzed in this study. Thus the most frequent collocations, i.e. the Top 5, are viewed to deliver the most conclusive data. However, the distribution of nodes is erratic, therefore not every node features five lexical collocates, a few only possess two or three lexical collocates.

Looking at the Top 5 collocates of each node, the data display that 34 (94.4%) of the 36 nodes share lexical collocates. The only exceptions are the nodes **BROKER** and **DEAL** which do not share any identical or equivalent collocates. A total of 17 nodes share identical collocates, 11 nodes feature identical and equivalent collocates and 6 nodes only equivalent collocates. The group of the Top 5 collocates of all nodes includes 48 identical and 26 equivalent collocates. Figure 25 below summarizes this information again.

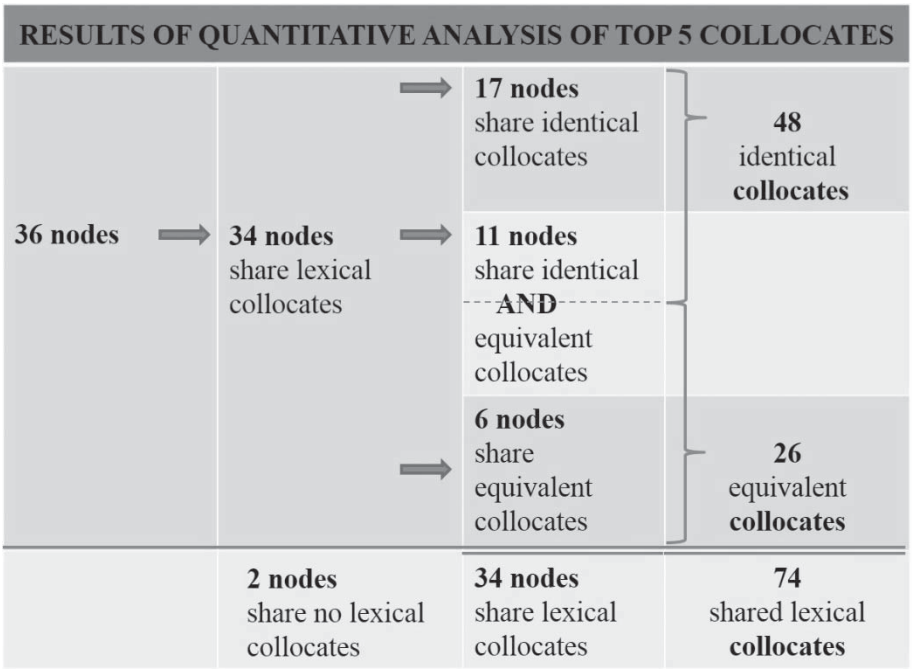

Figure 25 Results of the quantitative analysis of Top 5 business magazine collocates

The nodes feature different numbers of shared lexical collocates within the Top 5. ‘Shared lexical collocates’ captures the total of equivalent and identical collocates (cf. section 3.2). The distribution of the shared lexical collocates among the nodes, and more particularly the distribution of the 48 identical and 26 equivalent lexical collocates, is illustrated in Figure 26 below.

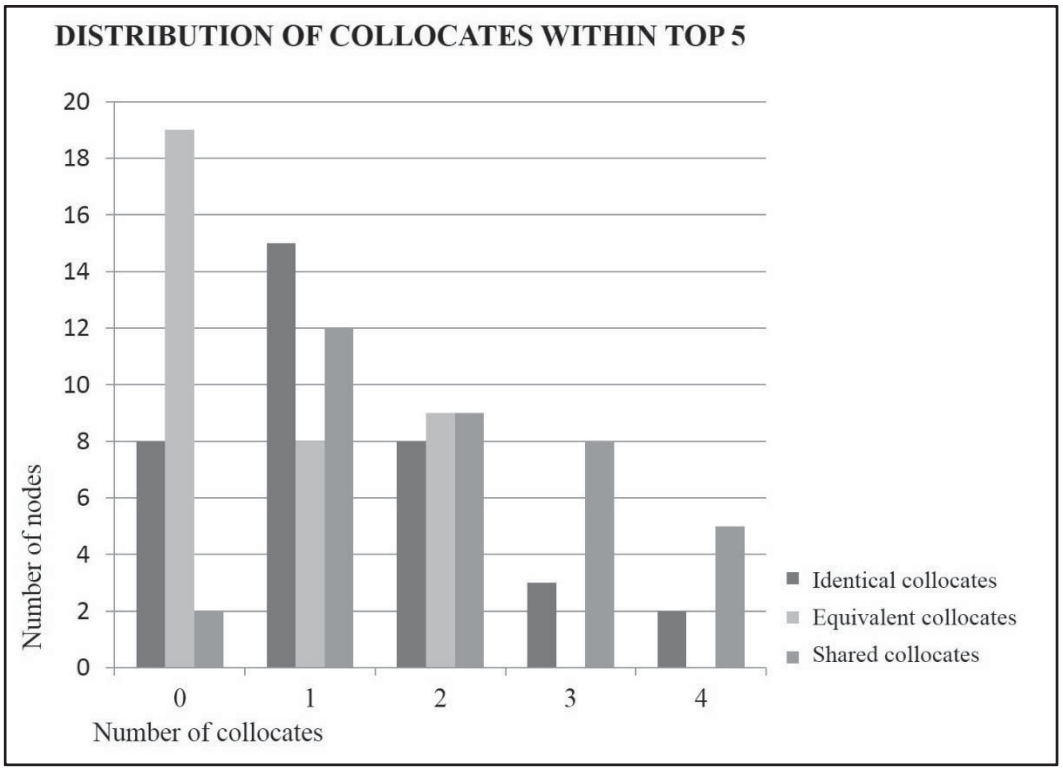

Figure 26 Distribution of business magazine collocates within Top 5

The x-axis in Figure 26 shows the number of collocates. As a result of the quantitative corpus analysis it can be stated that each of the 36 analyzed nodes features a number of 0 to 4 shared lexical collocates among the Top 5 collocates. Thus, the x-axis is labeled from 0 to 4. The nodes feature different numbers of identical and equivalent lexical collocates. This is shown with the different bars in different shades of blue. In addition, the bar in the lightest shade of blue provides the total number of shared collocates, i.e. all identical and equivalent collocates of a node are counted and summarized as ‘shared collocates’ (for a detailed definition of ‘identical collocate’ and ‘equivalent collocate’ cf. section 3.2). The y-axis shows the number of nodes from 0 to 20.

Moving from left to right on the x-axis, the following information is represented in the diagram: 8 nodes do not feature any identical collocates (but they may feature equivalent collocates), 19 nodes do not feature any equivalent collocates (but they may feature identical collocates), 2 nodes do not feature any shared (identical or equivalent) collocates at all.

15 nodes feature 1 identical collocate (and may feature additional equivalent collocates), 8 nodes feature 1 equivalent collocate (and may feature additional identical collocates). 12 nodes share a total of exactly 1 (either identical or equivalent) collocate.

8 nodes feature 2 identical collocates (and may feature additional equivalent collocates), 9 nodes feature 2 equivalent collocates (and may feature additional identical collocates). 9 nodes feature a total of exactly 2 shared collocates (either identical or equivalent or a combination of the two).

3 nodes feature 3 identical collocates (and may feature additional equivalent collocates), 0 nodes feature 3 equivalent collocates and may feature additional identical collocates), 8 nodes feature a total of exactly 3 shared collocates (as a result of the combination of identical and equivalent collocates of a given node). For example, the node **EQUITY** features 1 identical and 2 equivalent lexical collocates; in other words **EQUITY** features a total of 3 shared collocates.

2 nodes feature 4 identical collocates, 0 nodes feature 4 equivalent collocates (but they may feature identical collocates), 5 nodes feature a total of exactly 4 shared collocates (as a result of the combination of identical and equivalent collocates of a given node).

As an intermediate result and judging from this quantitative analysis of the business magazine corpora, it seems legitimate to assume that similarities considerably outnumber the differences of collocational behavior of Anglicisms in American and German business magazines. The qualitative analysis is expected to deliver more details on this outcome.

6.2.2 Qualitative analysis

The 34 nodes which evidence, according to the quantitative analysis, similar collocational behavior are subject to a supplementary qualitative analysis. Further investigation of the 2 nodes which do not share any lexical collocates is unlikely to lead to rewarding results.

According to the underlying methodology, source language collocates of the selected nodes are identified and translated. Afterwards, target language collocates of the same node are retrieved and contrasted with results from the source language corpus. For this purpose a chart, containing significant information for comparison, is constructed for each of the 36 nodes. The charts of the 2 nodes which do not share any lexical collocates are included in this subsection of the study, but they are not discussed in detail for the qualitative analysis. The nodes are analyzed in alphabetical order. In the analysis charts, ‘T’ stands for the total of instantiations of a given node, ‘LC’ indicates the total number of lexical collocates of a

given node. The numbers in parentheses '()' show how many times node and collocate co-occur within the designated (source or target language) corpus. Because the search for collocates is not case-sensitive, all nodes and collocates are given in capital letters. This is analogous to the modus operandi of *WordSmith Tools* (cf. Figure 22). Pursuing as clear an arrangement as possible, identical collocates are marked with a mathematical 'identical to' sign '≡' and equivalent collocates are marked with the mathematical 'corresponds to' sign '≙'.

One of the main tasks of the qualitative corpus analysis is to determine reliably which quantitatively identified shared collocates are truly of that kind. In the charts a checkmark '✓' is put by those collocates which prove to be identical or equivalent according to the quantitative and qualitative corpus analysis. The outcomes of the quantitative and qualitative corpus analysis differ significantly and a synopsis of the results is provided in subsection 6.2.3.

ANALYST

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
427	48	RESEARCH	(45)	Forschung	BANK	(7)	15	156
		SENIOR	(27)	Vorgesetzte/r	ERWARTET	(6)		
		GLOBAL	≡ (17)	weltweit	INSIGHT	(5)		
		FIRM	(15)	Firma/Kanzlei/ Unternehmen	GLOBAL	≡ (5)		
		CREDIT	≡ (14)	Anerkennung	CREDIT	≡ (5)		

The node **ANALYST** shares two identical lexical collocates. In the source and target language corpora the collocate GLOBAL is used in a number of different positions within the collocational span. GLOBAL is never used in R1 or L1 position. Thus it is likely that all instantiations are collocations, not compounds. Examples from both corpora are provided below.

SL[23] 67[24]:"[…] use the iPhone as a PC in their pocket," says analyst Trip Chowdhry of Global Equities Research.

SL 264: "Americans see hybrids as the darling," says Global Insight auto analyst Philip Gott, "and diesel as old-tech."

SL 291: […] says Trip Chowdhry, an analyst with Global Equities Research.

SL 326: […] says Christoph Stürmer, an analyst at researcher Global Insight in Frankfurt.

TL[25] 47: US-Analyst Nigel Gault von Global Insight hält […].

TL 104: […] urteilt Global-Insight-Analyst Mathyssek.

It becomes evident that in the source and target language examples GLOBAL forms part of the company names *Global Equities Research* and *Global Insight*. Unfortunately, this extends to all instantiations in which GLOBAL has been identified as a lexical collocate by the software. Thus, the qualitative analysis refutes the result of the quantitative analysis in this case and GLOBAL cannot be counted as an identical lexical collocate of **ANALYST**.

With the second identical collocate, CREDIT, the case is somewhat different. CREDIT is also used in a number of different positions within the collocational span of **ANALYST**. However, in the source language corpus it is frequently used in L1 position. To be precise, in 7 out of 14 co-occurrences CREDIT appears in L1. This leads to the assumption that CREDIT **ANALYST** might be a 'compound', i.e. a "[…] lexical unit functioning both grammatically and semantically as a single word" (Quirk 1985: 1567). An example is presented below.

SL 182: […] where he took a job as a credit analyst at North Carolina National Bank.

In German and English compounding is very common (cf. Yule 2003: 65). Naturally, when computing collocates within a predefined span, elements forming a compound with the node will turn up as results in L1 and R1. Compounds encompass many characteristics and are object to comprehensive research. A detailed definition of 'compound' would exceed

[23] In the provided examples 'SL' stands for 'source language'.
[24] The numbers indicate the number of the concordance line which the sample is taken from. In other words it also means that this example is the 67th instantiation of **ANALYST** in the source language corpus.
[25] In the provided examples 'TL' stands for 'target language'.

the scope of the study. For the purposes of the qualitative corpus analysis compound status will be determined based on three criteria. Firstly, a collocate which is potentially the constituent of a compound has to co-occur with the node in L1 or R1 position within the collocational span. Secondly, it has to co-occur with the node in L1 or R1 position at least one third of all instantiations in which it has been quantitatively qualified as a collocate. Lastly, a dictionary entry[26] has to exist for the potential compound. If all these three criteria are fulfilled, the collocate will not be counted as collocate any longer, but as a constituent of a compound. The present study aims at the evaluation of collocations and therefore constituents of compounds need to be excluded as collocates.

As for CREDIT, all three criteria are fulfilled. *Oxford Dictionaries* lists CREDIT **ANALYST** and defines it as "a person employed to assess the credit rating of people or companies." It should be noted that the compound CREDIT **ANALYST** is not used in the target language corpus.

Yet, CREDIT does not exclusively occur in L1 position. Only half of all instantiations in the source language corpus are compounds. The remaining 7 co-occurrences of **ANALYST** and CREDIT are potential collocations. According to the criteria (cf. sections 2.4 and 5.3) for the present analysis of collocations, node and collocate have to co-occur at least five times within a corpus. Thus, if the corpus evidence verifies that the remaining 7 co-occurrences of **ANALYST** and CREDIT are actual collocations, CREDIT could still be counted as an identical lexical collocate. However, in all cases in which CREDIT is used in other positions but L1 and R1 it is part of the company names *Uni Credit*, *Credit Sights* or *Credit Swiss*. This applies equally to the source and the target language corpora. One example per corpus is provided below.

SL 112: […] says David Hendler, a senior analyst at research firm Credit Sights.

TL 27: Doch Branchenkenner wie Koya Tabata, Analyst bei Credit Suisse […].

[26] Three mono- and bilingual dictionaries were chosen to determine compound status based on their lexical entries. As a monolingual British English dictionary the *Oxford English Dictionary* (OED) was chosen. Instead of working with the printed version the online version, which is constantly updated, is used. It is available at www.oxforddictionaries.com. This dictionary is either referenced by citing the aforementioned website or by referring to '*Oxford Dictionaries*'. As a monolingual American English dictionary *Merriam Webster* was chosen. Again, the present study refers to the online version rather than the printed one. The dictionary is available at www.merriam-webster.com and cited as such or '*Merriam Webster'*. In addition, one bilingual dictionary also backs the determining of compound status: *PONS*. The corresponding websites is www.pons.eu. The dictionary is cited by referring to this website or '*PONS*'.

To summarize the qualitative analysis of **ANALYST**, it can be stated that neither of the quantitatively identified shared lexical collocates proved to form a collocation with the node. The collocates GLOBAL and CREDIT are either part of company names which are also used in the target language or constituents of a compound. *WordSmith Tools* cannot differentiate between proper nouns/company names and collocates or constituents of a compound and collocates. Already the analysis of the first node demonstrates that the qualitative analysis of all collocations is mandatory to arrive at meaningful results for this study.

One additional comment is required when looking at the chart for **ANALYST** and the identified collocates. One of the Top 5 source language collocates of **ANALYST** is FIRM. It should be pointed out that the word-form FIRM can either be an adjective or a noun. Because the corpora are unannotated, *WordSmith Tools* cannot differentiate between the two word classes. The manual disambiguation of word class reveals that in all 15 instantiations of FIRM it is used as a noun. Had FIRM been used as both a noun and adjective, it would have been counted as two different lexical collocates.

ASSET

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
235	28	MANAGEMENT	≡ (52)	Geschäftsführung	MANAGEMENT	≡ (18)	7	50
		BACKED	≡ (15)	abgesichert	BACKED	≡ (16)		
		SECURITIES	≡ (12)	Wertpapiere	SECURITIES	≡ (16)		
		MANAGER	≡ (11)	Führungskraft	MANAGER	≡ (6)		
		FUND	(10)	Fonds	TROUBLED	(5)		

For the node **ASSET** four identical lexical collocates could be identified within the business magazine corpora. In the source language corpus, MANAGEMENT co-occurred 51 times in R1 position and once in L2 position with **ASSET**. In L2, MANAGEMENT is part of a department name and does thus not enter into a collocation with **ASSET**. The example at stake is provided below.

SL 151: Goldman Sachs Capital Research & Management, Iridian Asset Management, and Berkeley Capital […].

The frequent combination **ASSET** MANAGEMENT is listed in *PONS* and translates to "Vermögensverwaltung". **ASSET** MANAGEMENT has two general definitions, one relating to advisory services and the other relating to corporate finance. In the target language corpus, MANAGEMENT occurs 18 out of 18 times in R1 position within the collocational span of **ASSET**. Consequently, **ASSET** MANAGEMENT qualifies as a compound and not a collocation and this identical compound is also used in German. Examples from both corpora are given below.

SL 123: So he created a high-powered internal asset management unit that would employ […].

TL 33: Das vom Aachener Forschungsinstitut für Asset Management (Fifam) berechnete Insiderbarometer […].

The next identified collocate for **ASSET** is BACKED. In the source language corpus, BACKED co-occurs with **ASSET** exclusively in R1 position. This leads to the assumption that **ASSET** BACKED is a compound. *Oxford Dictionaries* lists **ASSET** BACKED as an adjective and defines it as "denoting securities having as collateral the return on a series of mortgages, credit agreements, or other forms of lending." The spelling varies and frequently it is spelled with a hyphen, which the corpus evidence supports. One corresponding example is provided below.

SL 36: On July 3, Tribune signed a $300 million asset-backed commercial paper deal with Barclays […].

In the target language corpus, BACKED also only occurs in R1 position. When looking at the target language corpus data another observation can be made. Whenever BACKED occurs in R1, SECURITIES occurs in R2 position. In the source language corpus SECURITIES is preceded by BACKED in all 12 occurrences of SECURITIES. **ASSET** BACKED SECURITIES is a compound in the source language and adopted in its identical form in German. Examples are provided below.

SL 167: […] while the loans to holders of asset-backed securities are temporary by design.

TL 6: Die Manager hatten versucht, die Rendite durch Investitionen in Asset Backed Securities aufzupeppen.

According to *PONS*, **ASSET** BACKED SECURITIES translates to German "forderungsbesicherte Wertpapiere". BACKED and SECURITIES do not qualify as identical lexical

collocates of **ASSET**. Instead, both collocates are constituents of a compound which is borrowed.

The last identical lexical collocate for **ASSET,** which was identified in the quantitative analysis, is MANAGER. In the source language corpus, MANAGER is identified in L3, L4 and most frequently in R1 position. **ASSET** MANAGER is a listed compound in *PONS* and translates to "Vermögensverwalter(in)". This compound also occurs 5 times in the target language corpus. This means that all co-occurrences of **ASSET** and MANAGER in the target language are occurrences of the compound **ASSET** MANAGER. For examples from both corpora see below.

SL 149: […] a Princeton (N.J.) asset manager, says the economy is getting all the help it needs […].

TL 10: […] einer der dienstältesten Asset Manager Deutschlands, der den Dax seit dessen Geburtsstunde analysiert.

It should be noted that all 7 identified lexical collocates of **ASSET** in the target language corpus are Anglicisms. In addition to the Top 5 collocates these were RELIEF and PROGRAM. They are mentioned here, because they also co-occur 5 times with the node. However, in the source language corpus neither RELIEF, nor PROGRAM were identified as collocates of **ASSET**.

BANKING

SOURCE LANGUAGE				TARGET LANGUAGE			
T	LC	Top 5	TRANSLATION	Top 5	LC	T
352	27	≡ INVESTMENT (140)	Investition	PRIVATE (10)	3	44
		FINANCIAL (86)	finanziell	ONLINE (9)		
		SIMILARLY (82)	ähnlich	≡ INVESTMENT (5)		
		RELATIONSHIPS (82)	Beziehung/ Verbindung	–		
		INDUSTRY (64)	Branche/Gewerbe	–		

As illustrated above, all lexical collocates of **BANKING** which were identified in the target language corpus are Anglicisms.

BANKING features one identical lexical collocate, namely INVESTMENT. In both corpora INVESTMENT was exclusively found in L1 position combining to INVESTMENT

BANKING. *Oxford Dictionaries* reveals that INVESTMENT **BANKING** is a derivative of 'investment bank' and the latter refers to "a bank that purchases large holdings of newly issued shares and resells them to investors." Clearly, INVESTMENT **BANKING** is a compound and it is used in both languages in identical form. Below corpus examples for INVESTMENT **BANKING** are provided.

SL 1: Similarly, they have no investment banking or other financial relationships with them.

TL 42: […] halb so viel Ertrag wie das Investment Banking und wurde bei der Deutschen Bank lange Zeit eher als Beiwerk abgetan.

The German translation of INVESTMENT **BANKING** is "Emissionsgeschäft" (www.pons.eu). Since a German equivalent exists, the English term is most likely preferred for stylistic reasons.

While the other source language collocates occupy different positions within the collocational span, PRIVATE and ONLINE were only detected in L1 position in the target language corpus. PRIVATE **BANKING** was also used 8 times in the source language corpus, whereas ONLINE **BANKING** did not occur at all. Examples of PRIVATE **BANKING** from both corpora are provided below.

SL 46: UBS stopped offering private banking to wealthy U.S. clients in […].

TL 19: […] die über das Kreditgeschäft hinaus bis in unser Private Banking reichen.

Apparently the frequent use of INVESTMENT **BANKING** in the source language and its adoption in German lead to further imitation of this structure in the target language, regardless of whether the same combination exists in the source language. Newly coined combinations are semantically transparent. ONLINE **BANKING** is banking via the Internet. Consider the following example.

TL 21: Von maßgeschneiderten Kreditlösungen oder Online-Banking ist die Sparkasse weit entfernt.

The second element provides the general meaning and the first element makes this meaning more specific. Steffens (2003: 5) refers to such combinations as 'Anglo neologisms'. Newly coined combinations like this are also referred to as 'pseudo-Anglicisms" (cf. subsection 5.1.5).

BROKER

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
101	9	REAL	(8)	Immobilie/n	PRIME	(7)	3	35
		ESTATE	(8)	Grundstück/e	HEDGEFONDS	(5)		
		INSURANCE	(6)	Versicherung	HABEN	(5)		
		DEALERS	(5)	Händler	-			
		GROUP	(5)	Gruppe/Konzern	-			

The node **BROKER** does not feature any identical or equivalent lexical collocates. One interesting fact should be pointed out though. The top target language collocate PRIME is an Anglicism. PRIME occurs exclusively in L1 position combining to PRIME **BROKER**. Consider the examples below.

TL 17: Es gibt Hedgefonds, die den Prime Broker nicht gewechselt haben, und nun sind ihre Vermögenswerte eingefroren.

TL 24: Hedgefonds, für die Lehman als sogenannter Prime Broker mit Wertpapieren handelte und Kredite vergab […].

This expression cannot be found in the source language corpus or any of the chosen dictionaries. However, the term PRIME **BROKER** exists in finance and is defined as follows:

> "[…] a large financial institution that offers services to large institutional clients or hedge funds. It is possible for a firm to have more than one prime broker." (www.businessdictionary.com)[27]

Despite the fact that PRIME does not turn up as a source language collocate for **BROKER** in the quantitative analysis and is never used in the source language corpus, it is used frequently in the target language corpus. Remarkably, PRIME **BROKER** seems to be used correctly in the target language corpus, i.e. it is not an Anglo neologism or pseudo-Anglicism.

[27] The website www.businessdictionary.com is consulted for explanations and definitions of technical terms which do not have an entry in the chosen dictionaries. *Business Dictionary* is not decisive to determine compound status, but merely referred to for explanatory purposes.

BUSINESS

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
2,641	248	SCHOOL	≡ (208)	Fakultät/ Hochschule	SCHOOL	≡ (108)	26	343
		MODEL	(54)	Modell	DESIGN	(23)		
		EXCHANGE	(50)	Tausch/Währung	OBJECTS	(21)		
		SMALL	(47)	klein	EUROPEAN	(20)		
		GRADUATE	≡ (23)	Absolvent(in)	GRADUATE	≡ (17)		

BUSINESS is among the most frequent nodes in the source and target language corpora (cf. Figure 23). The node has two identical lexical collocates among the Top 5. SCHOOL was detected 88 times in R1 position within the target language and 125 times in the source language corpus forming the compound **BUSINESS** SCHOOL. Its German translation "Fakultät der Wirtschaftswissenschaften" (www.pons.eu) did not occur in the target language corpus and neither did 'Business-Hochschule', 'Wirtschafts-Hochschule' or 'kaufmännische Fachschule'. The expression 'Business-Schule' was found once. The corresponding example is presented below.

TL 6: […] sowie an einer Business-Schule in Barcelona startete er beim US-Konzern Proctor & Gamble […].

It can be concluded that through its exceptionally frequent use the compound **BUSINESS** SCHOOL has gained acceptance to the degree that it is completely suppressing the use of target language equivalents. Below, examples are provided to demonstrate the use of **BUSINESS** SCHOOL in both corpora.

SL 885: […] served in the army as an officer for six years before going to business school.

TL 124: […] Berufserfahrung erforderlich, um an einer Business School angenommen zu werden.

Accordingly, SCHOOL is not a collocate of **BUSINESS**, but the compound **BUSINESS** SCHOOL is borrowed as a whole in German.

The remaining 20 co-occurrences of **BUSINESS** and SCHOOL in the target language corpus can be attributed to the expression 'Graduate SCHOOL of **BUSINESS'** with very few

exceptions. SCHOOL does not qualify as an identical collocate within the guidelines of the corpus analysis.

All 17 co-occurrences of GRADUATE and **BUSINESS** in the target language corpus are part of the expression 'GRADUATE School of **BUSINESS**'. All 23 instantiations of GRADUATE in the source language corpus also occurred in L3 position, forming the identical expression. Below, examples from both corpora are provided.

SL 409: […] a professor of organizational behavior at Stanford University's Graduate School of Business.

TL 277: Die Stanford Graduate School of Business geht noch weiter und limitiert den Zugang zu den Studenten im ersten Jahr.

'GRADUATE School of **BUSINESS**' is not listed in any of the chosen dictionaries. Yet, it is always preceded by a specific University's name thus combining to a proper name. In SL 409 and TL 277 above it is referred to *Stanford*. It can be summarized that the node **BUSINESS** and its collocate GRADUATE are in each instantiation borrowed as constituents of a proper name. As a consequence, GRADUATE does not qualifies as a shared lexical collocate of **BUSINESS**.

A total of 26 lexical collocates for **BUSINESS** were identified in the target language corpus. Nearly all of these collocates are Anglicisms. This leads to the assumption that particularly highly frequent Anglicisms foster the acceptance of additional Anglicisms as collocates, or generally in their co-text. Interestingly, none of the remaining three most frequent source language collocates of **BUSINESS**, i.e. MODEL, EXCHANGE or SMALL is among the 26 identified target language corpus collocates. The three remaining target language collocates, i.e. DESIGN, OBJECTS and EUROPEAN all qualified as collocates in the source language corpus according to quantitative criteria and thus shared collocates. It can be hypothesized that the particularly high frequency of an Anglicism leads to an increased adaptation of surrounding word-forms, i.e. Anglicisms as collocates.

BUYOUT

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5	TRANSLATION	Top 5	LC	T
143	16	✓ FIRMS ≙ (19)	Firmen	✓ FIRMEN ≙ (7)	3	49
		BILLION (12)	Milliarde	BRANCHE (5)		
		LEVERAGED (8)	fremdfinanziert	✓ GESCHÄFT ≙ (5)		
		✓ DEAL ≙ (6)	Geschäfte	–		
		TARGET (5)	Ziel	–		

The node **BUYOUT** shares two equivalent lexical collocates. Remarkably, no Anglicisms were identified as collocates within the target language corpus. The equivalent collocations are both semantically transparent and although DEAL is a widely used Anglicism, GESCHÄFT was preferred. All lexical target language collocates were located in R1 position. In many cases, collocates in the target language corpus in R1 position are connected with a hyphen. Here are some examples:

TL 2: Wenn die Buyout-Branche 2008 “noch ein Drittel der Deal-Aktivitäten aus dem Vorjahr“ schaffe […].

TL 3: Derzeit bestimmt im sogenannten Buyout-Geschäft vor allem die Höhe der Investitionen, was geht und was nicht.

TL 7: Wegen der Finanzkrise sind die Banken nicht mehr bereit, den Buyout-Firmen Geld zu leihen.

In the source language corpus, all instantiations of FIRMS and the majority of DEAL were also identified in R1 position combining to **BUYOUT** FIRMS and **BUYOUT** DEAL. Consider the examples below.

SL 19: […] to co-invest in a high profile buyout deal as well as in buyout firms themselves.

SL 23: As the buyout firms increased pressure on Mayer […].

SL 121: […] to avoid paying a fat premium, as they do in the buyout deal.

In none of the chosen dictionaries a corresponding entry for **BUYOUT** FIRMS or **BUYOUT** DEAL could be detected.

It seems that shared equivalent collocates of **BUYOUT** are preferably nouns behaving similarly within the source and target language. It is interesting that contrary to the popular practice not the entire combination is taken over, but that the source language collocates are substituted with target language equivalents. **BUYOUT**-GESCHÄFT and **BUYOUT**-FIRMEN exemplify some of the mechanisms of 'loan-translation'. 'Loan translation' is defined by Lyons as "[…] the translation of the constituent parts of a foreign word or phrase" (2004: 309). **BUYOUT**-GESCHÄFT and **BUYOUT**-FIRMEN can be perceived as a partial loan-translation. Because **BUYOUT** has already entered the German language as a frequently used Anglicism only its collocates, DEAL and FIRMS, are translated. It is somewhat surprising though, because DEAL is a frequently used Anglicism in the target language. It can be hypothesized whether this arises from the missing institutionalization of **BUYOUT**-DEAL and **BUYOUT**-FIRMS in the source language as a compound, because for neither a dictionary entry exists. The concepts of lexicalization and institutionalization, as an integral part of word-formation, are important in the context of lexical items that have not existed in a language for a very long time and are not yet fully established and accepted by the language community.

CASH

SOURCE LANGUAGE				TARGET LANGUAGE			
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
823	57	≡ FLOW (81)	(Zu)fluss	≡ FLOW (71)	12	172
		INVESTORS (24)	Anleger	OPERATIVE (11)		
		✓ ≙ STOCK (19)	Aktie/Wertpapier	✓ ≡ MANAGEMENT (10)		
		✓ ≡ MANAGEMENT (16)	Geschäftsführung	✓ ≙ AKTIE (6)		
		✓ ≙ CUSHION (15)	Kissen/Polster	✓ ≙ POLSTER (5)		

Naturally, the node **CASH** is often used to refer to monetary issues, in many cases denominating fixed amounts of money. Therefore, it is not surprising that generally measuring units and currencies are strongly represented among the lexical collocates in both corpora. One example which would qualify as an equivalent collocate is BILLION and MILLIARDEN. Below this is exemplified.

SL 2: Freescale expects to have $1.2 billion in cash on hand by June […].

TL 19: Pfizer sitzt auf zwölf Milliarden Dollar Cash und fährt im Jahr ebenso viel Cash-Flow ein.

Such identified equivalent collocates which are quantifiers and simply denote amounts of money or are names of currencies cannot be rated to indicate similar collocational behavior. Although they are not grammatical words, they come from a closed class and are thus not truly lexical collocates. Consequently, collocates which belong to this category were put on an additionally created stop list for the node **CASH**.

For the node **CASH** two identical and two equivalent lexical collocates were identified. The identical collocate FLOW always occurred in R1 position in the source and target language corpus without exception. Below one example from each corpus is given.

SL 665: […] on a calendar that displays estimated future cash flow based on when customers are paid[…].

TL 21: Er wird berechnet aus dem operativen Cash-Flow abzüglich von Posten wie zum Beispiel […].

There are different spelling conventions for **CASH** FLOW in the target language corpus. In many cases it is spelled with a hyphen (cf. TL 21 above). Clearly, the compound **CASH** FLOW is borrowed in the target language. **CASH** FLOW is defined by *Oxford Dictionaries* as "the total amount of money being transferred into and out of a business, especially as affecting liquidity." Thus, FLOW does not qualify as a shared identical lexical collocate. **CASH** FLOW is a source language compound which is also used in the target language corpus.

The second identical collocate MANAGEMENT is identified in both corpora in various positions within the collocational span, but most frequently in R1 position. Below, one example per corpus is given to illustrate this.

SL 613: […] the cash-management firm accused of mismanaging clients' money.

TL 105: Aber auch Mittelständler brauchen Leistungen wie Cash-Management oder Währungsabsicherung.

CASH MANAGEMENT is not listed in the chosen monolingual dictionaries, but *PONS* has a corresponding lexical entry and translates **CASH** MANAGEMENT with "Kassenhaltung" or simply the Anglicism "Cash-Management". Consequently, **CASH** MANAGEMENT qualifies as a compound according to the present study's criteria. As a result, MANAGEMENT cannot be counted as an identical lexical collocate whenever it

occurs in R1 position. This is the case for 2 instantiations in the source language corpus and 4 instantiations in the target language corpus.

Below, examples in which the collocate MANAGEMENT occupies different positions within the span in the source and target language corpora are provided.

SL 295: But when markets get frothy, management hoards cash rather than risk over-paying.

TL 98: Das Management brauchte Cash und verkaufte alles, was noch […].

In both examples MANAGAMENT occurs in L2 position. The remaining 6 co-occurrences of **CASH** and MANAGEMENT per corpus were analyzed qualitatively. As a result, and according to the concept of collocation used in this study, MANAGEMENT qualifies as an identical lexical collocate.

The first equivalent lexical collocate of **CASH** is STOCK and AKTIE. STOCK occurs in the source language corpus in L4, R2, R3 and R4 position of the collocational span. AKTIE occurs in L1, R2, R3 and R4 position in the target language corpus. For exemplification see below.

SL 172: Since 2003, some $490 billion in net new cash poured into international stock.

TL 117: Aktie statt Cash - viele Amerikaner investieren lieber in ein konservatives Portfolio […].

According to the concept of collocation used in this study and the study's definition of equivalent collocates, STOCK and AKTIE qualify as equivalent lexical collocates.

The second equivalent collocate CUSHION co-occurred 15 times with **CASH** in different positions within the collocational span. Here are a couple of examples from the source language corpus to exemplify this:

SL 89: Ross's cash cushion is also allowing for some opportunistic buying.

SL 183: You are building a cushion of cash so you don't have to rely on short-term funding.

CASH CUSHION is not listed as a compound in any of the chosen dictionaries, thus qualifies as a collocation. CUSHION is translated to German usually with "Kissen" or "Polster" (www.pons.eu). POLSTER was identified as a Top 5 lexical collocate in the target language corpus, usually in R1 position. Below a couple of examples are provided.

TL 49: Ein Cash-Polster auch für mögliche Verzögerungen könnte Schott gut gebrauchen.

TL 157: Immerhin hat die Aktie einige Vorteile zu bieten, das Cash-Polster von 1,1 Milliarden Euro zum Beispiel […].

CUSHION and POLSTER as lexical equivalent collocates of **CASH** can be perceived as another example of a partial loan translation (cf. discussion of **BUYOUT** above). One of the constituent parts of the foreign collocation **CASH** CUSHION is translated. Because **CASH** has already entered the German language as a frequently used Anglicism, only the second constituent of the source language collocation, CUSHION, is translated to POLSTER. Up to this point the qualitative analysis has already revealed that frequently source language compounds are used in identical form in the target language. Why the expression **CASH** CUSHION is not borrowed in its entirety remains unresolved. The qualitative corpus analysis allows for the educated guess that this is due to the missing institutionalization of **CASH** CUSION as a compound in the source language. At least in the chosen dictionaries no lexical entry exists. It cannot be ultimately dismissed that theoretically, **CASH** CUSHION and **CASH**-POLSTER may have arisen independently in English and German. As far as the qualitative analysis of the current study is concerned, CUSHION and POLSTER qualify as equivalent lexical collocates.

CORPORATE

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
667	112	MARKET	(26)	Markt/Handel	✓ GOVERNANCE ≡	(23)	6	56
		BONDS	(22)	Schuldverschreibung	RESPONSIBILITY	(9)		
		CUSTOMERS	(16)	Kunden	SOCIAL	(9)		
		WORLD	(14)	Welt/Erde	FINANCE	(5)		
		✓ GOVERNANCE ≡	(13)	Kontrolle	COLLECTION	(5)		

One identical lexical collocate was identified for **CORPORATE**. In the target language corpus, **GOVERNANCE** occurred exclusively in R1 position. Below this is illustrated with corpus examples.

TL 9: Dank der Bemühungen um eine gute Corporate Governance kam BMW auf Platz drei der […].

TL 11: Neben den Zahlen wird daher die Beurteilung der Corporate Governance immer wichtiger […].

In the source language corpus, GOVERNANCE occurred 9 out of 13 times in R1 position. See below for an example.

SL 127: […] and progress in corporate governance there has made it safer for investors, says the report.

Neither *Oxford Dictionaries,* nor *Merriam Webster,* or *PONS* lists **CORPORATE** GOVERNANCE as a compound. This might be owed to the fact that there is no universally accepted definition. However, the term is used frequently in the source and target language corpora and generally speaking **CORPORATE** GOVERNANCE "refers to the set of systems, principles and processes by which a company is governed" (www.businessdictionary.com). It may be best translated as 'Grundsätze der Unternehmensführung'. The status of **CORPORATE** GOVERNANCE as a compound seems arguable, but according to the criteria in the present study it does not qualify as a compound and GOVERNANCE is thus counted as an identical lexical collocate of **CORPORATE**.

All lexical collocates of **CORPORATE** in the target language corpus are Anglicisms. It should be noted that the collocates SOCIAL and RESPONSIBILITY combine in all instantiations to **CORPORATE** SOCIAL RESPONSIBILITY, which is a fixed expression, but not a compound according to the guidelines of the present study. This expression describes "a company's sense of responsibility towards the community and environment (both ecological and social) in which it operates" (www.businessdictionary.com). Below one instantiation from the target language corpus is presented.

TL 37: […] dass sie es mit ethischem Verhalten ernst meinen, dass sich hinter Corporate Social Responsibility nicht nur Marketing verbirgt.

SOCIAL and RESPONSIBILITY were also identified in the source language corpus as collocates, co-occurring 5 times each with **CORPORATE**. However, the expression **CORPORATE** SOCIAL RESPONSIBILITY only exists in the target language corpus and could not be identified in the source language corpus. This makes the term suspicious of being an Anglo neologism (cf. discussion of **BANKING** above) or maybe even a pseudo-Anglicism (cf. subsection 5.1.5). It is difficult to translate this expression, particularly because no lexical entry exists in the chosen dictionaries, but it may be translated as 'soziale Verantwortung des Unternehmens'.

The Top 5 target language collocate FINANCE was also identified as a collocate in the source language corpus and co-occurred 8 times with **CORPORATE**, mainly in R1 position. One example from each corpus is given below.

SL 42: […] you should also pursue positions in corporate finance to gain skills […].

TL 32: […] sowie Vorstandssprecher des dortigen Instituts für Rating und Corporate Finance im Mittelstand.

CORPORATE FINANCE is not listed as a compound in the chosen dictionaries. It refers to "the area that involves the financial aspects of a business or corporation. Financial aspects include accounting and investments" (www.businessdictionary.com) and can be translated as 'Unternehmensfinanzen' or 'Zentralabteilung Finanzen'. Thus, FINANCE could be counted as an identical lexical collocate. However, it did not co-occur with **CORPORATE** frequently enough in the source language corpus to make the Top 5.

COLLECTION is a frequent target language collocate, but could not be identified as a source language corpus collocate. It is used as in the following example:

TL 22: Er hat die entscheidenden Ideen für unsere Corporate Collection entwickelt.

A definition for **CORPORATE** COLLECTION could not be found. Judging from the context of TL 22 it refers to a collection of clothing which is embroidered with the company's logo. **CORPORATE** COLLECTION is very likely to be a pseudo-Anglicism (cf. subsection 5.1.5).

CRASH

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
55	5	✓ ≡ 1987 (7)	1987	JANUAR (9)	12	111
		✓ ≙ STOCK MARKET (7)	Aktienmarkt	✓ ≙ AKTIENMARKT (6)		
		FINANCIAL (6)	finanziell	GROßEN (6)		
		REAL ESTATE (5)	Immobilien/ Grundstücke	CASH (6)		
		PRICES (5)	Preise	✓ ≡ 1987 (5)		

The chart above shows that **CRASH** features one identical and one equivalent collocate. **CRASH** is more frequent in the target language corpus than in the source language one. In the target language corpus, **CRASH** does not collocate with Anglicisms at all, i.e. the node

only co-occurs with target language word-forms. This is typical of nodes which are more frequent in the target language corpus than in the source language corpus.

In both corpora **CRASH** co-occurred frequently with the year 1987. Below this is exemplified.

SL 16: In contrast, the crash of October 1987 was followed by a clear and quick recovery.

TL 86: […] Handelsprogramme, die schon 1987 für einen Crash gesorgt hatten.

In finance '1987' refers to the year in which stock markets crashed around the world for the first time after World War II. The crash began in Hong Kong and spread westward to Europe and the USA. This explains why the collocation is identified frequently in German and American business magazines alike.

The literal translation of STOCK MARKET is AKTIENMARKT and they are equivalent lexical collocates of **CRASH**. Both collocates occur in different positions within the collocational span in the source and target language corpora. Examples from each corpus are given below.

SL 5: The disasters of 2001--the terrorist attacks, the stock market crash, the collapse of Enron […].

SL 33: The crash of the stock market and a barrage of lawsuits […].

TL 2: Der Aktienmarkt-Crash schlägt deshalb auf die Policen durch.

TL 20: Der tiefe Fall der Häuserpreise und der Crash am Aktienmarkt haben die US-Konsumenten […].

STOCK MARKET and AKTIENMARKT qualify as equivalent lexical collocates.

DEAL

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
780	57	GOOD	(25)	gut	NEW	(19)	3	150
		BILLION	(22)	Milliarde	KÖNNTE	(6)		
		GREAT	(13)	groß(artig)	NEUE	(5)		
		MILLION	(13)	Million	-			
		BETTER	(12)	besser	-			

The node **DEAL** does not feature any shared identical or equivalent lexical collocates. Moreover, in the target language corpus **DEAL** almost exclusively features grammatical collocates. To some extent, this can also be seen from the chart above. It is unusual for a node with 150 corpus occurrences to only feature 3 lexical collocates. It should be noted that the most frequent lexical target language collocate is an Anglicism. NEW exclusively co-occurred with **DEAL** in L1 position. NEW **DEAL** is an expression used especially in politics. It is listed as a compound in *Oxford Dictionaries* with the following definition: "The economic measures introduced by Franklin D. Roosevelt in 1933 to counteract the effects of the Great Depression." *Merriam Webster* adds that "the term was taken from Roosevelt's speech accepting the 1932 presidential nomination, in which he promised "a new deal for the American people." NEW **DEAL** is in *PONS* translated as "Reformprogramm". **DEAL** can also refer to a single round of a card game, in the given context, NEW **DEAL** then also has a metaphoric meaning. NEW was not identified as a collocate of **DEAL** in the source language corpus. It is interesting that the English expression NEW **DEAL** which was coined by an American President is not used in the American English source language corpus, but instead used frequently in the German target language corpus.

DESIGN

SOURCE LANGUAGE				TARGET LANGUAGE			
T	LC	Top 5	TRANSLATION	Top 5	LC	T
385	29	NEW (18)	neu	✓ ≡ BUSINESS (23)	7	210
		✓ ≙ GOOD (10)	Gut	SERVICE (16)		
		FIRM (10)	Kanzlei/Firma/ Unternehmen	KUNST (7)		
		✓ ≡ BUSINESS (10)	Handel/Gewerbe	EURO (5)		
		INDUSTRIAL (9)	Industrie-	✓ ≙ GUTES (5)		

The word-form **DESIGN** can either be a noun or verb. 22 out of the 385 instantiations of **DESIGN** are verbs and were not counted as instantiations of the key word.

The node **DESIGN** features one identical and one equivalent lexical collocate. The identical collocate BUSINESS occurs in R2, R3 and R4 position in the source language corpus and in L2, L1, R2 position in the target language corpus. One example per corpus of the co-occurrence of BUSINESS and **DESIGN** is presented below.

SL 107: […] aimed to emphasize the importance of design to Brazilian business, which has surged in […].

TL 14: Der Apparat aus dem Business Design Laboratorium erkennt die Gesichter von bis zu zehn Menschen.

In 12 of the target language occurrences of BUSINESS it is used as follows:

TL 143: Wir hätten doch nicht in Business By Design investiert, wenn wir nicht der Überzeugung wären […].

Business By Design is the name of a software which is distributed by the company *SAP*. Thus, in those instantiations BUSINESS cannot be counted as an identical collocate. However, the remaining 11 co-occurrences of **DESIGN** and BUSINESS in the target language corpus allow for classification as an identical lexical collocate.

GOOD and GUTES is the identified Top 5 equivalent lexical collocate of the node **DESIGN**. In the target language corpus GUTES occurs in L1 position. In the source language corpus it occurs in L2, L1 and R2 position. Below, one example per corpus is provided.

SL 9: Good design also teaches us about your values.

SL 149: Two years ago we brought the awards to China, believing the message "good design is good business" would find fertile ground there.

SL 256: Yes, good hardware design is critical.

TL 27: Nur durch gutes Design wird eine architektonische Hülle überhaupt erst lebenswert.

The inflectional variants of the German adjective GUT, as shown in TL 61 below, were considered as translation equivalents of GOOD. In the chart for each node only the most frequent adjective form is given.

TL 61: […] oder die Bürokultur mit gutem Design verändern.

DESIGN is usually associated with positive meaning in the Western World. Therefore, **DESIGN** co-occurs often with other positively connoted adjectives like 'energy-saving', 'great', 'ideal' in the source language corpus or 'modernes', 'exzellentes', 'lebendiges' in the target language corpus. Because they do not co-occur frequently enough with **DESIGN**, none of these adjectives qualify as equivalent collocates though, except for GOOD and GUTES.

DEVELOPMENT

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
318	27	RESEARCH	(48)	Forschung	BUSINESS ≡	4	29
		PRODUCT	(24)	Produkt	BANK (7)		
		ECONOMIC	(21)	wirtschaftlich	CHINA (5)		
		BUSINESS ≡	(19)	Handel/Gewerbe	MANAGEMENT (5)		
		DIRECTOR	(16)	Direktor/Leiter	-		

As illustrated above, **DEVELOPMENT** shares one identical lexical collocate. In both corpora BUSINESS was typically identified in L1 position of the collocational span. See below for examples.

SL 92: A former NAF staff employee familiar with its business development efforts says […].

TL 10: […] sagt der Unternehmensberater für Business Development, der seit seinem ersten Handy […].

PONS translates BUSINESS **DEVELOPMENT** to German as "Geschäfts(feld)entwicklung". No entry in the monolingual dictionaries was found. BUSINESS **DEVELOPMENT** means different things to different people, thus a precise definition is not attempted. Mostly, BUSINESS **DEVELOPMENT** "comprises a number of tasks and processes generally aiming at developing and implementing growth opportunities" (www.businessdictionary.com). BUSINESS co-occurs frequently in L1 position with the node **DEVELOPMENT**. In addition, a dictionary entry in *PONS* exists for BUSINESS **DEVELOPMENT**. As a consequence, BUSINESS cannot be counted as an identical lexical collocate of **DEVELOPMENT.**

The Anglicism and target language collocate MANAGEMENT was not identified as a source language collocates.

ECONOMY

SOURCE LANGUAGE					TARGET LANGUAGE		
T	**LC**	**Top 5**		**TRANSLATION**	**Top 5**	**LC**	**T**
1,250	215	GLOBAL	(57)	weltweit	NEW ≡	7	68
		WEAK	(41)	schwach	BLASE	(48)	
		FUEL	(28)	Treibstoff/ Benzin	PLATZEN	(15)	
		RECESSION	(24)	Rezession	OLD	(10)	
		NEW	≡ (22)	neu	CLASS	(9)	
						(5)	

The node **ECONOMY** shares one identical lexical collocate in the source and target language corpora. In the source language corpus NEW was identified in L4, L2, L1, R1, R2, R3 and R4 position. Most frequently NEW co-occurred with **ECONOMY** in L2 and L1. In the target language corpus NEW occurred exclusively in L1 position forming NEW **ECONOMY**. An example from both corpora in which NEW occurs in L1 is given below.

SL 584: The Nasdaq, the index of our New Economy hopes and dreams, crashed.

TL 9: Das Ende der New Economy und damit verbundene Milliardenverluste bereiteten allen Fantasien ein jähes Ende.

Oxford Dictionaries defines NEW **ECONOMY** as "new industries, such as biotechnology or the Internet, that are characterized by cutting-edge technology and high growth." *PONS* lists as a translation for NEW **ECONOMY**: "New Economy (Wirtschaftszweige der modernen Technologien)". Thus, a German equivalent does not even exist and is unlikely to

be coined. However, although NEW and **ECONOMY** may actually collocate in the source language, NEW **ECONOMY** is borrowed as a compound and thus NEW does not qualify as an identical lexical collocate of **ECONOMY**.

CLASS and OLD, which are Anglicisms and Top 5 target language corpus collocates, were not identified as collocates in the source language corpus.

EQUITY

SOURCE LANGUAGE				TARGET LANGUAGE			
T	LC	Top 5	TRANSLATION	Top 5		LC	T
539	67	≡ PRIVATE (276)	privat	≡ PRIVATE (178)		14	192
		HOME (78)	Eigenheim/ Zuhause	✓ ≙ FONDS (27)			
		✓ ≙ FIRMS (53)	Firmen	INVESTOREN (19)			
		✓ ≙ FUNDS (42)	Fonds	✓ ≙ FIRMEN (17)			
		LOANS (27)	Anleihen	MANAGER (13)			

The chart above shows that **EQUITY** shares one identical and two equivalent lexical collocates. The extraordinarily frequent co-occurrence of PRIVATE and **EQUITY** already hints to the fact that the combination forms a compound. Furthermore, PRIVATE occurs exclusively in L1 position in the source and target language corpora. Below examples are provided.

SL 10: Private equity, venture capital, real estate, and other risky plays account for over 30% of its portfolio […].

TL 79: Dank Private Equity hat Gerresheimer bei einem Börsenwert von 980 Millionen Euro heute 460 Millionen Euro Nettoschulden an der Backe.

PRIVATE is thus not counted as an identical lexical collocate.

EQUITY tends to co-occur with PRIVATE in L1 and another collocate in R1 simultaneously. Some R1 collocates qualify as equivalent lexical collocates according to the quantitative analysis. Examples are PRIVATE **EQUITY** FIRMS and PRIVATE-**EQUITY**-FIRMEN or, less frequently, PRIVATE-**EQUITY**-UNTERNEHMEN. Below, examples from both corpora are provided.

SL 43: While private equity firms control just a tiny fraction of U.S. corporations, their companies are disproportionately troubled.

TL 36: So musste Boss früher oder später auf dem Radar von Private-Equity-Firmen erscheinen.

TL 67: […] vermehrt in das Segment vorwagen, das sonst eher die turbokapitalistischen Private-Equity-Unternehmen betreiben.

FIRMS is identified as a Top 5 source language collocate. Equivalent target language collocates thus need to occur also in the plural form. The singular and plural form of UNTERNEHMEN is identical in German, but only those instantiations in which the plural is used are counted. PRIVATE **EQUITY** FIRMS is not listed in the chosen mono- or bilingual dictionaries (and neither is the singular form). FIRMS and FIRMEN collocate exclusively in R1 with **EQUITY**. UNTERNEHMEN collocates with **EQUITY** in many instantiations in different positions. An example is given below.

TL 54: Diese Normalisierung wird für die Unternehmen in Private-Equity-Hand eine harte Zeit.

Thus, the co-occurrence of FIRMS, FIRMEN/UNTERNEHMEN with **EQUITY** shows similar collocational behavior.

The second equivalent lexical collocate FUNDS and FONDS also shows a strong tendency to co-occur with PRIVATE **EQUITY** primarily in R1 position. Below, one example per corpus is given.

SL 449: That could hurt private equity funds – and their investors.

TL 130: Zu hohe Ausgabepreise hatten vor allem Papiere, die von Private-Equity-Fonds bei Investoren abgeladen wurden.

It should be noted that **EQUITY** and FUNDS also co-occurred without being preceded by PRIVATE. This is exemplified below.

SL 125: […] which formed six equity funds that acquired more than 50 companies worth $7 billion.

PRIVATE **EQUITY** FUND has an entry in *PONS* and translates to “Private-Equity-Fonds”. In those instantiations in which FUNDS and FONDS occur in R1 position they can therefore not be counted as equivalent lexical collocates of **EQUITY**. However, FUNDS and FONDS were also identified in all remaining positions of the collocational span in both corpora. To be precise, only in 13 out of 42 co-occurrences FUNDS was identified in R1 position in the source language corpus and in 17 out of 27 co-occurrences

FONDS in the target language corpus. Examples of FUNDS and FONDS in different positions with the span are provided below.

SL 194: To raise funds from banks or equity markets, entrepreneurs often join the Party […].

SL 274: […] public investment funds that take equity stakes in alternative energy ventures.

TL 72: Geschlossene Fonds, Hedgefonds oder Private-Equity-Beteiligungen unterliegen keinerlei Sicherung.

TL 161: Wir diskutieren auch über Investitionen in Infrastrukturprojekte und Beteiligungen an Private Equity, wie das andere Fonds auch machen.

As illustrated above, in such instantiations FUNDS and FONDS qualify as equivalent lexical collocates of **EQUITY**.

FINANCE

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
414	70	PROFESSOR	(36)	Professor	MANAGEMENT	(6)	8	43
		MORTGAGE	(14)	Hypothek	WEEK	(6)		
		≡ INTERNATIONAL	(14)	international	≙ SCHOOL	(6)		
		≙ UNIVERSITY	(13)	Universität/ Hochschule	✓ ≡ CORPORATE	(5)		
		✓ ≡ CORPORATE	(12)	Konzern-	≡ INTERNATIONAL	(5)		

FINANCE features two identical collocates and one equivalent lexical collocate according to the quantitative analysis. Interestingly, the collocates UNIVERSITY and SCHOOL which qualify as equivalent collocates are both Anglicisms, i.e. source language word-forms but not identical. UNIVERSITY, much like SCHOOL, is translated by *PONS* to "Hochschule". The *WordSmith Tools* analyses reveal that UNIVERSITY and SCHOOL take different positions within the span in the source and target language corpora. UNIVERSITY occurs preferably in R3 position and SCHOOL most frequently in L2 position. See below for examples.

SL 266: Says Michael Pettis, a professor of finance at Beijing University.

TL 15: […] sagt Uwe Wystup, Professor an der Frankfurt School of Finance & Management.

In the target language corpus, SCHOOL was only identified as a constituent of the proper name *Frankfurt School of Finance*. Thus, it cannot be counted as an equivalent lexical collocate of **FINANCE**. It should be mentioned that SCHOOL could be identified among the less frequent source language collocates of **FINANCE**. The same applies vice versa to UNIVERSITY in the target language corpus.

The identical collocate INTERNATIONAL was identified in various positions within the collocational span of **FINANCE** in the source and target language corpora. In the target language corpus, however, INTERNATIONAL was only identified in proper names such as:

TL 8: […] und das Institute of International Finance eine "Überprüfung" der Pflicht […].

TL 37: […] des Shanghai International Banking and Finance Institute (SIBFI) […].

TL 43: […] das Gipfeltreffen von Weltbank und International Finance Corporation (IFC) […].

Thus, INTERNATIONAL cannot be counted as an equivalent lexical collocate of **FINANCE**.

The other identical collocate of **FINANCE**, i.e. CORPORATE, qualifies as an identical lexical collocate although CORPORATE occurs preferably in L1 position. This was discussed previously (cf. analysis of **CORPORATE** above) and is repeated here for convenience:

> CORPORATE **FINANCE** is not listed as a compound in the chosen dictionaries. It means “the area that involves the financial aspects of a business or corporation. Financial aspects include accounting and investments” (www.businessdictionary.com).

Again, a couple of examples of the use of CORPORATE **FINANCE** in the source and target language corpus are given below.

SL 149: A move by hedge funds into traditional corporate finance would mean even less transparency […].

TL 23: Zuletzt wurden die in den Einzelunternehmen agierenden Finanzabteilungen zur Schörghuber Corporate Finance gebündelt […].

GROUP

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
996	158	INTERNATIONAL	(41)	international	≡ CONSULTING (14)	11	156
		✓ ≡ INVESTMENT	(36)	Investition	≡ CAPITAL (11)		
		FINANCIAL	(34)	finanziell	✓ ≡ INVESTMENT (5)		
		≡ CONSULTING	(33)	beratend/ Beratung	EUROPEAN (5)		
		≡ CAPITAL	(31)	Vermögen/ Kapital	DERIVATIVES (5)		

For the node **GROUP**, three identical lexical collocates were identified. The collocate INVESTMENT was identified in all positions of the span within the source language corpus and in L3, L1 and R2 position in the target language corpus. A few examples are given below.

SL 288: Morgan Stanley (MS), and Dubai's state-owned investment group[28] took a 6.5% stake in MGM Mirage (MGM).

SL 447: […] chief Ken Lewis wanted an investment banking group of his own, especially after closing […].

TL 66: KPN ist bisher ein gutes Investment für die Group.

TL 70: Ng, 59, ist Managing Director und Group Chief Investment Officer beim Staatsfonds […].

In the chosen dictionaries no entry exists for INVESTMENT **GROUP**. The fact that INVESTMENT and **GROUP** co-occur also in different positions of the collocational span in both corpora proves that INVESTMENT is an identical collocate of the node **GROUP**.

The second identical lexical collocate of **GROUP** is CONSULTING. CONSULTING occurs in the source language corpus in the positions L4, L1 and R2, but in the target language corpus only in L1 position. In the source language corpus, CONSULTING also collocates with **GROUP** most frequently in L1 position. A dictionary entry for CONSULTING **GROUP** in the chosen dictionaries does not exist. It is thus not a compound. Examples of the co-occurrence of CONSULTING and **GROUP** from the corpora are presented below.

[28] It should be noted that in all corpus examples GROUP refers to a "Konzern" or "Unternehmensgruppe" (www.pons.eu). This is not obvious with such limited context given.

SL 282: […] they hired Virtue Ventures, a consulting group specializing in social enterprises […].

SL 283: Tim Armstrong, emerging markets director at consulting group Global Insight.

TL 123: Als Hans-Paul Bürkner vor vier Jahren an die Spitze der Unternehmensberatung Boston Consulting Group (BCG) rückte […].

The target language example above shows that in this case CONSULTING and **GROUP** are constituents of a company name. While the company names vary, this holds true for all instantiations of CONSULTING in the target language corpus. Thus, it cannot be counted as an identical lexical collocate. In the source language corpus this can also be observed. For examples in which either just **GROUP** or CONSULTING and **GROUP** both form part of a company name see below.

SL 244: […] a study done by consulting firm Oliver Wyman Group showing the company's manufacturing efficiency […].

SL 504: At Boston Consulting Group, employees with 18 months at the firm […].

BusinessWeek and *WirtschaftsWoche* both cover issues concerning the same enterprises. Consequently, it is not surprising that identical company names turn up during the corpus analysis. In many instantiations CONSULTING is a collocate of **GROUP** in the source language corpus. But this has not been taken over in the German language yet.

The third and last identical collocate among the Top 5 for **GROUP** is CAPITAL. It needs to be pointed out that CAPITAL has at least two meanings. *Oxford Dictionaries* defines them as follows:

capital$_1$ - (also capital city or town) the city or town that functions as the seat of government and administrative centre of a country or region.

capital$_2$ - wealth in the form of money or other assets owned by a person or organization or available for a purpose such as starting a company or investing.

At first, it might seem rather unlikely that especially in a business magazine ‘capital$_1$’ is used frequently. However, in the source language corpus in which CAPITAL was identified as a collocate for **GROUP** it is used once with the meaning of ‘capital$_1$’. See below for the example.

SL 306: Consumers United for Rail Equity, a shippers' lobbying group in the capital.

CAPITAL occurs in L3, L2, L1, R2, R3 and R4 position of the span within the source language corpus. Within the target language corpus it was identified in L2 and L1 position. In all L positions in the source language corpus **GROUP** and CAPITAL only co-occurred as constituents of company names such as *Capital and Blackstone Group*, *Capital Advisors Group*, *Avenue Capital Group*. The same is true for all instantiations in which CAPITAL co-occurs in R2, R3 and R4 position with **GROUP**:

SL 230: […] half the level of the same period in 2007, research group Real Capital Analytics estimates.

SL 791: […] Dubai investment group Istithmar World Capital, has provided […].

SL 806: Other big investors include Dubai Financial Group and Dubai International Capital, controlled by […].

In the target language corpus, **GROUP** and CAPITAL were also exclusively identified as constituents of company names. Below, corresponding examples of CAPITAL in L2 and L1 position are given.

TL 129: HSBC-Trinkaus hat sich deshalb mit der Capital Efficiency Group aus dem schweizerischen Zug zusammengetan […].

TL 21: Inzwischen hält die Capital Group rund 30 Prozent und ist damit zum einzigen Großaktionär […].

In sum, CAPITAL does not qualify as a source language collocate or a shared identical lexical collocate for the study.

The remaining two Top 5 target language collocates EUROPEAN and DERIVATIVES are Anglicisms. None of the two were identified as collocates in the source language corpus.

IMAGE

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5	TRANSLATION	Top 5	LC	T
105	6	✓ BRAND (9)	Marke	✓ MARKE (6)	15	201
		FRIENDLY (7)	freundlich	✓ UNTERNEHMEN (5)		
		PROBLEMS (7)	Probleme	GRÜNES (5)		
		✓ COMPANIES (5)	Unternehmen	HÖHERE (5)		
		BUILDING (5)	Gebäude/Bau	KONZERN (5)		

The node **IMAGE** has two equivalent lexical collocates. Interestingly, **IMAGE** is among the few nodes which feature German collocates only. In addition, neither source nor target language collocates were identified directly adjacent to the node, i.e. in L1 or R1 position. Because their constituents may not be separated by other forms (cf. Adams 1973: 30), borrowing in this case is definitely not motivated by the adoption of compounds. **IMAGE** is used more frequently in the target language corpus than in the source language corpus. As the analysis advances, it will show whether similar observations also apply to **INSIDER** and **MANAGER** (cf. subsection 6.2.1). Noticeably, **CRASH** featured almost exclusively German lexical collocates in the target language corpus; with the exception of CASH.

The first equivalent collocate of **IMAGE** is BRAND and MARKE. BRAND occurred in L4, L1, R3, R4 in the source language corpus and MARKE in R2 and R3 in the target language corpus. Compounding can be excluded as a borrowing mechanism here. Below, examples of the co-occurrence of **IMAGE** and BRAND/MARKE are given.

SL 20: "Where we fell short was in the image of the brand."

SL 85: Eager to reinvigorate its image as a basketball brand, Converse this fall began organ […].

SL 88: […] selling their products or services, or creating a positive brand image with their customers.

TL 89: Das Image einer Marke hängt unserer Erfahrung nach dem Produkt und dessen Qualität […].

TL 121: Das Image der eigenen Marke könnte unter der Kooperation leiden […].

TL 146: Sie setzen auf Design und das Image der Marke.

BRAND and MARKE qualify as equivalent lexical collocates of **IMAGE**.

The second equivalent lexical collocate of **IMAGE** which was identified in the quantitative corpus analysis is COMPANIES and UNTERNEHMEN. In the source language corpus, COMPANIES was identified in different positions left and right of the node. The same can be said for UNTERNEHMEN and the target language corpus. As has been mentioned previously (cf. **EQUITY** above), the singular and plural form of UNTERNEHMEN is identical in German. Again, only those instantiations in which the plural meaning is used were counted as equivalents to COMPANIES. Below one example per corpus is provided.

SL 7: The public has a poor image of drug companies.

TL 177: könnte dies auch das Image der Unternehmen in Mitleidenschaft ziehen, die mit ihren Sponsorengeldern […].

The qualitative analysis confirms that COMPANIES and UNTERNEHMEN are equivalent collocates of **IMAGE**. Thus, similar collocational behavior of **IMAGE** in the source and target language corpora is verified.

INSIDER

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
30	3	TRADING (≙ (6))	Handel	BERICHTEN (20)	13	195
		FORMER (5)	ehemalig	SAGEN (8)		
		CHARGES (5)	Kosten/Gebühren	HABEN (7)		
		–	–	HANDEL (≙ (5))		
		–	–	WISSEN (5)		

The node **INSIDER** is one of the four nodes which occur more frequently in the target language than the source language corpus. The node features only German collocates in the target language corpus. This had already been observed with the node **CRASH** and **IMAGE** and seems to be typical. Apparently, nodes collocate typically with target language word-forms, instead of additional Anglicisms, if their overall frequency is higher in the target language than in the source language corpus.

INSIDER has been lexicalized in the target language. *PONS* does not even list a translation for **INSIDER** other than the Anglicism “Insider”. Other attempts to translate **INSIDER** result in ‘Eingeweihter’ or ‘Zugehöriger’ and are rather bulky. Thus, possibly lexicalized nodes in the target language usually co-occur with target language word-forms rather than

additional Anglicisms. Perhaps the Anglicism **INSIDER** and its collocates are not encountered frequently enough in the source language to be adopted or imitated in the target language.

The node **INSIDER** features one equivalent collocate. TRADING and HANDEL were only identified in R1 position forming the compounds **INSIDER** TRADING and **INSIDER**-HANDEL. The less frequent translation of **INSIDER** TRADING, namely **INSIDER**-GESCHÄFT was also identified twice in the target language corpus. Below examples are presented.

SL 14: […] to challenge a plan by bringing an insider trading case against an executive […].

TL 98: Alter: 61. Frühere Position: EADS-Chef. Verdacht: Insider-Geschäft.

TL 131: Wegen Verdachts auf Insider-Handel und Kursmanipulation muss Straub nun sein Amt […].

Oxford Dictionaries defines **INSIDER** TRADING as "the illegal practice of trading on the stock exchange to one's own advantage through having access to confidential information." TRADING and HANDEL/GESCHÄFT do not qualify as equivalent lexical collocates, because they form compounds. It should also be noted that **INSIDER** is one of the few nodes which takes a very high number of verbs as lexical collocates in the target language corpus, i.e. SAGEN, HABEN, WISSEN for example.

INSIGHT

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
39	4	≡ GLOBAL (7)	weltweit	≡ GLOBAL (33)	7	35
		FIRM (5)	Kanzlei/Firma/ Unternehmen	BERATUNGSUNTERNEHMEN (6)		
		RESEARCH (5)	Forschung	ANALYST (6)		
		NEW (5)	neu	EXPERT (5)		
		-	-	FINANZDIENSTLEISTER (5)		

The node **INSIGHT** features one identical collocate. In both corpora GLOBAL co-occurs with INSIGHT exclusively in L1 position. When looking at the *WordSmith Tools* analyses of INSIGHT, it is obvious that GLOBAL and **INSIGHT** only co-occur as constituents of the company name *Global Insight*. See below for corpus examples.

SL 2: […] says Christoph Stürmer, an analyst at researcher Global Insight.

TL 4: […] sagt Roman Mathyssek, Autoanalyst beim Beratungsunternehmen Global Insight in London.

GLOBAL is not an identical lexical collocate of **INSIGHT.**

INVESTMENT

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5	TRANSLATION	Top 5	LC	T
1,240	162	≡ BANKING (139)	Bankwesen	FONDS (9)	21	177
		BANK (103)	Bank	CORPORATION (7)		
		BANKS (97)	Banken	≡ BANKING (5)		
		FINANCIAL (96)	finanziell	AUTHORITY (5)		
		FIRM (79)	Kanzlei/Firma/ Unternehmen	RETURN (5)		

The node **INVESTMENT** features one identical collocate according to the quantitative corpus analysis. The analysis of **BANKING** above has already shown that **INVESTMENT** and **BANKING** co-occur very frequently and form the compound INVESTMENT BANKING. Thus, BANKING does not qualify as an identical lexical collocate of **INVESTMENT**.

Four of the Top 5 target language collocates of **INVESTMENT** are Anglicisms. AUTHORITY and RETURN were also identified as source language collocates and co-occurred with **INVESTMENT** 18 and 13 times respectively. This holds true for additional collocates below the Top 5 threshold and is illustrated in Figure 27 below.

SOURCE LANGUAGE		TARGET LANGUAGE	
ADDITIONAL COLLOCATES		ADDITIONAL COLLOCATES	
GROUP	(36)	GROUP	(5)
GOVERNMENT	(13)	GOVERNMENT	(5)
FUND	(42)	FUND	(5)
ESTATE	(24)	ESTATE	(5)

Figure 27 Additional collocates of INVESTMENT

In addition to the Top 5 collocates, four more Anglicisms were identified as target language lexical collocates. And although they did not rank among the Top 5, in the source language all of them were identified as source language collocates as well.

When looking at Figure 27 it becomes clear that **INVESTMENT** features more identical lexical collocates than were allowed in the Top 5 group. This hints at a weak spot of the methodology used in this study for very frequent nodes. It is more difficult to analyze them reliably with the chosen Top 5 methodology.

INVESTOR

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
247	24	FUND	(7)	Fonds	≡ RELATIONS (13)	16	199
		GROUP	(7)	Gruppe/Konzern	✓ ≡ PRIVATE (7)		
		≡ RELATIONS	(6)	Beziehungen	STRATEGISCHEN	(5)	
		SECURITIES	(6)	Wertpapiere	VERKAUFT	(5)	
		✓ ≡ PRIVATE	(5)	privat	EQUITY	(5)	

The node **INVESTOR** shares two identical lexical collocates. The collocate RELATIONS occurs exclusively in R1 position in the source language corpus and the same holds true for the target language corpus. The compound **INVESTOR** RELATIONS is not listed in the monolingual dictionaries, but in *PONS*. *PONS* provides three translations for **INVESTOR** RELATIONS: 1) “Investor Relations”, 2) “Aktionärspflege”, 3) “Kapital-geberbezie-hungen”. Neither ‘Aktionärspflege’, nor ‘Kapitalgeberbezie-hungen’ is used in the target language corpus. The examples below illustrate the use of **INVESTOR** RELATIONS in both corpora.

SL 48: Stephen Campbell, vice-president for investor relations at Newfield, says that its portfolio […].

SL 195: It's almost like having an extra staff member on my investor relations team.

TL 20: In Frankfurt wird der "Deutsche Investor Relations Preis 2008" vergeben.

TL 49: Die Folge: Management und Investor Relations der großen Banken verfielen teilweise in […].

The qualitative corpus analysis thus shows that RELATIONS does not qualify as an identical lexical collocate.

With the second identical Top 5 collocate, PRIVATE, the case is different. PRIVATE is identified in the source language corpus in L3, L2, L1 and R2 position and in the target language corpus in L3, L2, R2, R3, R4 position. Compounding can be excluded. Below, corpus examples of the co-occurrence of **INVESTOR** and PRIVATE are given.

SL 56: Tata International, India's largest private investor in Africa […].

SL 76: In June, Swedish private equity firm investor AB bought a majority stake […].

SL 167: […] says one investor in private equity funds.

SL 174: […] bringing in $360 million in new money from private equity investor MacFarlane Partners.

TL 33: Einen Aufkauf des weltgrößten Stahlkonzerns könnte ein Private-Equity-Investor binnen nur sechs Jahren […].

TL 51: […] sagt Investor Albrecht, dessen Private-Equity-Firma dem Vernehmen nach […].

It should be mentioned that the English and German word-form PRIVATE are identical. However, the qualitative analysis revealed that only two of the concordance lines in which PRIVATE was identified as a collocate in the target language corpus are about the German word-form PRIVATE. Both examples in question are given below.

TL 150: Der gesuchte private Investor soll sich nun verpflichten, das Straßennetz 25 Jahre lang […].

TL 165: Als strategischen Investor empfehlen sie eine private Krankenhauskette.

Nevertheless, PRIVATE qualifies as an identical lexical collocate because in 5 instantiations in the target corpus it was clearly the English word-form PRIVATE which collocates with **INVESTOR** (cf. examples TL 33 an TL 51 above).

MANAGEMENT

SOURCE LANGUAGE				TARGET LANGUAGE			
T	LC	Top 5	TRANSLATION	Top 5	LC	T
840	76	≡ CAPITAL (81)	Vermögen/ Kapital	TOP (35)	53	523
		≡ ASSET (52)	Vermögenswert	SCHOOL (34)		
		INVESTMENT (49)	Investition	≡ ASSET (18)		
		≡ SCHOOL (36)	Fakultät/ Hochschule	≡ CAPITAL (13)		
		✓ ≙ FIRM (39)	Firma/Kanzlei/ Unternehmen	✓ ≙ UNTERNEHMEN (13)		

The node **MANAGEMENT** features the exceptional number of three identical collocates. In addition, the node shares one equivalent collocate.

The identical collocate CAPITAL was identified in L3, L2, L1 and R2, R3 position in the source language corpus. CAPITAL was identified 71 times in L1 position. In the target language corpus CAPITAL only co-occurs with **MANAGEMENT** in L1. Despite this high frequency of the combination, a dictionary entry for CAPITAL **MANAGEMENT** does not exist in the chosen mono- and bilingual dictionaries. CAPITAL **MANAGEMENT** is used as follows in the source and target language corpora:

SL 680: […] the 1998 collapse of Long-Term Capital Management

TL 482: […] als der Hedgefonds Long-Term Capital Management kollabierte und die Finanzmärkte […].

It is obvious that, in the chosen corpus examples, CAPITAL and **MANAGMENT** form part of the identical company name *Long-Term Capital Management*. In all cases in which CAPITAL co-occurred with **MANAGEMENT** in L1 both word-forms are part of a company name. This extends to the target and source language corpus alike. CAPITAL is not an identical lexical collocate of **MANAGEMENT**.

Within both corpora ASSET and **MANAGEMENT** form the compound ASSET **MANAGEMENT**, "Vermögensverwaltung" (www.pons.eu), in German. The two terms are used synonymously. 'Vermögensverwaltung' occurs 33 times in the target language corpus and ASSET

MANAGEMENT 18 times. One example per corpus of the use of ASSET **MANAGEMENT** is provided below.

SL 66: Not surprisingly, that money pot – and the fat asset management fees it will generate – has financial-services firms salivating.

TL 497: Michael Keppler, der Gründer der bankunabhängigen Fondsgesellschaft Keppler Asset Management in New York.

When looking at the two examples above it becomes clear immediately that in TL 497 ASSET **MANAGEMENT** forms part of a company name: *Keppler Asset Management Inc.*, which was founded in 1992 by Michael Keppler. In all 18 instantiations in which ASSET **MANAGEMENT** is used in the target language corpus this is the case. This can also be observed in the source language corpus, but at the same time the collocation ASSET **MANAGEMENT** occurs frequently outside of company names. In sum, ASSET does not qualify as an identical lexical collocate of **MANAGEMENT**.

The third identical collocate SCHOOL is retrieved in the source language corpus in L3, L2 and R4 position.

SL 229: […] at the Johnson Graduate School of Management at Cornell University.

SL 614: […] professionalization of their career management starting in high school.

SL 635: […] at the Thunderbird School of Global Management.

Out of 36 co-occurrences of **MANAGEMENT** and SCHOOL in the source language corpus, the collocate SCHOOL was identified 31 times in L2 position forming the expression SCHOOL of **MANAGEMENT**. In the target language corpus, SCHOOL was identified in L4, L3, L2, R1 and R4 position as a collocate of **MANAGEMENT**. 26 out of 34 times SCHOOL co-occurred with **MANAGEMENT** in L2. Examples from the target language corpus for the co-occurrence with SCHOOL in the different span positions are presented below.

TL 130: […] Kühne School of Logistics and Management […].

TL 137: […] Otto Beisheim School of Management […].

TL 387: […] Lancaster University Management School […].

TL 394: […] Frankfurt School of Finance & Management.

TL 502: Institut für Private Wealth Management der European Business School […].

Because SCHOOL and **MANAGEMENT** co-occur in the target language corpus only as constituents of the names of educational institutions, SCHOOL does not qualify as an identical collocate.

The equivalent collocate UNTERNEHMEN does not occur in L1 or R1 position. Thus, the borrowing of hybrid compounds can be excluded. In the source language corpus **MANAGEMENT** and FIRM collocate in all positions of the span except for L1. Examples are presented below.

SL 296: […] CEO of money management and advisory firm Hamilton Lane.

SL 314: BusinessWeek has learned that a prominent money management firm plans to file suit […].

SL 485: […] according to the management consulting firm.

SL 576: […] co-head of portfolio management at the investment firm BlackRock.

SL 591: Abernathy's firm is a crisis-management boutique that is currently helping Yahoo!

For all instantiations of FIRM which were considered for the qualitative analysis it was secured that they concern the noun FIRM, not the adjective. Related to this issue was the disambiguation of the identical German singular and plural form of UNTERNEHMEN. Only the singular instantiations of UNTERNEHMEN were counted, because the plural meaning would not be equivalent to FIRM. In the target language corpus, UNTERNEHMEN was identified as a collocate in L4, L3, R2, R3 and R4. See below for examples.

TL 74: […] gut geführtes Unternehmen mit einem herausragenden Management.

TL 465: In Misskredit ist ein großes Unternehmen und sein Management in jüngster Zeit vor allem in Deutschland durch […].

TL 454: Das Management hat das Unternehmen nicht mehr richtig im Griff.

Consequently it can be recorded that MANAGEMENT shares one equivalent lexical collocate according to the quantitative and qualitative corpus analysis. Note that in the target language corpus neither FIRM, nor FIRMA qualified as collocates.

The node **MANAGEMENT** is one of the most frequent nodes in this study. It has been pointed out earlier that analysis of particularly frequent nodes according to the Top 5 collocate methodology has its disadvantages. In the following, a few remarks on the nature of the remaining 48 lexical target language corpus collocates are made. First of all, the top target language collocate TOP, was not identified as a collocate in the source language corpus at all. Additional 16 target language lexical collocates of **MANAGEMENT** are

Anglicisms: CONSULTANTS, EUROPEAN, CONSULTING, PRIVATE, WEALTH, FINANCE, BUSINESS, GRADUATE, SUPPLY, HOLDING, CASH, ACQUISITIONS, DEVELOPMENT, LONG, MERGERS, TERM. Out of these, 13 are also collocates in the source language corpus. They will not count towards the qualitative collocate count because only Top 5 collocates are considered. However, it can be stated on the basis of the numerous identical collocates that **MANAGEMENT** does show similar collocational behavior in the source and target language corpora. With so many instantiations of a node it is very useful to compare all target language collocates which are Anglicisms to all source language collocates. This also applies to the next node.

MANAGER

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
488	57	✓ ≙ FUND (44)	Fond	TOP (108)	135	1,330
		MONEY (42)	Geld	✓ ≙ FONDS (53)		
		PORTFOLIO (36)	Wertpapierbestand	DEUTSCHE (30)		
		GENERAL (30)	allgemein	UNTERNEHMER (26)		
		INVESTMENT (20)	Investition	EHEMALIGE (16)		

The node **MANAGER** is much more frequent in the target language corpus than in the source language one. In line with earlier results this brings about mostly German collocates in the target language corpus. One equivalent collocate was identified. FUND and FONDS were habitually retrieved in L1 position in both corpora, but can be identified in all other positions of the collocational span within both corpora as well. Examples are given below.

SL 19: Mohamed El-Erian was a top-ranked fund manager at Pacific Investment Management before joining […].

SL 206: Another recently reopened fund whose manager is a stickler for strong balance sheets is the […].

SL 321: […] was merged into PMFM, a fund provider and money manager.

SL 328: […] Ivy's chief investment officer and manager of the fund since 1997.

TL 846: […] sagt der Manager eines Londoner Fonds […].

TL 857: Das wichtigste Kriterium bei der Auswahl eines Fonds ist deshalb ein Manager, der über Jahre bewiesen hat […].

TL 868: Einige Fonds-Manager wetteten früh gegen den Subprime-Markt und machten damit Milliarden.

FUND and FONDS are equivalent lexical collocates of **MANAGER**.

Again (cf. discussion of **MANAGEMENT** above), the target language collocate TOP was not identified in the source language corpus as a collocate of **MANAGER**. A number of Anglicisms were detected among the less frequent target language lexical collocates of **MANAGER**. They were identified in the collocational span between 11 and 5 times: SCHOOL (11), SUMMER (10), TRAVEL (9), EQUITY (8), DOLLAR (6), GENERAL (6), PORTFOLIO (5), COACH (5), ASSET (5), MARKETING (5). Except for GENERAL and ASSET, which form the compounds GENERAL **MANAGER** and ASSET **MANAGER** with the node, all of the above are collocates of **MANAGER**. Although these target language collocates are all Anglicisms, only 2 of them were also identified as source language collocates, i.e. ASSET and MARKETING. This exemplifies a tendency of **MANAGER** to attract Anglicisms as collocates, but does not necessarily indicate similar collocational behavior of the node in both languages.

MARKET

SOURCE LANGUAGE				TARGET LANGUAGE			
T	LC	Top 5	TRANSLATION	Top 5	LC	T
2,441	314	≡ STOCK (185)	Aktien	✓ ≡ EMERGING (19)	6	55
		SHARE (90)	Aktie/Anteil	GLOBAL (8)		
		✓ ≡ EMERGING (83)	aufstrebend	MAKING (5)		
		GLOBAL (70)	weltweit	≡ RESEARCH (5)		
		≡ RESEARCH (56)	Forschung	≡ STOCK (5)		

The node **MARKET** features three identical lexical collocates in both corpora. The first identical collocate STOCK was identified 179 times in L1 position in the source language corpus and 5 times in the target language corpus in L1. The node **MARKET** and the collocate STOCK in L1 form the compound STOCK **MARKET** which *PONS* translates to "Aktienmarkt" or "Börse". Thus, STOCK is not an identical lexical collocate according to the guidelines of the qualitative corpus analysis.

The second identical collocate EMERGING was identified in L1 position in the target language corpus. In the source language corpus, it is also most frequently identified in L1, but also in other positions. However, the combination EMERGING **MARKET** is borrowed, but it does not qualify as a compound because no corresponding lexical entries in the chosen dictionaries exist. It may be translated with "Neuer Markt". Below, an example of EMERGING **MARKET** in both corpora is given.

SL 832: […] want to take a flyer on a bank stock or gamble on an emerging market.

TL 31: […] fünfmal stärker in den Emerging Market investiert als japanische und US-Banken.

In both corpora the plural form EMERGING MARKETS was also used. EMERGING qualifies as an identical lexical collocate of **MARKET**.

The third identical lexical collocate, RESEARCH, is found in all positions of the span except for L2 and L1 in the source language corpus. In the target language corpus, RESEARCH occurs in L3 and more frequently in R1. **MARKET** RESEARCH is defined by *Oxford Dictionaries* as "the action or activity of gathering information about consumers' needs and preferences." The term is usually translated with "Marktforschung". The latter is also identified 10 times in the target language corpus. Marktforschung and **MARKET** RESEARCH are used interchangeably, but the German term still occurs more frequently. RESEARCH does not qualify as an identical lexical collocate, because it forms part of the compound **MARKET** RESEARCH.

MARKET is among the most frequent nodes in the source language corpus (cf. Figure 23). All identified lexical target language collocates are Anglicisms. More importantly, they all were also identified as source language collocates of **MARKET**. This supports the hypothesis that nodes which are particularly frequent in the source language are likely to lead to adoption and imitation of their collocational structures in the target language.

MARKETING

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
413	34	✓ CHIEF	≙ (24)	Chef/Haupt-	VERTRIEB	(22)	13	207
		PRESIDENT	(21)	Vorstands- vorsitzender	✓ CHEF	≙ (16)		
		DIRECTOR	(14)	Leiter	✓ MANAGER	≡ (6)		
		✓ MANAGER	≡ (13)	Führungskraft	VORSTAND	(6)		
		✓ STRATEGY	≙ (9)	Strategie	✓ STRATEGIE	≙ (5)		

As illustrated in the chart above, **MARKETING** shares three collocates; one is identical and two are equivalent. At first glance, the collocates reveal that **MARKETING** is associated repeatedly with positions within a company in the source and target language corpora, such as CHIEF, PRESIDENT, MANAGER, VORSTAND.

The first equivalent lexical collocate of **MARKETING** is CHIEF and CHEF. CHIEF was used in positions L1 and R1 in the source language corpus and in L4, L3 and R1 position in the target language corpus. Below some examples from both corpora illustrate this.

SL 97: Then he went looking for a new marketing chief.

SL 303: As any chief marketing officer knows, this generation believes in […].

TL 26: […] sagt Oliver Mihm, Chef von Investors Marketing.

TL 179: […] sagt Oliver Mihm, Chef der Beratung Investors Marketing.

TL 182: Der stellvertretende Marketing-Chef Bruno Matheu gilt mit 44 als zu jung.

It should be pointed out that in concordance lines in which CHIEF co-occurred with the node in L1 position it is used as an adjective. *Oxford Dictionaries* defines the adjective CHIEF as either "most important" or "having or denoting the highest rank or authority". In those concordance lines in which CHIEF co-occurred with the node in R1 position, it is used as a noun. The noun CHIEF is defined by *Oxford Dictionaries* as "a leader or ruler of a people or clan." The target language word-form CHEF only exists as a noun. Furthermore, the English word-form CHIEF is unisex, the female form of German CHEF is CHEFIN. The latter was also identified as a collocate in the target language corpus. Consider the following example.

TL 317: […] Marketing-Chefin Kirsten Willenborg konferiert mit ihrem Team wann immer es geht im Kaminzimmer.

CHEFIN was also counted among the 16 co-occurrences of CHEF and **MARKETING** and qualifies as an equivalent lexical collocate.

The other equivalent lexical collocate of **MARKETING**, according to quantitative analysis, is STRATEGY and STRATEGIE. In the source language corpus STRATEGY occurs in L4, L2 and R1 position. In the target language corpus, STRATEGIE was identified in R1, R2 and R4 position. Below examples from both corpora are given.

SL 54: […] adapts its marketing strategy to appeal to all kinds of grocery shoppers.

SL 83: […] this strategy evinces the new marketing order.

SL 324: […] is shaking up a calcified strategy and marketing culture, including giving dealers a real voice.

TL 71: Heute haben uns die Werber ihren Vorschlag für eine Marketing-Strategie präsentiert.

TL 159: Er bearbeitete in lokalen Teams Projekte zu den Themen Marketing, Strategie und Produktion.

STRATEGY and STRATEGIE qualify as equivalent lexical collocates of **MARKETING**.

Finally, there is an identical lexical collocate for **MARKETING** which was identified. MANAGER occurs in the source language corpus twice in L4 position, but most frequently in R1. In the target language corpus, MANAGER co-occurs with **MARKETING** in R1 and R2 position. Example from both corpora show this below.

SL 200: […] Dave Fester, general manager of PC product marketing at Microsoft.

SL 235: […] Liz Friedman, a Microsoft group marketing manager who […].

TL 77: […] und dem für das zentrale Marketing zuständigen Manager Hans-Christian Schwingen.

TL 106 Dabei hat der frühere Marketing-Manager bei Blendax nicht einmal die Rezeptur selbst entwickelt.

There is no lexical entry for **MARKETING** MANAGER in the chosen mono- and bilingual dictionaries. Thus, MANAGER qualifies as an identical lexical collocate of **MARKETING**. Overall, the qualitative corpus analysis confirms similar collocational behavior of **MARKETING**.

It should be mentioned in passing that it is generally debatable whether PRESIDENT and VORSTAND qualify as equivalent collocates. But because the translation of PRESIDENT is rather "Vorstandsvorsitzender" (www.pons.eu) and VORSTAND is normally translated as "Management Board", this was disregarded. In addition, PRESIDENT always refers to a single person whereas VORSTAND can refer to a single person or a group of people which make up the management board of a company.

One additional target language collocate which is an Anglicism was identified, i.e. INVESTORS. INVESTORS does not occur as a collocate within the source language corpus.

OFFICER

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
262	21	✓ CHIEF ≡ (243)	Chef/Haupt-	✓ CHIEF ≡ (18)	7	41
		FINANCIAL (55)	finanziell	OPERATING (8)		
		✓ INVESTMENT ≡ (50)	Investition	✓ INVESTMENT ≡ (6)		
		EXECUTIVE ≡ (45)	leitender Angestellter	EXECUTIVE ≡ (5)		
		✓ TECHNOLOGY ≡ (27)	Technologie/ Technik	✓ TECHNOLOGY ≡ (5)		

The node **OFFICER** features four identical lexical collocates, as the chart above demonstrates. The collocate CHIEF was identified in the source language corpus in all positions of the span, except for L1 and R1. CHIEF collocates with **OFFICER** 237 times in L2 position. In the target language corpus, CHIEF collocates with **OFFICER** exclusively in L2 position. Examples of this are presented below.

SL 12: Still, it requires a gutsy chief marketing officer to ask the boss to invest […].

SL 16: […] who had worked at Monsanto (MON) and Allergan (AGN), as chief scientific officer.

TL 11: […] haben wir die Funktion eines global verantwortlichen Chief Compliance Officer geschaffen, der direkt an mich berichtet.

TL 19: […] ist Managing Director und Group Chief Investment Officer beim Staatsfonds […].

The combination CHIEF ____ **OFFICER** accounts in both corpora for all instantiations in which CHIEF collocates with **OFFICER** in L2. The blank space can be filled with various word-forms. All other Top 5 identical lexical collocates of **OFFICER** fulfill this function in the source and target language corpora. Consider the examples below.

SL 13: […] when a Boeing chief financial officer and a Pentagon official were found guilty of […].

TL 2: […] war bis Januar 2008 als Chief Operating Officer die Nummer zwei bei der Geschäftskundensparte.

SL 24: […] and global chief investment officer of equities at investment management firm BlackRock.

TL 12: Horacio Valeiras, Chief Investment Officer des Vermögensverwalters Nicholas Applegate […].

SL 9: Two years ago, Discovery got a new chief executive officer in the person of David Zaslav […].

TL 7: […] sieht sie ganz von selbst wie der ideale Chief Executive Officer der Vereinigten Staaten aus.

SL 49: Chris Bruzzo, Starbucks' chief technology officer, who oversees MyStarbucksIdea, adds […].

TL 6: […] wurde auf ein Inserat des Weltwirtschaftsforums aufmerksam, das einen Chief Technology Officer suchte.

The combination of CHIEF and **OFFICER** is used in both corpora to describe positions of high-level managers. CHIEF qualifies as an identical lexical collocate of **OFFICER**.

As far as the other three identical lexical collocates of **OFFICER** are concerned, it has to be said that there exists a dictionary entry for (CHIEF) EXECUTIVE **OFFICER**. *Merriam Webster* defines CEO as follows "the executive with the chief decision-making authority in an organization or business." The abbreviation CEO is used 54 times in the target language corpus and 1,056 times in the source language corpus. CHIEF EXECUTIVE **OFFICER** is, according to *PONS*, translated with "Hauptgeschäftsführer/in". EXECUTIVE is not counted as an identical lexical collocate of **OFFICER**.

For (CHIEF) INVESTMENT **OFFICER** or (CHIEF) TECHNOLOGY **OFFICER** no lexical entries were found. Thus, the identical lexical collocates INVESTMENT and

TECHNOLOGY affirm similar collocational behavior and qualify as identical lexical collocates of **OFFICER**. It should be noted that **OFFICER** collocates exclusively with additional Anglicisms in the target language corpus. Compared to other nodes under investigation **OFFICER** is not used very frequently. To arrive at a conclusion, there is a tendency for low-frequency nodes to co-occur with Anglicisms in the target language corpus.

The node **OFFICER** features two more lexical collocates in addition to the Top 5 target language collocates which are displayed in the chart above. They are also Anglicisms - RISK and COMPLIANCE. None of the two is identified as a source language corpus collocate.

PORTFOLIO

SOURCE LANGUAGE				TARGET LANGUAGE			
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**	
323	35	≡ MANAGER (36)	Führungskraft	MANAGER ≡ (5)	4	101	
		BILLION (19)	Milliarde	✓ AKTIEN ≙ (5)			
		INVESTMENT (17)	Investition	✓ FONDS ≙ (5)			
		✓ STOCKS ≙ (16)	Aktien	PROZENT (5)			
		✓ FUND/FUNDS ≙ (14)	Fonds	–			

As represented in the chart above, the node **PORTFOLIO** shares three lexical collocates; one identical and two equivalent. The identical lexical collocate MANAGER occurs in L3, R4, but most frequently in R1 position of the span of **PORTFOLIO** in the source language corpus. In the target language corpus, MANAGER was identified as a collocate only in R1 position. **PORTFOLIO** MANAGER is thus likely to form a compound and does have an entry in *PONS* with the translation "Portfolio-Manager(in)". Interestingly, **PORTFOLIO** does translate to 'Aktienbestand' or 'Wertpapierbestand', but this is not considered in the translation of **PORTFOLIO** MANAGER. Instead it is opted for the identical adoption of the compound. The German translations 'Wertpapierbestand' and 'Aktienbestand' do not occur in the target language corpus. Again, it seems like an Anglicism is preferred over a target language expression for stylistic reasons. Examples of the compound **PORTFOLIO** MANAGER in both corpora are given below. MANAGER does not qualify as an identical lexical collocate of **PORTFOLIO**.

SL 162: But last month, the portfolio manager at the Evergreen Strategic Growth (ESGIX) fund cut back on his fund's holdings in HP.

TL 9: Das Lernziel: Portfolio-Manager sollen frühzeitig erahnen, wohin sich Märkte entwickeln.

As far as the equivalent lexical collocate AKTIEN and STOCKS is concerned, they co-occur with **PORTFOLIO** in various positions in both corpora. This is exemplified below.

SL 58: […] start building an income portfolio with blue-chip stocks, many of which yield 3.1% or more.

SL 67: […] because the holdings of the fund may not mirror the stocks in an investor's portfolio.

SL 85: […] founded as a railroad holding company in 1929, also owns a portfolio of stocks and bonds.

SL 145: […] 60% of the portfolio is in stocks—while putting a sliver into riskier ventures like […].

SL 262: What does the possibility of a stable-to-stronger dollar mean for the international stocks in your portfolio?

TL 1: […] Fondsmanager zusätzliche Mittel frei, die er in sein Aktien-Portfolio steckt.

TL 98: […] und sein ganzes Portfolio mit günstigen Aktien zu bestücken ist auch nicht jedermanns Sache.

TL 101: Er würde stattdessen auf die im Portfolio enthaltenen Aktien Optionen verkaufen […].

AKTIEN and STOCKS qualify as equivalent lexical collocates of **PORTFOLIO**. In both corpora, STOCK is also among the collocates of **PORTFOLIO** which did not make the Top 5.

The second equivalent lexical collocate FUND/FUNDS and FONDS is identified in various positions left and right of the node in both corpora. For the source language corpus the singular and plural form were counted, because FONDS can refer to a single FUND or a number of FUNDS and thus both translate to FONDS. The corpus examples below will also show that FUND and FUNDS are used similarly.

SL 107: […] now isn't easy, regardless of the marquee names in his fund portfolio.

SL 191: […] my well-balanced portfolio of mutual funds wasn't doing so hot […].

SL 216: It has also halved the number of funds in its portfolio.

SL 314: So her model portfolio uses active mutual fund managers, often with […].

TL 32: […] und im Vergleich zu anderen europäischen Fonds hat das Portfolio Boden gutgemacht.

TL 37: […] mit einer Milliarde Euro im Portfolio mischt der Fonds von Union Investment.

TL 63 […] das Vermögen eines kleinen oder erfolglosen Fonds auf ein anderes Portfolio zu übertragen.

All in all, FUND/FUNDS and FONDS qualify as equivalent lexical collocates of **PORTFOLIO**.

PROVIDER

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5	TRANSLATION	Top 5	LC	T
73	12	SERVICE ≡ (12)	Dienst(leistung)	INTERNET ≡ (7)	2	28
		CABLE (7)	Kabel	SERVICE ≡ (5)		
		LARGEST (7)	größte	–		
		WIRELESS (6)	drahtlos	–		
		INTERNET ≡ (6)	Internet	–		

The node **PROVIDER** features two identical lexical collocates. Clearly, **PROVIDER** is of comparatively low-frequency in the target language corpus. Nodes of this kind tend to collocate in the target language corpus exclusively with Anglicisms. **PROVIDER** is typically modified by another noun, resulting in the target and source language corpora in combinations like INTERNET **PROVIDER** and SERVICE **PROVIDER**. Both of these quantitatively identified collocates are borrowed as constituents of a compound and thus not decisive for assigning similar collocational behavior. *Oxford Dictionaries* defines SERVICE **PROVIDER** as "a company that provides its subscribers access to the Internet." And it is translated to "Internet-(Dienst)anbieter" (www.pons.eu) just like INTERNET **PROVIDER**. Below, one example per collocate and corpus is given to demonstrate its use.

SL 13: […] with revenues split between the carrier and a service provider such as TeleNav or Networks in Motion.

TL 10: Um einen Anschlussinhaber zu ermitteln, verlangt der Internet-Provider bis zu 30 Euro.

SL 41: […] a partnership with Deutsche Telekom, Germany's biggest Internet provider […].

TL 12: Doch die Service-Provider wie T-Home oder 1&1 geben die dazugehörigen Namen und Adressen […].

In both corpora the compound INTERNET SERVICE **PROVIDER** was also identified; occasionally it is used as an alternative to SERVICE **PROVIDER** or INTERNET **PROVIDER**, but means the same. For exemplification, see below.

SL 44: […] or another Internet service provider to find out who was using that IP address […].

TL 17: Die Internet-Service-Provider sollen jetzt die Hilfssheriffs spielen.

PUBLIC

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5 ≡	TRANSLATION	Top 5 ≡	LC	T
544	89	PRIVATE (34)	privat	PRIVATE (16)	3	26
		OFFERING (25)	Spende/Angebot	PARTNERSHIP (10)		
		COMPANY (24)	Unternehmen	PARTNERSHIPS (5)		
		POLICY (21)	Programm/Grundsatz	–		
		RELATIONS (19)	Beziehungen	–		

From the chart above, it can be seen that **PUBLIC** shares one identical lexical collocate and attracts Anglicisms as collocates in the target language corpus. **PUBLIC** collocates in most instantiations with its opposite PRIVATE in both corpora. Sinclair states that contrasts have a strong chance of being mentioned together (cf. 1991: 110). Contrasting collocational behavior in two languages, this information is important. The Anglicism PRIVATE in this case is chosen over the German equivalent ‘privat’. Theoretically, PRIVATE could also be a German inflectional variant of the adjective ‘privat’ (cf. analysis of **INVESTOR** above). It was verified for all instantiations in which PRIVATE co-ocurred

with **PUBLIC** that the English word-form was used. In practice this means that all 16 co-occurrences of **PUBLIC** and PRIVATE were manually retrieved and it was determined whether the English or German word-form was used.

In the source language corpus, PRIVATE collocates with **PUBLIC** in L4, L3, L2, R1, R2 and R4 position. In the target language corpus only in R1 position. Every time **PUBLIC** is followed by PRIVATE in the target language corpus, it is also followed by PARTNERSHIP in R2. PARTNERSHIP and PARTNERSHIPS only occur in R2 position without exception. Consider the examples below.

TL 14: Das können etwa komplizierte Infrastrukturprojekte wie ein Flughafenausbau im Rahmen einer Public Private Partnership sein.

TL 16: Bei Public Private Partnership, das die indische Regierung derzeit gezielt fördert […].

The expression **PUBLIC** PRIVATE PARTNERSHIP is translated by *PONS* as "öffentlich-private Partnerschaft". However, the chosen English monolingual dictionaries do not mention the expression. Thus, the term is suspicious of being a pseudo-Anglicism (cf. subsection 5.1.2). In addition, and despite its high frequency in the source language corpus, PARTNERSHIP or PARTNERSHIPS did not co-occur with **PUBLIC**. The quantitatively identified lexical collocate PRIVATE is not an identical collocate of the node **PUBLIC**, because PRIVATE forms part of a fixed expression which is unknown in the source language, thus qualifies as a pseudo-Anglicism.

RATING

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	**LC**	**Top 5**			**Top 5**		**LC**	**T**
132	10	CREDIT	(41)	Kredit	AGENTUREN	≙ (14)	5	115
		AGENCY	≙ (18)	Agentur	TOP	(10)		
		AGENCIES	≙ (15)	Agenturen	AGENTUR	≙ (5)		
		CUT	(5)	Kürzung	BANK	(5)		
		STOCK	(5)	Aktien	ANLEIHEN	(5)		

The node **RATING** shares two equivalent collocates. All instantiations of AGENCY/IES and AGENTUR/EN were identified in R1 position of the collocational span. For examples from both corpora see below.

SL 107: Until late last year, the rating agency had adjusted its models only to account for […].

TL 67: In eine andere Richtung geht die Forderung nach einer europäischen Rating-Agentur […].

SL 117: […] and a longtime critic of the rating agencies.

TL 101: Bisher sind die Rating-Agenturen rein angelsächsisch geprägt.

RATING AGENCY is listed in *PONS* as a compound and translated as "Ratingagentur". The frequent use of compounding and their similar structures in the source and target language have led to the formation of hybrid compounds like RATING-AGENTUREN in this case. Unfortunately, the quantitatively retrieved collocates of **RATING** are not significant for affirming or refuting the initial hypotheses because they are not collocates, but form compounds with the node.

RESEARCH

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
848	101	✓ FIRM	≙ (119)	Firma/Kanzlei/ Unternehmen	CENTER	(6)	11	96
		DEVELOPMENT	(48)	Entwicklung	INSTITUTE	(6)		
		ANALYST	(45)	Börsenfachmann/ Analytiker	ECONOMIC	(5)		
		MARKET	(40)	Markt	BUREAU	(5)		
		INVESTMENT	(38)	Investition	✓ FIRMA	≙ (5)		

The chart above illustrates that **RESEARCH** has one equivalent lexical collocate in the target and source language corpus. Only instantiations in which FIRM is used as a noun are counted (cf. **ANALYST** and **MANAGEMENT** above). Although the collocate FIRM was in many concordance lines situated in R1, it was also identified in R3 and R4 position. In the target language corpus, it was identified in L3 and R1 positions of the span. The node **RESEARCH** collocates almost exclusively with Anglicisms in the target language corpus. The only exception is FIRMA. The counterpart of the only German collocate of **RESEARCH** in the target language, is at the same time the most frequent collocate in the source language corpus. Examples from both corpora are presented below.

SL 259: Tom Gallagher, the head of Washington research for financial services firm ISI Group […].

SL 412: New research from executive search firm Egon Zehnder International finds that fewer […].

SL 468: […] macro strategist at CreditSights, a capital-structure research firm in New York.

SL 607: Platt, director of the Platt Retail institute, a retail research and consulting firm.

TL 48: Bereits 2001, bei seiner ersten Firma für Management und Research, deren Börsenwert […].

TL 60: Sie stammt von der unabhängigen Research-Firma Credit Derivatives Research (CDR).

Although FIRM is used frequently in R1 position in both corpora, a lexical entry in the chosen mono- and bilingual dictionaries does not exist for **RESEARCH** FIRM. Thus it can be concluded that source language collocational behavior of **RESEARCH** is imitated in the target language in this case. FIRM and FIRMA qualify as equivalent lexical collocates.

The remaining four Top 5 lexical target language collocates of the node **RESEARCH** are all Anglicisms. Remarkably, they were all also identified as significant lexical collocates in the source language corpus. In the source language corpus, CENTER co-occurred with **RESEARCH** 18 times, INSTITUTE 25 times, ECONOMIC 14 times, and BUREAU 11 times.

SERVICE

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5	TRANSLATION	Top 5	LC	T
710	65	✓ CUSTOMER (106) ≙	Kunde	DESIGN (16)	23	268
		NEW (24)	neu	✓ KUNDEN (10) ≙		
		✓ COMPANIES (20) ≙	Unternehmen	BIETEN (7)		
		✓ INTERNET (19) ≡	Internet	✓ UNTERNEHMEN (7) ≙		
		FREE (18)	kostenlos/frei	✓ INTERNET (5) ≡		

The node **SERVICE** shares two equivalent and one identical lexical collocate in the source and target language corpora. The equivalent collocate CUSTOMER was mainly encountered in L1 position, but also in L3, L2 and R4 in the source language corpus. See below for a few examples.

SL 179: For example, a customer seeking service too often faces a predictable progression from […].

SL 187: Overall, customer service has dipped, perhaps a victim of spending cutbacks.

SL 517: "We offer every customer the same service," he says.

SL 535: […] by offering aisles of service to every potential customer.

The collocate KUNDEN was identified in L3, L2, R2 and R4 position. Below this is exemplified.

TL 104: Bei der Ergo Versicherungsgruppe legen wir großen Wert auf Service und beraten unsere Kunden umfassend […].

TL 123: […] internes Chaos und für Kunden schmerzliche Service-Engpässe beim Aufbau neuer Unternehmensabläufe.

TL 135: "Die Kunden schätzen wieder Service, Qualität und Bequemlichkeit", sagt Breuer.

TL 161: Die drastischen Kostensenkungsmaßnahmen werden umfassende Auswirkungen auf den Service haben, Kunden werden unzufrieden und gehen.

CUSTOMER and KUNDEN qualifies as an equivalent lexical collocate of **SERVICE**.

The second equivalent lexical collocate of **SERVICE** is COMPANIES and UNTERNEHMEN. Again, it was verified that only instantiations are considered for the analysis in which the plural meaning of UNTERNEHMEN is used (cf. **EQUITY**, **IMAGE**, **MANAGEMENT** above). COMPANIES co-occurs with **SERVICE** in L4, L3, L2 R1, R3 and R4 position in the source language corpus. For a number of examples from the source language corpus see below.

SL 143: To create meaningful local service, companies must nurture the empathetic connections […].

SL 144: In connection with the report's release, service companies voluntarily agreed to stop using […].

SL 171: Just ask those who responded to our second annual ranking of the best companies for customer service, which uses data from J.D. Power […].

SL 253: The company provides warranty support and customer service for U.S. companies including […].

SL 351: […] a set of 10 specifications for customer service phone systems that companies could adopt to promote their […].

SL 543: […] 10 new names made our list this year, including companies well-known for service, like L.L. Bean and Amazon.com (AMZN).

SL 647: These instances exemplify the most common ways companies get service seriously wrong, and the surprisingly easy ways they can […].

UNTERNEHMEN was identified in L4, L3 and R1 position in the target language corpus. Corresponding examples are presented below.

TL 41: […] wo zu DDR-Zeiten der Rechnerriese Robotron saß, wickeln die Unternehmen Wartung und Service ab.

TL 113: […] Mindestmaß an Service, Unternehmen kämpfen daher mit ausgefeilten Methoden um den Nachwuchs.

TL 226: Gute Umgangsformen geben inzwischen den Ausschlag bei Bewerbungen, sie heben Unternehmen im Vertrieb oder Service positiv hervor […].

According to the standards of the quantitative and qualitative analysis, COMPANIES and UNTERNEHMEN are equivalent lexical collocates of **SERVICE**.

The node **SERVICE** shares one identical lexical collocate, i.e. INTERNET. In the source language corpus, INTERNET was identified in L4, L3, L2, L1 and R4 position. In the target language corpus, INTERNET co-occurred with **SERVICE** in L1, R2 and R4 position. The examples from both corpora below illustrate this.

SL 321: […] TV, Internet, and phone service over Comcast's 125,000 miles of pipe.

SL 324: Internet service to some homes may be as high as $416 a month.

SL 383: […] the PayPal online payment operation and the Internet calling service Skype.

SL 394: Internet companies with a service up and running and millions of users should fare […].

SL 406: […] bring Sprint new subscribers from deals to bundle wireless service with TV and Internet access.

TL 126: […] eigene Abteilungen zu diesem Zweck aufgebaut und werben mit dem Service im Internet.

TL 144: Die Internet-Service-Provider sollen jetzt die Hilfssheriffs spielen.

TL 153: Damit dieser Service funktioniert, unterhalten die Internet-Banken beträchtliche IT-Abteilungen.

INTERNET is used in different positions of the span, illustrating similar collocational behavior in the source and target language corpora. INTERNET qualifies as an identical lexical collocate of **SERVICE**. It should be mentioned that although **SERVICE** is highly frequent in German it does not merely collocate with target language expressions.

The target language collocate DESIGN was not identified as a source language collocate. Other target language collocates which are Anglicisms and which co-occur 5 times with **SERVICE** are CHAMPION, PRICING and PROVIDER. Out of these three only PROVIDER qualified also a significant lexical collocate in the source language corpus.

SOFTWARE

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
625	116	NEW	(21)	neu	✓ HERSTELLER	≙ (17)	31	382
		✓ COMPANIES	≙ (17)	Unternehmen	✓ UNTERNEHMEN	≙ (14)		
		✓ COMPUTER	≡ (17)	Computer/Rechner	✓ COMPUTER	≡ (7)		
		✓ MAKER	≙ (17)	Hersteller	KUNDEN	(7)		
		BUSINESS	≡ (15)	Geschäftsbetrieb	BUSINESS	≡ (6)		

The node **SOFTWARE** shares four lexical collocates; half of them are identical and the other half is equivalent. When looking at the Top 5 collocates within the target language corpus, it shows that **SOFTWARE** collocates with Anglicisms and target language wordforms alike.

The *WordSmith Tools* analyses show that COMPANIES and UNTERNEHMEN (only the instantiations with plural meaning were considered, cf. above **EQUITY**, **IMAGE**, **MANAGEMENT**, **SERVICE**) take varied positions left and right of the node within the

target and source language corpora. Below a number of source and target language corpus examples are provided to illustrate this.

SL 51: Gaming companies are using the software to more swiftly and cheaply distribute their wares.

SL 117: We like software and IT service companies with businesses that are […].

SL 226: […] a high-end purveyor of software that helps companies analyze their markets.

SL 328: Dell will give other companies the software to help establish the standard and will make […].

SL 412: Even in a slowing market, though, M&A can be extremely profitable for software companies.

SL 488: Open-source software or companies like InnoCentive, which encourages outside […].

TL 12: […] wichtige Einnahmequelle geworden, anderen Unternehmen die eigene Software zur Verfügung zu stellen.

TL 47: Immer mehr Unternehmen nutzen Gebraucht-Software - zum Ärger der Softwarehersteller, die […].

TL 56: Und bei den indischen Software-Unternehmen herrscht vollends verkehrte Welt […].

TL 132: […] die ursprünglich für 2009 geplanten Großeinkäufe von Hard- und Software schieben viele Unternehmen erst einmal vor sich her.

TL 243: Derzeit hofft die Branche auf Investitionen der Unternehmen in Software und IT-Services.

TL 344: […] können Unternehmen die Software nicht auf eventuelle Sicherheitslücken hin überprüfen.

TL 375: […] vor allem um Unternehmen den Weiterverkauf ihrer Software zu erschweren.

UNTERNEHMEN and COMPANIES qualify as equivalent lexical collocates of **SOFTWARE**.

The second equivalent lexical collocate of **SOFTWARE** which was identified during the quantitative corpus analysis is MAKER/HERSTELLER. Much like with UNTERNEHMEN, the singular and plural form of HERSTELLER is identical. In accordance with MAKER, which is the singular form, only instantiations in which HERSTELLER occurs with singular meaning as well were counted for the quantitative analysis and taken into consideration for the qualitative analysis. In both corpora MAKER and UNTERNEHMEN were identified in L4, L3, L2 and R1 position. Consider the following examples:

SL 326: […] the leading maker of operating system software for advanced mobile phones.

SL 401: […] the top maker of virtualization software.

SL 564: SAP (SAP), a maker of software for corporations.

SL 603: […] won't preclude moves by the software maker to bypass the board entirely.

TL 164: […] in den Neunzigerjahren ein vielversprechender Hersteller von Hard- und Software für Datenübertragung sowie Grafikkarten.

TL 211: Nicht minder rigide versucht der Hersteller von Büro-Software, Datenbank-Programmen und […].

TL 205: Statt von einem Consultinghaus wird der Betrieb nun vom Hersteller der Software direkt betreut.

TL 202: […] SAP - den zukünftigen größten Software-Hersteller in Europa.

MAKER and HERSTELLER qualify as equivalent lexical collocates of **SOFTWARE** according to the quantitative and qualitative corpus analysis.

The first identical lexical collocate of the node **SOFTWARE** is COMPUTER. In the source language corpus, COMPUTER collocates with **SOFTWARE** in all positions of the collocational span except for R1 and R4. Examples are provided below.

SL 110: […] can use them to hire software programmers or computer scientists with rare skills.

SL 135: And it was a chance to test their world-beating computer systems and software in some of the most demanding business settings.

SL 237: […] is also creating computer software and devices to help consumers monitor their […].

SL 481: […] cost savings that made the subcontinent a leader in software and computer services.

SL 489: Those industries include computer and communications hardware, software and […].

SL 525: Yes, Andersen had installed on her computer a software program, KaZaA, for sharing music […].

In the target language corpus, COMPUTER collocates with **SOFTWARE** in all positions except for L4 and L1. For according examples from the target language corpus see below.

TL 48: […] weil sie sich die notwendigen Computer oder die Software nicht leisten können.

TL 67: […] hatte den Konzern über drei Jahrzehnte mit seiner Formel "Windows-Software auf alle Computer" zu einer der wertvollsten […].

TL 152: Solarwerte dominieren längst den Technologie-Index TecDax, haben Software, Computer und Halbleiter buchstäblich in den Schatten gestellt.

TL 214: […] der sich auch nach seiner Zeit bei SAP auf Software- und Computer-nahe Projekte konzentriert.

TL 327: Nicht minder rigide versuchen Hersteller von Büro-Software, Datenbank-Programmen oder Computer-Betriebssystemen […].

TL 338: […] fast alles drehte sich um Computer und Software.

The collocate COMPUTER is an identical lexical collocate of the node **SOFTWARE**.

The other identical lexical collocate of **SOFTWARE** is BUSINESS. In the source language corpus, BUSINESS occurs in all positions of the span except for R2. The following examples confirm this.

SL 97: […] he co-founded a company, BitTorrent, to build a business around his software.

SL 138: Some say the problem is the business model, not the software.

SL 249: […] then send the contents to small-business accounting software such as QuickBooks […].

SL 276: Microsoft's Kevin Johnson talks about where the software king's online business is headed now.

SL 313: […] for selling consumer and business software, it represents a threat to Cisco's hegemony.

SL 416: […] markets for cloud search and software tools - a natural business for Google and its competitors.

SL 552: […] adding $1 billion each month to the cash hoard from its lucrative software business, it faces a serious long-term threat.

In the target language corpus, BUSINESS was identified in L3, L1, R1 and R3 position. Below corresponding corpus examples are given.

TL 42: Der Computer - eine sogenannte Appliance - mit der Business-By-Design-Software steht beim Kunden […].

TL 65: […] im vergangenen Herbst präsentierten Software Business By Design.

TL 148: Die Software ist das erste Komplettpaket für Business-Software, das zur Miete über das Internet angeboten werden soll.

TL 330: […] schließlich haben wir bereits die Software-Lösungen SAP Business One, SAP Business All-in-One und […].

It can be seen from the examples above that BUSINESS and **SOFTWARE** co-occur in the target language corpus in many cases as constituents of software or company names, i.e. *Business-By-Design-Software*. This holds true for 4 out of the 6 instantiations in which the node and BUSINESS co-occurred. Thus, BUSINESS is not counted as an identical lexical collocate of **SOFTWARE**. Nevertheless, with two equivalent and one identical lexical collocate, the node displays similar collocational behavior in the source and target language corpora.

Also note that further target language collocates which are Anglicisms and which co-occur at least 5 times with the node **SOFTWARE** are DESIGN (5), MAIL (5), OPEN-SOURCE (5), HARDWARE (6), WEB (6) and INTERNET (16). All of the above, except for MAIL, are also significant lexical collocates of **SOFTWARE** in the source language corpus: DESIGN (8), OPEN-SOURCE (12), HARDWARE (10), WEB (8), INTERNET (5).

SUPPLY

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
222	29	CHAIN	≡ (32)	(Liefer)kette	CHAIN	≡ (8)	3	42
		DEMAND	(28)	Nachfrage	MANAGEMENT	≡ (7)		
		✓ GLOBAL	≙ (15)	weltweit	✓ WELTWEITEN	≙ (5)		
		MANAGEMENT	≡ (14)	Betriebsleitung	–			
		SHORT	(7)	kurz	–			

The node **SUPPLY** features two identical and one equivalent collocate. The most frequent source and target language collocate CHAIN was only identified in R1 position forming the compound **SUPPLY** CHAIN in both corpora. See below for one example per corpus.

SL 2: With a supply chain like that, even labeling a line of packaged noodles is a chore.

TL 9: […] glaubt Guido Boehnke, Supply-Chain-Spezialist der Unternehmensberatung Celerant.

The compound **SUPPLY** CHAIN is defined by *Oxford Dictionaries* as "the sequence of processes involved in the production and distribution of a commodity." Consequently, CHAIN does not qualify as an identical lexical collocate of **SUPPLY**.

The other identical lexical collocate MANAGEMENT was also most frequently identified in R1 position and occasionally in R2 in the source and target language corpus. No lexical entry for SUPPLY MANAGEMENT was found in the chosen mono- or bilingual dictionaries. The examples below illustrate the use of MANAGEMENT in R1 in both corpora.

SL 96: On Oct. 1 in the U.S., the Institute of Supply Management announced a plunge in its key manufacturing index

SL 109: […] says Michael Lewis, a supply-management professor at University of Bath's School of Management.

TL 3: Seine Schwerpunkte sind Supply Management, insbesondere Global Sourcing und E-Procurement.

TL 8: Der viel beachtete Einkaufsmanagerindex für das verarbeitende Gewerbe, den das Institute of Supply Management (ISM) monatlich ermittelt […].

In 9 out of 10 co-occurrences of **SUPPLY** and MANAGEMENT in R1 in the source language corpus, node and collocate formed part of the proper name *Institute of Supply Management.* The only exception is given above in SL 109. In the target language corpus, **SUPPLY** and MANAGEMENT in R1 co-occurred 4 times; twice as part of the proper name *Institute of Supply Management.* The following examples illustrate the co-occurrence of **SUPPLY** and MANAGEMENT in R2 position in both corpora.

SL 97: Nortel Networks (NT) has 500 workers at a supply-chain management center in Monkstown.

SL 126 […] have taken over key positions including purchasing, supply-chain management, and human resources.

TL 5: […] ist Chef des Supply Chain Management von VF, das ein System von Lieferketten betreut […].

TL 17: Supply Chain Management (SCM): Steuerung der gesamten Wertschöpfungskette vom Lieferanten bis zum Kunden.

Thanks to the definition given in TL 17, an additional dictionary definition of **SUPPLY**-CHAIN MANAGEMENT is not necessary. In the source language corpus, **SUPPLY** and MANAGEMENT in R2 co-occur 4 times. In the target language corpus this is the case twice; both instantiations are given in the examples above. MANAGEMENT does not qualify as an identical lexical collocate of **SUPPLY** for two reasons. Firstly, the node and/or collocate are used as constituents of proper names frequently. Secondly, the remaining instantiations in which MANAGEMENT co-occurs with **SUPPLY** are not enough, i.e. below five, to determine collocation status according to the guidelines of the analysis. In addition, it could be argued that **SUPPLY** CHAIN MANAGEMENT qualifies as a proper name.

The equivalent lexical collocate GLOBAL/WELTWEITEN turns up in different positions of the collocational span. Some corpus examples of this are provided below.

SL 89: […] the emerging market economies are providing the biggest boost to both global supply and global demand since the Industrial Revolution.

SL 137: […] will begin to pass global supply and push prices near or even above $100 a barrel, say analysts.

SL 151: But unlike the Great Depression, there's a problem on the supply side of the global economy.

SL 190: The global money supply was growing at its fastest pace since […].

SL 216: We expect global coking-coal supply to trail underlying demand in […].

TL 2 […] im Sinne eines weltweiten Supply Managements als eine strategische Herausforderung zu verstehen.

TL 6: […] verantwortlich für Supply Chain Excellence im weltweiten Logistik-netzwerk bei Bosch.

TL 14: […] Praktikum im weltweiten Supply Chain Management.

The adjective WELTWEIT is identified in the target language corpus only in its inflected form WELTWEITEN. In the analysis chart the most frequent is listed. GLOBAL and WELTWEIT qualify as equivalent lexical collocates of **SUPPLY**.

VALUE

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
662	101	MARKET	(37)	Markt/Handel	FAIR (11)	6	71
		BILLION	(27)	Milliarde	MANAGEMENT (8)		
		ASSET ≡	(22)	Vermögenswerte	PRINZIP (5)		
		DOLLAR	(21)	Dollar	RISK (5)		
		FUND	(19)	Fonds	ASSET ≡ (5)		

Among the Top 5, the node **VALUE** features one identical collocate - ASSET. This is interesting, because **VALUE** collocates also with several other Anglicisms in the target language. Unfortunately, ASSET co-occurs with **VALUE** only in L1 position and as part of the compound 'Net ASSET **VALUE**'. Thus, it does not qualify as an identical lexical collocate. Consider the corpus examples below.

TL 18: […] werden derzeit mit einem Abschlag auf den Nettovermögenswert (Net Asset Value, NAV) ihrer Reserven gehandelt.

TL 29: Hohe Abschläge auf den Nettovermögenwert (Net Asset Value, NAV) von mehr als 30 Prozent boten oft gute Einstiegsgelegenheit […].

Oxford Dictionaries defines 'Net ASSET **VALUE**' as "the value of a mutual fund that is reached by deducting the fund's liabilities from the market value of all of its shares and

then dividing by the number of issued shares." It is translated by *PONS* with "Nettoinventarwert" or "Liquidationswert".

The top target language collocate of **VALUE** is an Anglicism, namely FAIR. FAIR is at the same time the only other target language collocate (Anglicism) which is also identified as a source language collocate. Consider the following examples.

SL 293: Its sector analysts calculate fair value for every stock they cover based on fundamentals […].

SL 522: Fair value is more of an art than a science […].

TL 10: […] und da ist Fair Value das grundsätzlich richtige Prinzip.

TL 11: Während einer Krise soll auf Fair Value verzichtet werden, wenn es gut läuft aber nicht?

FAIR only occurs in L1 position in the target language corpus and is only identified as part of the compound FAIR **VALUE**. *PONS* translates FAIR **VALUE** as "Marktwert" and *Merriam Webster* defines FAIR **VALUE** as "the estimated value of all assets and liabilities of an acquired company used to consolidate the financial statements of both companies." The other identified target language collocates which are Anglicisms, but which do not occur in the source language corpus, evidence that VALUE has developed distinctive collocations with source language word-forms (cf. subsection 5.1.2).

WORK

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	**LC**	**Top 5**			**Top 5**		**LC**	**T**
1,118	148	PEOPLE	(33)	Leute/Menschen	LIFE	≡ (6)	5	62
		MAKE	(29)	machen	BALANCE	(6)		
		HARD	(24)	hart/schwierig	EASY	(5)		
		LIFE	≡ (21)	Leben	DIVERSITY	(5)		
		TIME	(21)	Zeit	FOUNDATION	(5)		

The node **WORK** features one identical lexical collocate according to the quantitative corpus analysis. The word-form **WORK** can be a noun or a verb. The corpus analysis focuses on nouns as nodes (cf. section 5.2). Therefore only such instantiations in which **WORK** is

used as a noun were taken into consideration for the quantitative and qualitative corpus analysis.

The identical lexical collocate LIFE occurs in the source language corpus in L3, L2, R1, R2 and R4 position. In the target language corpus, LIFE co-occurs with **WORK** only in R1 position. More precisely, node and collocate form part of the expression '**WORK**-LIFE-balance' in each instantiation in the corpus. Consider the examples below.

TL 11: Er sucht nach einem Job, bei dem die "Work-Life-Balance" stimmt […].

TL 16: Andere Firmen erlauben eine ausgewogene Work-Life-Balance.

The term '**WORK**-LIFE-balance' is also identified in the source language corpus 12 times and is translated by *PONS* as "Vereinbarkeit von Beruf und Familie". For examples see below.

SL 564: I have achieved a great work-life balance through strategic planning.

SL 979: […] especially for a generation for whom work-life balance is a priority.

The frequent co-occurrence of **WORK** and LIFE in the source and target language corpus thus does not evidence similar collocational behavior. It illustrates rather the adoption of an entire expression which is not flexible to the degree collocations are.

A few additional comments on the node **WORK** are in order. Firstly, all Top 5 target language collocates are Anglicisms. Out of them only BALANCE (as part of work-life balance) was also identified as a collocate in the source language corpus. Secondly, **WORK** shows a tendency to collocate preferably with grammatical collocates (different from most other nodes). Consequently, despite a generally high number of instantiations of **WORK** in the source language corpus, the frequency of its Top 5 collocates is comparatively low. In the source language corpus for example, the top 18 collocates are all grammatical. This is not something that is copied in the target language corpus. One of the Top 5 source language collocates of **WORK** is HARD, whereas EASY is a Top 5 collocate in the target language corpus. This seems to bear the potential for meaningful conclusions at first. However, EASY only collocates with **WORK** in the target language corpus in the following expression:

TL 8: […] EeePC genannten Rechner - das Kürzel steht für "Easy to Learn, Easy to Work, Easy to Play".

EeePCs were often covered in WirtschaftsWoche articles of 2008, thus EASY turns up erroneously as a frequent identical lexical collocate.

6.2.3 Summary of results and evaluation

Two hypotheses were formulated to guide the present study. The first is that the use of Anglicisms produces replications of source language collocational structures within the target language, i.e. that Anglicisms feature identical or equivalent lexical collocates in the source and target language corpora. Secondly, it was hypothesized that this is true for general and specialized journalistic texts. Thus, the purpose of the quantitative and qualitative corpus analysis of the business magazine corpora was to determine to which extent identical and equivalent lexical collocates are used with the 36 selected nodes.

The findings of the quantitative corpus analysis (cf. subsection 6.2.1) indicated that similarities considerably outnumber the differences of the collocational behavior of Anglicisms in American and German business magazines. According to the quantitative corpus analysis of the business magazines, 34 out of 36 nodes shared identical and/or equivalent lexical collocates. The qualitative interpretation of the quantitative results has shown that actually only 20 nodes out of the 36 nodes share identical and/or equivalent lexical collocates. A summary of the results of the qualitative analysis of the Top 5 collocates is given in Figure 28 below. Figure 28 is similarly constructed as Figure 25 (cf. subsection 6.2.1).

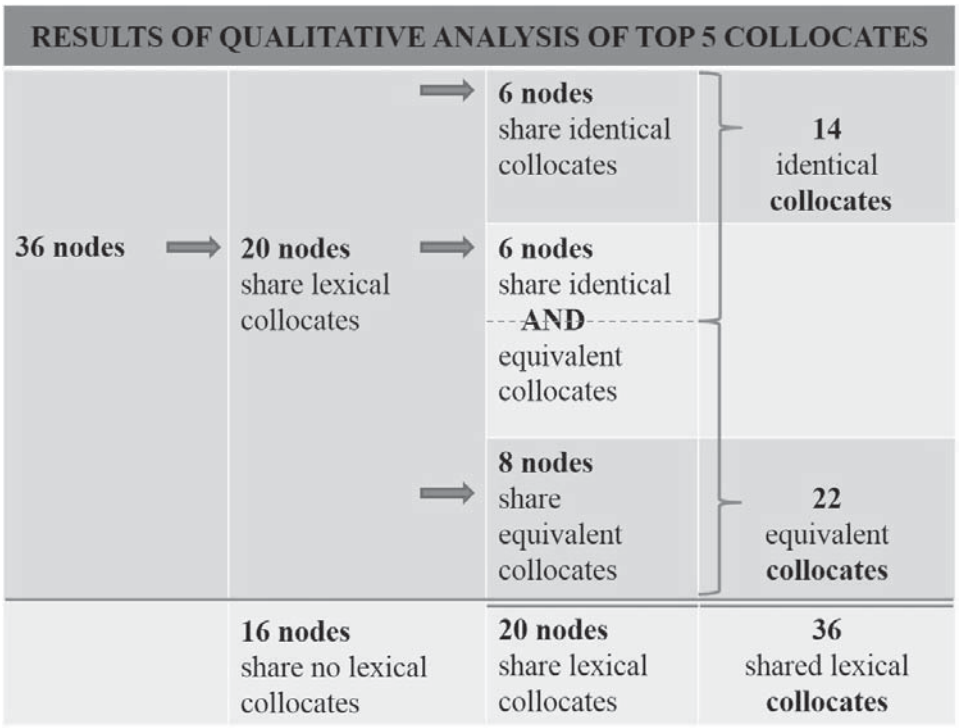

Figure 28 Results of the qualitative analysis of Top 5 business magazine collocates

When comparing Figures 25 (see page 100) and 28, the different results leap out immediately. A synopsis of the results of the quantitative and qualitative analysis of the business magazine corpora is provided in Figure 29 below.

SYNOPSIS OF QUANTITATIVE AND QUALITATIVE ANALYSIS OF THE BUSINESS MAGAZINE CORPORA									
				QUANTITATIVE ANALYSIS		QUALITATIVE ANALYSIS			
NODE		FREQUENCY SL	FREQUENCY TL	IDENTICAL COLLOCATES	EQUIVALENT COLLOCATES	IDENTICAL COLLOCATES		EQUIVALENT COLLOCATES	REASON FOR DEVIATION
(1) ANALYST	x	427	156	2	0	0		0	Compound/ Proper name
(2) ASSET	x	235	50	4	0	0		0	Compound
(3) BANKING	x	352	44	1	0	0		0	Compound
(4) BROKER		101	35	0	0	0		0	-
(5) BUSINESS	x	2,641	343	2	0	0		0	Compound/ Proper name
(6) BUYOUT		143	49	0	2	0		2	≙ -
(7) CASH		823	172	2	2	1	≡	2	≙ Compound
(8) CORPORATE		667	56	1	0	1	≡	0	-
(9) CRASH		55	111	1	1	1	≡	1	≙ -
(10) DEAL		780	150	0	0	0		0	-
(11) DESIGN		385	210	1	1	1	≡	1	≙ -
(12) DEVELOPMENT	x	318	29	1	0	0		0	Compound
(13) ECONOMY	x	1,250	68	1	0	0		0	Compound
(14) EQUITY		539	192	1	2	0		2	≙ Compound
(15) FINANCE		414	43	2	1	1	≡	0	Proper name
(16) GROUP		996	156	3	0	0	≡	0	Proper name
(17) IMAGE		105	201	0	2	0		2	≙ -
(18) INSIDER	x	30	195	0	1	0		0	Compound
(19) INSIGHT	x	39	35	1	0	0		0	Proper name
(20) INVESTMENT	x	1,240	177	1	0	0		0	Compound
(21) INVESTOR		247	199	2	0	1	≡	0	Compound
(22) MANAGEMENT		840	523	3	1	0		1	≙ Compound/ Proper name
(23) MANAGER		488	1,330	0	1	0		1	≙ -
(24) MARKET		2,441	55	3	0	1	≡	0	Compound
(25) MARKETING		413	207	1	2	1	≡	2	≙ -
(26) OFFICER		262	41	4	0	3	≡	0	Compound
(27) PORTFOLIO		323	101	1	2	0		2	≙ Compound
(28) PROVIDER	x	73	28	2	0	0		0	Compound
(29) PUBLIC	x	544	26	1	0	0		0	Pseudo-Anglicism
(30) RATING	x	132	115	0	2	0		0	Compound
(31) RESEARCH		848	96	0	1	0		1	≙ -
(32) SERVICE		710	268	1	2	1	≡	2	≙ -
(33) SOFTWARE		625	382	2	2	1	≡	2	≙ Proper name
(34) SUPPLY		222	42	2	1	0		1	≙ Compound/ Proper name
(35) VALUE	x	662	71	1	0	0		0	Proper name
(36) WORK	x	1,118	62	1	0	0		0	Compound

Figure 29 Synopsis of the results of the quantitative and qualitative analysis of business magazine corpora

Figure 29 facilitates the comparison between the results from the quantitative and qualitative corpus analysis. The qualitative corpus analysis revealed that 14, out of the originally 34 nodes thought to display similar collocational behavior do not. These are marked with a small 'x' in figure 29. Identical collocates are marked with the mathematical 'identical to' sign '≡' and equivalent collocates are marked with the mathematical 'corresponds to' sign '≙'.

There is one main reason for the deviation in the identical and equivalent lexical collocate count: The statistical corpus analysis of collocations also identifies constituents of a compound and constituents of proper names as collocates of a given node. It can be seen in Figure 29 that for 19 nodes the quantitatively identified collocates did not ultimately qualify as such because of compounding. In 9 cases the quantitatively identified collocates actually formed a proper name with the given node, such as a company name. In one case, the quantitatively identified identical lexical collocates co-occurred with the node in the target language corpus only in the form of a pseudo-Anglicism.

The deviation from quantitative and qualitative corpus analysis is directly connected to one weakness of the chosen methodology. The lexical analysis software cannot differentiate between constituents of compounds and collocates. As a result, successive qualitative analysis is required. In addition, *WordSmith Tools* cannot differentiate between proper nouns, such as company names, and collocates. This is closely related to the type of corpora which was used. One advantage of an annotated corpus would have been that the software would have been able to identify proper names and differentiate them from collocates.

The qualitative analysis of lexical collocates allows for a few obvious and some subtler conclusions. First of all, it can be stated that when dealing with collocations in large corpora, quantitative analysis is highly advantageous for a pre-selection of decisive material from the corpora. Nevertheless, the statistical corpus analysis will determine words which frequently occur in L1 and R1 position of the collocational span as collocates, although some of them are actually constituents of a compound formed with the node. Since compounding is a productive source of new words in German and English this happens quite frequently (cf. Figure 29 above). In any case, such erroneously purely statistically identified collocates need to be eliminated by manual analysis since they cannot be used to evaluate similarities in collocational behavior. In sum, the results from the quantitative corpus analysis would have been less meaningful without the support of the qualitative analysis.

Strikingly, many target language collocates are Anglicisms. The following numbers substantiate this claim. 163 lexical Top 5 collocates of the 36 nodes were identified in the target language corpus. Out of these 163 lexical target language collocates 100, or 61.3%, are Anglicisms. Only 63 Top 5 target language collocates, or 38.7%, are German collocates. The fact that numerous Anglicisms were identified as collocates clearly shows that Anglicisms are rarely borrowed as isolated items. Instead, Anglicisms are borrowed along with identical collocates or as a constituent of a source language compound which is used in identical form in the target language.

As far as word classes are concerned, nouns constitute by far the largest group of the Top 5 collocates. This and the frequencies of adjectives, verbs and adverbs among the Top 5 collocates in both corpora are summarized in Figure 30 below.

TOP 5 COLLOCATES ACCORDING TO WORD CLASSES		
WORD CLASS	**SOURCE LANGUAGE**	**TARGET LANGUAGE**
NOUNS	129	118
ADJECTIVES	41	34
VERBS	4	11
ADVERBS	1	0

Figure 30 Top 5 collocates according to word classes in business magazines

It has been stated earlier that German regularly employs English loans to denote new items, concepts or thoughts. Nouns, especially, can describe new items more readily than can adjectives or verbs (cf. Carstairs-McCarthy 2006: 62). Therefore, it is not surprising that 29 nouns were identified as shared, i.e. identical or equivalent, collocates but only 6 adjectives, according to the qualitative corpus analysis. The analysis of the news magazine corpora will have to show whether this is typical of business magazine articles and thus domain-dependent. At this point, it is hypothesized that the analysis of the news magazine corpora will reveal more adjectives and verbs as shared collocates than were identified in the business magazine corpora. This is expected, because technical terms are mostly nouns and more likely to occur in specialized text than in general news reporting.

From the outset of the present study, it was assumed that highly frequent Anglicisms in the target language corpus would tend to co-occur with German collocates. This was hypothesized because the overall frequency of an Anglicism can be seen as an indicator of its degree of integration and the same can be assumed for the number of German collocates of an Anglicism in the target language corpus. The corpus analysis revealed that nodes collocate typically with target language word-forms, instead of additional Anglicisms, if their overall frequency is higher in the target language than in the source language corpus. Vice versa there is a tendency for low-frequency nodes, i.e. whose overall frequency is below 100 in the target language corpus, to co-occur with numerous, sometimes exclusively, Anglicisms in the target language corpus. Relevant examples are **ASSET**, **BANKING**, **BROKER**, **CORPORATE**, **DEVELOPMENT**, **ECONOMY**, **FINANCE**, **MARKET**, **OFFICER**, **PROVIDER**, **PUBLIC**, **RESEARCH**, **VALUE**, **WORK**. However, a small number of low-frequency nodes was discovered which feature only German collocates in the target language corpus. Equally, highly frequent Anglicisms also collocate with additional Anglicisms in the target language corpus. Generally speaking then, a given

node's overall frequency in the target language corpus does not reliably indicate whether it will co-occur with additional Anglicisms or target language word-forms.

Another aim of the corpus analysis was to analyze the factors which influence the use of identical and equivalent collocates. The corpus analysis did not expose a conclusive relation between the frequency of a given node and its number of identical and equivalent collocates. It can be concluded that high-frequency and low-frequency nodes alike display similar collocational behavior in the languages contrasted in the present study.

The number of quantitatively identified identical collocates was drastically reduced during the qualitative corpus analysis. To be exact, of the originally 48 identified identical collocates, only 14 proved to be collocates after the qualitative corpus analysis (cf. Figures 25 and 28). This illustrates the frequency of borrowing identical compounds. In contrast, the number of quantitatively identified equivalent lexical collocates remained almost the same even after their qualitative analysis. That is to say that 22 of the 26 quantitatively identified equivalent lexical collocates proved to be collocates. More to the point, quantitatively identified equivalent lexical collocates were more likely to illustrate similar collocational behavior than identical ones.

The corpus analysis revealed another shortcoming of the methodology used in the present study which concerns, above all, the exceptionally frequent nodes. The difficulty with very frequent nodes is that they usually feature many more lexical collocates in the source and target language corpus than are analyzed by looking only at the Top 5. The more instantiations of a given node there are, the more likely it is to have more than five lexical collocates in both corpora. This means that the risk of missing significant identical or equivalent collocates increases the more lexical collocates a node features, if the corpus analysis is limited to the Top 5 collocates. As a consequence, the method of investigation becomes less reliable the more lexical collocates a node features in addition to the Top 5 collocates. Very frequent nodes are thus not ideal for Top 5 analysis. Because of the large number of additional (in addition to Top 5) lexical target language collocates, overall certainly more shared collocates of the 36 nodes could have been identified (cf. for example **INVESTMENT** and Figure 27).

In general, many target language collocates are highly frequent Anglicisms (some even serve as nodes for the study). This does not mean though that such identified target language collocates are also source language collocates. The sheer number of Anglicisms which are identified as significant lexical target language collocates does not necessarily indicate similar collocational behavior of a node. For some cases, it can be recorded that nodes have developed distinctive target language collocations with source language word-

forms (cf. subsection 5.1.2). Consider for example the node **VALUE** which co-occurs with numerous Anglicisms in the target language which are not identified as source language collocates (cf. subsection 6.2.2). Certainly, the corpora are not necessarily representative of all collocations a given node may enter in the source language. However, in many cases, those Anglicisms were not identified as source language collocates because they do not co-occur with the node in the source language. In other words, the target language uses source language word-forms to create a new collocation or compound in the target language with source language word-forms that is unknown in the source language. This phenomenon can be described as 'pseudo-Anglicism' or 'Anglo neologism'.

A differentiation between the two is difficult because essentially they refer to the same phenomenon. However, Anglo neologisms are always semantically transparent and modeled along existing source language structures. ONLINE **BANKING** can be seen as an example of an Anglo neologism. It imitates the structure of PRIVATE **BANKING** which exists in the source language and is semantically transparent; ONLINE **BANKING** is banking via the Internet. However, it is difficult to prove that the imitation of the source language model actually influences the creation of the Anglo neologism. **CORPORATE** COLLECTION can be described as a pseudo-Anglicism. Its meaning cannot be derived from its single constituents outside of a given target language context. It could refer to a collection (of some kind) which belongs to a company, a corporation. In the target language corpus though, **CORPORATE** COLLECTION refers to a collection of clothing which is embroidered with the company's logo.

Sometimes compound Anglicisms are used very frequently in the target language corpus, but do not even occur in the source language corpus and are not defined in any of the chosen dictionaries. Generally speaking, they are then suspicious of being pseudo Anglicisms. Only in a very limited number of cases the reason for this is that they are very specific technical terms such as PRIME **BROKER** which exist also in the source language, but which are rarely used. Because some of such terms are missing institutionalization as compounds, they cannot yet be found in dictionaries.

The fact that compounding is such a productive source of new words in German and English possibly leads to an increased formation of hybrid compounds in the target language corpus. The frequent use of compounding and the similar structures of compounds in the source and target language, have led to the formation of hybrid Anglicisms like **RATING**-AGENTUREN from **RATING** AGENCIES, **INSIDER**-HANDEL from **INSIDER** TRADING, or **CASH**-POLSTER from **CASH** CUSHION. Hybrid Anglicisms are based on a combination of English borrowings with German elements (cf. Onysko 2007: 55).

Zindler (cf. 1959: 16) and Carstensen (cf. 1965: 39) try to divide hybrids into compounds that follow an English model and compounds without an English model. However, this is difficult to put into practice and Busse and Carstensen themselves state that hybrid compounds are particularly difficult to classify (cf. 2001: 66). Hybrid compounds can also be called 'partial loan translation' due to the fact that one element from the original lexical unit is retained, whereas the other element is substituted with a German translational equivalent. The high frequency of hybrid compound nouns and hybrid collocations, for example **BUYOUT** GESCHÄFT, **BUYOUT** FIRMEN, underlines the fact that hybrid formation is a major process of lexical innovation with nominal Anglicisms in *WirtschaftsWoche 2008*.

The qualitative corpus analysis illustrates that frequently source language compounds are used in identical form in the target language, i.e. **ASSET** MANAGER, **PORTFOLIO** MANAGER and many more. Why **BUYOUT** DEAL, for example, is not borrowed as one lexical item, but as **BUYOUT** GESCHÄFT remains unresolved. The qualitative corpus analysis allows for the educated guess that this is due to the missing institutionalization of **BUYOUT** DEAL as a compound in the source language. At least in the chosen dictionaries no lexical entries exist. This again frequently leads to the translation of collocates and results in hybrid target language collocations like **BUYOUT** GESCHÄFT. It cannot be ultimately dismissed that theoretically **BUYOUT** DEAL and **BUYOUT** GESCHÄFT may have arisen independently in English and German.

Often, the use of source language compounds is preferred although target language equivalents exist. It should be noted that some source language compounds, for example **BUSINESS** SCHOOL, or INVESTMENT **BANKING** have gained acceptance in the target language to the degree that they effectively suppress the use of their target language equivalents. All the same, sometimes source language compounds and their target language equivalents are used interchangeably in the target language corpus, i.e. **MARKET** RESEARCH and 'Marktforschung' and the target language term still dominates. This is a popular strategy for stylistic variation in business magazine articles. It may be speculated whether this is also to bestow an international appearance on *WirtschaftsWoche*.

The qualitative corpus analysis suggests the possibility that semantically less transparent compounds are borrowed habitually as one lexical item, giving up the institutionalization of an equivalent target language term. Consider for example the compound NEW **ECONOMY**. Remarkably, even in bilingual dictionaries, for some Anglicisms a German translation does not even exist or is given as a translation equivalent (anymore). Once such institutionalization has taken place, an equivalent German term is unlikely to be ever coined. And in most such cases never has existed. Renowned bilingual dictionaries give

Anglicisms as translations of Anglicisms. Consequently “INSIDER”, “ANALYST”, or “**INVESTOR** RELATIONS” are listed as German translations for example (www.pons.eu).

The corpus analysis shows that there is continuous variation in the extent to which a collocation is fixed or variable. That is to say, whether a collocate co-occurs with a node in a restricted or larger number of span positions is a reference to the degree of fixedness of a collocation (cf. also Biber 2009). The extent to which a collocation is variable increases by the number of collocational span positions a given collocate can take.

6.3 News magazines

6.3.1 Quantitative analysis

The nodes for the analysis of the news magazine corpora were also identified via close reading (cf. section 5.2). The primary difficulty for the analysis of the news magazine corpora was to identify nodes which feature five lexical collocates. This depends of course on the frequency of the nodes. However, it became clear immediately that more general nodes have to be much more frequent in order to feature sufficient analyzable collocates. This applies to the news magazine source and target language corpora alike. Therefore, and in contrast to the analysis of the business magazine corpora, an overall minimum frequency of each node of 100 is determined. Still, this does not result in five lexical collocates for each node. Even nodes with 150 instantiations and more do not all feature a minimum of five lexical collocates. This is very different from the more specialized nodes which were selected for the analysis of the business magazine corpora. However, because not every single node for the business magazines features five lexical collocates, it is deemed adequate not to raise the minimum frequency of news magazine nodes further, but to stick with a minimum frequency of 100 and accept that “some words are lonelier than others” (Firth 1966: 183). As a first result, it can thus be recorded that more general nodes in more general corpora feature less lexical collocates than more technical nodes in specialized text. It should also be noted that in contrast to the business magazine nodes, the news magazine nodes generally featured more verbs and adjectives as collocates.

When calculating the frequencies of the nodes and their collocates, a few opening observations were made. Again, highly frequent nodes in the source language do not necessarily turn into highly frequent Anglicisms and there are four nodes which, judging strictly from absolute numbers (cf. Figure 32), are more frequent in the target language corpus than in the source language corpus: **BESTSELLER**, **CLAN**, **COMEBACK** and **KILLER**. Busse

and Carstensen's *Anglizismenwörterbuch* (2001) states that **BESTSELLER** was introduced as an Anglicism already early in the twentieth century. **CLAN** became known even earlier in the second half of the eighteenth century. The first use of **COMEBACK** dates back to 1947 and first occurred in *Der Spiegel*. Back then, **COMEBACK** was either still spelled as two words separated by a blank space or a hyphen. The first use of **KILLER** is recorded for 1959 in *Stuttgarter Zeitung*. In sum, it can be stated that nodes which occurred more often in the target language news magazine corpus than in the source language one were comparatively early introduced to German. This confirms the results from the analysis of the nodes which are more frequent in the target language business magazine corpus than in the source language one (cf. subsection 6.1.1). It will be interesting to see whether the results for nodes like this in the business magazine corpora also hold true for such nodes in the news magazine corpora.

In order to facilitate comparison, the following figures display the frequencies of the nodes. Figure 31 displays the nodes in both corpora, descending from highest to lowest frequency.

FREQUENCY OF NODES			
SOURCE LANGUAGE		TARGET LANGUAGE	
TREND	(2,069)	JOB	(443)
LIFE	(1,652)	TREND	(417)
STORY	(1,162)	E-MAIL	(339)
TALK	(1,052)	TEST	(274)
COMMUNITY	(783)	STAR	(256)
JOB	(718)	CLAN	(217)
NEWS	(629)	STORY	(199)
TEST	(535)	TALK	(198)
MEETING	(497)	INTERVIEW	(195)
SOCIETY	(459)	BOOM	(180)
STAR	(279)	KILLER	(180)
INTERVIEW	(398)	SOUND	(171)
REALITY	(358)	COMEBACK	(165)
E-MAIL	(345)	LIFE	(160)
PERFORMANCE	(335)	FAN	(160)
MODEL	(331)	MODEL	(159)
SOUND	(307)	BESTSELLER	(134)
NETWORK	(279)	REALITY	(132)
EVENT	(202)	ENTERTAINMENT	(130)
FARM	(198)	SOCIETY	(129)
BOOM	(191)	COMMUNITY	(127)
ENTERTAINMENT	(184)	EVENT	(120)
FAN	(176)	FREAK	(119)
CAMP	(165)	NETWORK	(119)
FASHION	(156)	FARM	(116)
COMEBACK	(149)	GANGSTER	(115)
FREAK	(144)	CAMP	(115)
KILLER	(144)	MEETING	(115)
GANGSTER	(142)	FASHION	(114)
LUNCH	(121)	LUNCH	(110)
KNOW-HOW	(118)	COMEDY	(109)
BESTSELLER	(117)	PERFORMANCE	(109)
DOTCOM	(115)	NEWS	(107)
COMEDY	(113)	DOTCOM	(106)
AIRPORT	(111)	KNOW-HOW	(104)
CLAN	(105)	AIRPORT	(101)

Figure 31 Frequency of nodes in the news magazine corpora

Figure 32 below shows the nodes and their frequencies in alphabetical order. Frequencies range from 105 to 2,069 within the source language and from 101 to 443 in the target language corpus. As a reminder, the total word count of the *Newsweek* corpus is 1,823,194 words, whereas the *Der Spiegel* corpus comprises a total of 3,332,784 words (cf. section 4.6). Because of the difference in corpus size, the frequencies of all nodes are also given per 1 million words. This again facilitates comparability of frequencies of the identical nodes in differently sized corpora.

FREQUENCY IN ALPHABETICAL ORDER				
NODE	SOURCE LANGUAGE	FREQUENCY/ 1 MIO WORDS	TARGET LANGUAGE	FREQUENCY/ 1 MIO WORDS
AIRPORT	(111)	61	(101)	30
BESTSELLER	(117)	64	(134)	40
BOOM	(191)	105	(180)	54
CAMP	(165)	91	(115)	35
CLAN	(105)	58	(217)	65
COMEBACK	(149)	82	(165)	50
COMEDY	(113)	62	(109)	33
COMMUNITY	(783)	429	(127)	38
DOTCOM	(115)	63	(106)	32
E-MAIL	(345)	189	(339)	102
ENTERTAINMENT	(184)	101	(130)	39
EVENT	(202)	111	(120)	36
FAN	(176)	97	(160)	48
FARM	(198)	109	(116)	35
FASHION	(156)	86	(114)	34
FREAK	(144)	79	(119)	36
GANGSTER	(142)	78	(115)	35
INTERVIEW	(398)	218	(195)	59
JOB	(718)	394	(443)	133
KILLER	(144)	79	(180)	54
KNOW-HOW	(118)	65	(104)	31
LIFE	(1,652)	906	(160)	48
LUNCH	(121)	66	(110)	33
MEETING	(497)	273	(115)	35
MODEL	(331)	182	(159)	48
NETWORK	(279)	153	(119)	36
NEWS	(629)	345	(107)	32
PERFORMANCE	(335)	184	(109)	33
REALITY	(358)	196	(132)	40
SOCIETY	(459)	252	(129)	39
SOUND	(307)	168	(171)	51
STAR	(279)	153	(256)	77
STORY	(1,162)	637	(199)	60
TALK	(1,052)	577	(198)	59
TEST	(535)	293	(274)	82
TREND	(2,069)	1,135	(417)	125

Figure 32 Alphabetical order of nodes in the news magazine corpora

According to the definition of collocation in this study (cf. section 2.4), a total of 1,031 lexical collocates of the 36 nodes was detected in the news magazine source language corpus. This results in averagely 29 collocates per node. But the number of lexical collocates per node differs heavily. Within the source language corpus the number of lexical collocates per node ranges from 3 to 117. Within the target language business magazine corpus the number of lexical collocates per node ranges from 3 to 31. A total of 289 collocates was identified in the target language corpus, resulting in an average of 8 collocates per node. The current study aims at analyzing decisive data o nly. The analysis of the news magazine corpora will also concentrate on the 'Top 5' collocates, i.e. the five collocates which co-occur most frequently with a node.

Looking at the Top 5 collocates of each node, the evidence displays that 20 (55.5%) of the 36 nodes share lexical collocates. A total of 11 nodes share identical collocates, 2 nodes feature identical and equivalent collocates and 7 nodes only equivalent collocates. The group of the Top 5 collocates of all nodes includes 22 identical and 12 equivalent collocates. This information is summarized in Figure 33 below.

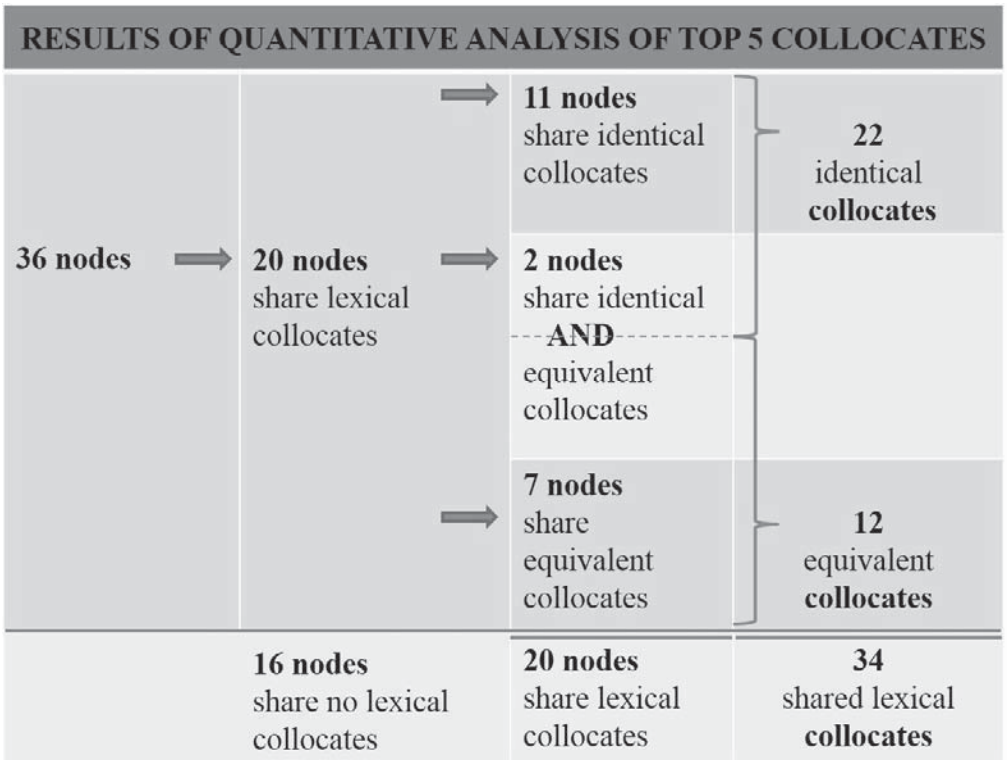

Figure 33 Results of the quantitative analysis of Top 5 news magazine collocates

The different nodes feature different numbers of shared lexical collocates within the Top 5. Their distribution among the 36 nodes, and more particularly the distribution of the 22 identical and 12 equivalent lexical collocates among the 36 nodes, is illustrated in Figure 34 below.

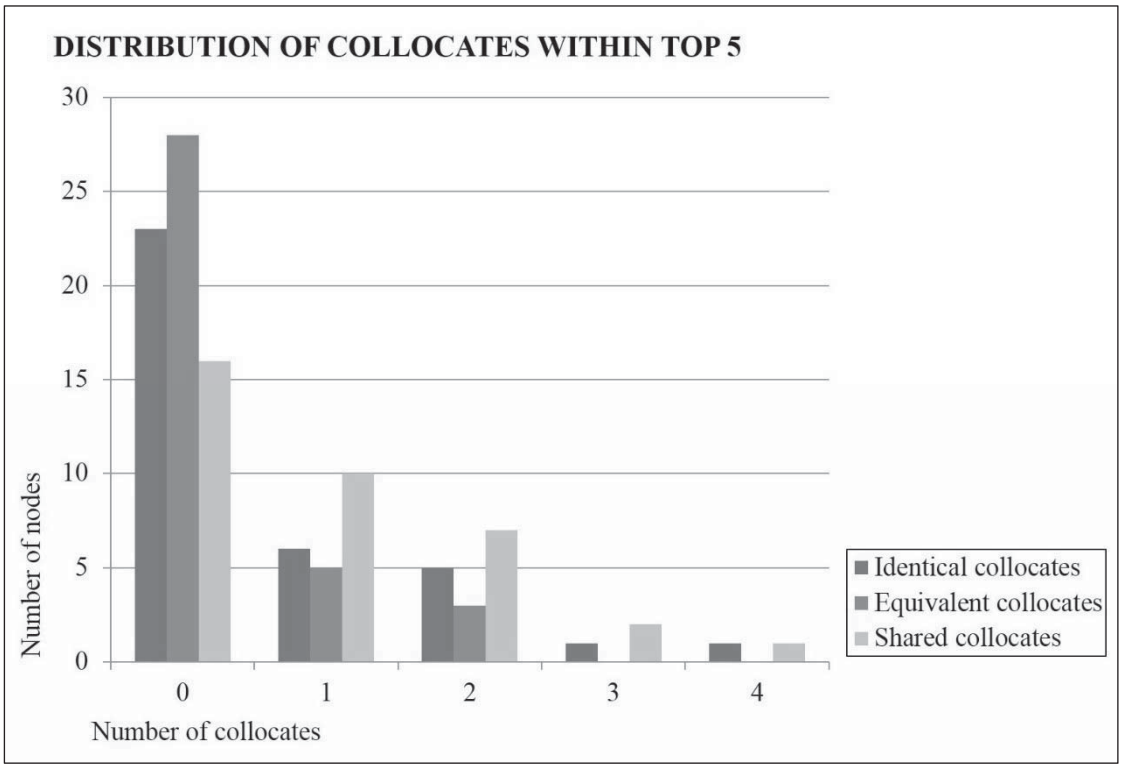

Figure 34 Distribution of news magazine collocates within 'Top 5'

The x-axis in Figure 34 shows the number of collocates. As a result of the quantitative corpus analysis it can be stated that each of the 36 analyzed nodes features a number of 0 to 4 shared lexical collocates among the Top 5 collocates. Thus, the x-axis is labeled from 0 to 4. As has been mentioned above, the nodes feature different numbers of identical and equivalent lexical collocates. This is shown with the different bars in different shades of blue. In addition, the bar in the lightest shade of blue provides the total number of shared collocates, i.e. identical and equivalent collocates of a node are counted. The y-axis shows the number of nodes from 0 to 30.

Moving from left to right on the x-axis, the following information is represented in the diagram: 23 nodes do not feature any identical collocates (but they may feature equivalent collocates), 28 nodes do not feature any equivalent collocates (but they may feature identical collocates), and 16 nodes do not feature any shared (identical or equivalent) collocates at all.

6 nodes feature 1 identical collocate (and may feature additional equivalent collocates), 5 nodes feature 1 equivalent collocate (and may feature additional identical collocates). 10 nodes share a total of exactly 1 (either identical or equivalent) collocate.

5 nodes feature 2 identical collocates (and may feature additional equivalent collocates), 3 nodes feature 2 equivalent collocates (and may feature additional identical collocates).

7 nodes feature a total of exactly 2 shared collocates (either identical or equivalent or a combination of the two).

1 node features 3 identical collocates (and may feature additional equivalent collocates), 0 nodes feature 3 equivalent collocates and may feature additional identical collocates), 2 nodes feature a total of exactly 3 shared collocates (as a result of the combination of identical and equivalent collocates of a given node). For example the node **DOTCOM** features 2 identical and 1 equivalent lexical collocates; in other words **DOTCOM** features a total of 3 shared collocates.

1 node features 4 identical collocates, 0 nodes feature 4 equivalent collocates, 1 node features a total of exactly 4 shared collocates (as a result of the combination of identical and equivalent collocates of a given node).

As an intermediate result and judging from this quantitative analysis of the news magazine corpora, it seems legitimate to assume that similarities and differences in the collocational behavior of Anglicisms apply approximately to the same extent in American and German news magazines. The qualitative analysis is expected to deliver more details on this outcome.

6.3.2 Qualitative analysis

The 20 nodes which evidence, according to the quantitative analysis, similar collocational behavior are subject to supplementary qualitative analysis. Initially, it was hypothesized (cf. section 1.1) that the selected nodes feature identical or equivalent collocates in the target and source language corpora. Further detailed investigation of the 16 nodes which do not share any lexical collocates is thus unlikely to lead to rewarding results. The charts of the 16 nodes which do not share any lexical collocates are included in this subsection of the study, but they are not discussed in detail for the qualitative analysis.

The qualitative analysis of the news magazine corpora is conducted analogously to the qualitative analysis of the business magazine corpora. The nodes are analyzed in alphabetical order. In the analysis charts, ‘T’ stands for the total number of instantiations of a given node, ‘LC’ indicates the total number of lexical collocates of a given node. The numbers in parentheses ‘()’ show how many times node and collocate co-occur within the designated (source or target language) corpus. Because the search for collocates is not case-sensitive, all nodes and collocates are given in capital letters. This is analogous to the modus operandi of *WordSmith Tools* (cf. Figure 21). Pursuing as clear an arrangement as

possible, identical collocates are marked with a mathematical 'identical to' sign '≡' and equivalent collocates are marked with the mathematical 'corresponds to' sign '≙'.

One of the main tasks of the qualitative corpus analysis is to determine reliably which quantitatively identified shared collocates are truly of that kind. In all charts a checkmark '✓' is put by those collocates which prove to be truly identical or equivalent according to the quantitative and qualitative corpus analysis. The outcomes of the quantitative and qualitative corpus analysis differ significantly and a synopsis of the results is provided in subsection 6.3.3.

AIRPORT

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
111	4	MAJOR (6)	bedeutend/ wichtig	CITY (8)	3	101
		NEW (5)	neu	LOUNGE (6)		
		INTERNATIONAL (5)	international	KRISEN (5)		
		SECURITY (5)	Sicherheit	-		
		-	-	-		

The node **AIRPORT** does not feature any identical or equivalent lexical collocates. The node occurs with a similar frequency in both corpora. It is somewhat surprising that the source language collocates NEW and SECURITY were not identified as lexical collocates in the target language corpus because they are generally popular Anglicisms in German and in the news magazine target language corpus. Interestingly, **AIRPORT** features two Anglicisms as collocates in the target language corpus which were not identified as collocates in the source language corpus.

BESTSELLER

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5	≙		Top 5	≙	LC	T
117	5	LIST	(24)	Liste	LISTE	(19)	9	134
		LISTS	(8)	Listen	AUTOR	(14)		
		BECAME	(5)	wurde	PLATZ	(8)		
		ENTERED	(5)	(hin)eingehen	BUCH	(6)		
		PERENNIAL	(5)	mehrjährig	JAHREN	(5)		

The node **BESTSELLER** is more frequently identified in the target language corpus than in the source language corpus. As has been illustrated in subsection 6.1.2 above, nodes with this quality prove to feature almost exclusively target language word-forms as collocates in the business magazine corpus analysis. This is also valid for the news magazine corpora. **BESTSELLER** features exclusively German lexical collocates in the news magazine target language corpus.

It may be speculated that the Anglicism **BESTSELLER** is particularly popular because its German equivalent "Verkaufsschlager" (www.pons.eu) is rather long. In addition, **BESTSELLER** usually refers to books in the source and target language alike. This is also evident from its definitions in *Oxford Dictionaries*: "a book or other product that sells in very large numbers"; and *Merriam Webster*: "an article (as a book) whose sales are among the highest of its class." The German translation 'Verkaufsschlager' would thus have to be further specified to achieve translation equivalency with **BESTSELLER** and clear reference to books. This is necessary because 'Verkaufsschlager' frequently also refers to different products in German. This can be exemplified with news magazine target language corpus material:

> TL 1: […] entdeckt Lena ganz zufällig die Rezeptur für die Currywurst. Die wird prompt ein Verkaufsschlager, und schon hat Lenas Leben wieder Würze.
>
> TL 4: Der Käfer wird heute nicht mehr produziert, als Verkaufsschlager hat ihn der Golf längst abgelöst.

The corpus examples above demonstrate that 'Verkaufsschlager' is used to describe 'Currywurst', 'Käfer' and 'Golf'. **BESTSELLER** refers in both corpora exclusively to books.

The dictionary *PONS* lists as a translation of **BESTSELLER** merely the Anglicism "Bestseller". This indicates its institutionalization. According to Carstensen and Busse's *Anglizismenwörterbuch* (2001), the Anglicism **BESTSELLER** was introduced early in the twentieth century. This is comparatively early and in line with the results for business magazine corpus nodes which were more frequent in the target language corpus than in the source language corpus (cf. subsection 6.2.1).

The node **BESTSELLER** features one equivalent lexical collocate according to the quantitative corpus analysis, i.e. LIST and LISTE. Both were exclusively identified in R1 position of the node **BESTSELLER**. This is illustrated with the following examples:

SL 15: […] has topped The New York Times's paperback nonfiction bestseller list for more than a year.

TL 8: […] steht derzeit in den Niederlanden auf den ersten Plätzen der Bestseller-Liste […].

LIST and LISTE do not qualify as equivalent lexical collocates of **BESTSELLER**. Instead, *PONS* lists **BESTSELLER** LIST as a compound which translates to "Bestsellerlisten". This spelling variety of the compound was also identified three times in the target language corpus.

BOOM

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
191	20	OIL	(39)	Öl	GESCHÄFT	(18)	14	180
		HOUSING	(18)	Wohnungen	JAHREN	(12)		
		BABY	≡ (14)	Baby/ Säugling	INTERNET	(12)		
		TIME	(11)	Zeit	BABY	≡ (10)		
		GENERATION	(10)	Generation	CRASH	(6)		

The node **BOOM** features one identical lexical collocate. In the source language and target language corpus, BABY was identified exclusively in L1 position. Consider the following corpus examples:

SL 19: Despite our size (the peak of the baby boom was 1957, the year I was born), we spent years […].

TL 4: Ein Baby-Boom sieht anders aus.

Oxford Dictionaries defines BABY **BOOM** as "a temporary marked increase in the birth rate, especially the one following the Second World War." Consequently, BABY is a constituent of the compound BABY **BOOM** and does not qualify as an identical lexical collocate of **BOOM** as of the qualitative corpus analysis.

Two further Anglicisms were identified among the Top 5 target language collocates of **BOOM**: INTERNET and CRASH. An additional Anglicism, DOTCOM, was identified as a lexical target language collocate. All of these were identified as collocates of **BOOM** in the source language corpus, but they did not occur frequently enough to make the Top 5 source language collocates. Examples for the co-occurrence of INTERNET and **BOOM** as well as DOTCOM and **BOOM** from both corpora are provided below.

SL 102: The two universities have become intellectual centers of the Internet boom, doing their best to attract […].

SL 146: Just as in the original dotcom boom of the late 1990s, Valley types claimed these valuations were perfectly reasonable, and insist […].

TL 47: […] die ihren großen Reichtum aus dem Dotcom-Boom nicht mit dem Fiskus teilen wollten.

TL 59: Viele, die heute an der Wall Street arbeiten, wissen nicht mal, was 2000 nach dem Internet-Boom passierte.

DOTCOM and INTERNET typically collocate with **BOOM** in L1 position in both corpora. However, no lexical entry for either DOTCOM **BOOM** or INTERNET **BOOM** exists in the chosen dictionaries. The collocate CRASH usually appears further left or right of **BOOM** in the collocational span. Consider the following examples from the source and target language corpora.

SL 38: How to Prosper in the Crash Following the Greatest Boom in History By Harry S. Dent Jr.

TL 15: Diese Dynamik gehört zum Betriebssystem des Kapitalismus: Zyklen von Boom und Crash, von Größenwahn und Panik.

In sum, the node **BOOM** shows potential for three more identical lexical collocates, but they cannot be counted due to lack of frequency and the chosen Top 5 methodology.

CAMP

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
165	17	TRAINING	(10)	Ausbildung/ Schulung	ROCK	(11)	3	115
		CONCENTRATION	(10)	Konzentrations-	NEUEN	(7)		
		BOOT	(10)	Ausbildungs-	FILM	(5)		
		SUMMER	(8)	Sommer	-			
		PRISON	(7)	Gefängnis	-			

As the chart above illustrates, **CAMP** does not share any identical or equivalent lexical collocates in the source and target language corpora. The Anglicism ROCK which was identified as a significant target language collocate, does not co-occur with **CAMP** in the source language corpus. The reason for this is simple. In each instantiation in which **CAMP** and ROCK co-occurred in the target language corpus, they formed part of a movie title. For exemplification, see the corpus example below.

TL 32: Im Film "Camp Rock", der in Deutschland Mitte September bei ProSieben zu sehen sein wird […].

CLAN

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
105	3	THREAT	(5)	Drohung	MITGLIEDER	(16)	10	217
		PROTECTIVE	(5)	Schutz-/ fürsorglich	CHEF	(9)		
		MUSICAL	(5)	musikalisch	INDUSTRIELLEN	(9)		
		-		-	MÄCHTIGEN	(8)		
		-		-	KRIMINELLEN	(8)		

The node **CLAN** does not feature any identical or equivalent lexical collocates. Two things are worth mentioning about this node. Firstly, **CLAN** is one of the few nodes which is more frequent in the target language corpus than in the source language one. Secondly, all 10 identified lexical collocates of **CLAN** in the target language corpus are target language word-forms. This is typical for nodes of this kind and can by now be regarded as a proven fact.

COMEBACK

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
149	6	MAKES	(12)	macht	ALBUM (20)	3	165
		KID	(10)	Kind/ Jugendlicher	HIT (7)		
		SURPRISING	(10)	Überraschend	WERK (6)		
		MAKING	(10)	machen	-		
		RECENT	(5)	kürzlich	-		

The node **COMEBACK** does not feature any identical or equivalent lexical collocates. The chart above illustrates this. The Anglicism HIT, which is defined by *Oxford Dictionaries* as "a successful venture, especially a film, pop record, or song" and translated by *PONS* as "Riesenerfolg", was not identified as a source language corpus collocate.

COMEDY

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
113	8	NEW	(8)	neu	KULTUR (9)	6	109
		✓ SHOW	≡ (7)	Schau/ Sendung	GESCHÄFT (9)		
		TRAGEDY	(5)	Tragödie	KLASSIKER (7)		
		ROMANTIC	(5)	romantisch	✓ SHOW ≡ (6)		
		STAND-UP	≡ (5)	Stegreif	STAND-UP ≡ (5)		

The node **COMEDY** features two identical lexical collocates in the source and target language corpora according to the quantitative corpus analysis. In the source and target language corpora SHOW co-occurs with **COMEDY** in R1 and R2 positions. Below, this is illustrated with corpus examples.

SL 31: "State of the Union," Ullman's first sketch-comedy show since becoming a Yank, is full of celebrity takedowns.

SL 80: Indeed, both candidates have made frequent appearances on the quip-laden, BBC comedy game show Have I Got News for You.

TL 15 […] widersprach Hitchens' Diagnose mit Stand-up-Auftritten und einer gefeierten Comedy-TV-Show.

TL 26: Die makabre Inszenierung mit einem Schauspieler des örtlichen Theaters ist keine Comedy-Show, sondern eine politische Aktion der Kommunisten.

In the chosen mono- and bilingual dictionaries, no lexical entry for **COMEDY** SHOW was detected. Thus, and although SHOW is at times identified in R1 position, it qualifies as an identical collocate of **COMEDY** rather than a compound.

The second identical lexical collocate, STAND-UP, co-occurs with **COMEDY** in both corpora exclusively in L1 position. Below, examples of this are provided.

SL 110: I started reading books on stand-up comedy.

TL 31: Ihr Weg zum Ziel der Sehnsucht führte über Stand-up-Comedy und einen fünf Jahre gespielten Broadway-Theaterhit nach Hollywood.

STAND-UP **COMEDY** is translated by *PONS* to "Stegreifkomödie" or "Improvisationskomödie". STAND-UP is not a collocate of **COMEDY**, but one constituent of the compound STAND-UP **COMEDY**.

COMMUNITY

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
783	67	BLACK	(45)	schwarz	INTERNET (22)	4	127
		INTERNATIONAL	(43)	international	NETZ (11)		
		ORGANIZER	(29)	Organisator	WEB (10)		
		✓ GAY	≡ (25)	homosexuell/ schwul	✓ GAY ≡ (8)		
		BUSINESS	(23)	Geschäft/ Handel	-		

The node **COMMUNITY** shares one identical lexical collocate, i.e. GAY. In the source language corpus, GAY co-occurs with **COMMUNITY** occasionally in R3, but most frequently in L1 position. In the target language corpus, GAY is only identified in L1 position of **COMMUNITY**. Consider the examples below.

SL 292: He's resentful that the gay community has appropriated his son's murder as part of a larger cause.

SL 542: But not everyone in the community backs the gay parade.

TL 10: […] Hobby-Sportler Putin wird von vielen Frauen, aber auch in der Gay-Community verehrt.

Although GAY co-occurs with **COMMUNITY** most frequently in L1 position in both corpora, it does not qualify as a compound. Lexical entries for GAY **COMMUNITY** were not found in the chosen mono- and bilingual dictionaries. GAY thus proves to be a collocate of **COMMUNITY** according to the quantitative and qualitative corpus analysis.

DOTCOM

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
115	5	✓ BUBBLE ≙ (12)	Blase	✓ BLASE ≙ (9)	4	106
		✓ BOOM ≡ (10)	Aufschwung	✓ BOOM ≡ (7)		
		✓ CRASH ≡ (10)	Zusammenbruch	✓ CRASH ≡ (6)		
		BUST (7)	Niedergang/ Pleite	PHANTASIALAND (5)		
		MILLIONAIR (5)	Millionär	-		

The node **DOTCOM** features two identical and one equivalent lexical collocates. In both corpora all of the shared collocates are exclusively identified in R1 position. Beginning with the equivalent lexical collocate BUBBLE and BLASE, this is exemplified below.

SL 4: By the 1990s, it was the rise of the Internet, which collapsed with the dotcom bubble and gave way to the housing boom and the financing that paid for it.

TL 2: Der US-Ökonom Robert Shiller, der zur Jahrhundertwende bereits das Platzen der Dotcom-Blase prognostiziert hatte […].

DOTCOM BUBBLE and **DOTCOM** BLASE do not have lexical entries in the chosen monolingual and bilingual dictionaries. Thus, BUBBLE and BLASE qualify as equivalent lexical collocates of **DOTCOM**.

The first identical lexical collocate of the node **DOTCOM** is BOOM. According to the known criteria, **DOTCOM** BOOM does not qualify as a compound and BOOM is counted as an identical lexical collocate of **DOTCOM**. In order to illustrate the co-occurrence of **DOTCOM** and BOOM, examples from both corpora are provided below.

SL 3: But penny-pinching went out of style once the dotcom boom started.

TL 43: […] die bei Merrill Lynch und Morgan Stanley jahrelang mit optimistischen Prognosen den Dotcom-Boom befeuerten.

The other identical lexical collocate of **DOTCOM** is CRASH. In both corpora CRASH occurs most frequently in R1 position of the collocational span of **DOTCOM**. One example per corpus of the use of **DOTCOM** CRASH is given below.

SL 5: Especially at risk are the new startups created in the wake of the last dotcom crash in 2001.

TL 4: Nach dem Dotcom-Crash und dem 11. September sorgte er für eine regelrechte Geldschwemme […].

No lexical entries for **DOTCOM** CRASH were found in the chosen mono- and bilingual dictionaries. Thus, CRASH qualifies as an identical lexical collocate of **DOTCOM**.

E-MAIL

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
345	32	SENT	(21)	gesendet	VERKEHR	24	339
		✓ PHONE	≙ (9)	Telefon	ADRESSEN		
		WROTE	(8)	geschrieben	WERDEN		
		ACCOUNT	(8)	Konto	✓ TELEFON		
		ADDRESSES	≙ (6)	Adressen	COMPUTER		
							(5)

The node **E-MAIL** features two equivalent lexical collocates in the source and target language corpora. In the source language corpus, three spelling variations, ‘E-mail’, ‘e-mail’ and ‘email’ were identified, with a preference for ‘e-mail’. In the target language corpus, only the spelling variant ‘E-Mail’ can be found.

Originally it was considered to analyze the node **MAIL**. However, 234 out of 345 instantiations in the source language corpus, and 122 out of 339 in the target language corpus concerned **E-MAIL** rather than **MAIL**. In addition, all instantiations which are of importance to determine the status of equivalent collocates concern **E-MAIL**. Thus, the qualitative corpus analysis revealed that the node **E-MAIL** is more promising than **MAIL** and accordingly the analyzed node is modified to be **E-MAIL** instead of **MAIL**.

The German translation of **E-MAIL**, according to *PONS*, is either the Anglicism "E-Mail" or "elektronische Post". The latter is not used at all in the target language corpus. The original node **MAIL** has two meanings in the source language. The meaning of **MAIL** as "letters and parcels sent by post" (www.oxforddictionaries.com), is not used in the target language corpus. Instead, **MAIL** is used with the meaning of **E-MAIL** in the target language corpus. For reasons of comparability, it is thus advantageous that in both corpora PHONE and TELEFON merely co-occur with **E-MAIL**, not with **MAIL**.

The first equivalent lexical collocate of **E-MAIL** is PHONE and TELEFON. In the source language corpus, PHONE co-occurs with **E-MAIL** in L3, L2 and R2 position. The examples below illustrate this.

SL 139: Fenty and Rhee communicate several times a day by e-mail and phone.

SL 203: That means the best online teachers are easily accessible, if not by phone, then by E-mail, instant message, or some other method.

SL 226: […] had begun collecting vast amounts of information about the phone and e-mail records of American citizens.

As two popular ways of modern communication PHONE/TELEFON and **E-MAIL** are likely to be mentioned together. In the target language corpus, TELEFON co-occurs with **E-MAIL** in L2 and L1 position. Below corpus examples are provided.

TL 23: […] war binnen einer Woche per Telefon und E-Mail verhandelt […].

TL 142: Zu diesem Zweck sollen Betroffene Marks & Spencer per Telefon, E-Mail oder Postkarte auffordern, eine gerechtere Preispolitik zu betreiben.

In summary, **PHONE** and **TELEFON** qualify as equivalent collocates of the node **E-MAIL**.

The second equivalent lexical collocate of **E-MAIL** is ADDRESSES and ADRESSEN. Much like PHONE and TELEFON, ADDRESSES and ADRESSEN only co-occur with **E-MAIL**, rather than **MAIL**. In the source language corpus, ADDRESSES co-occurs with **E-MAIL** in L4 and most frequently R1 position. Corresponding examples are provided below.

SL 20: […] providing addresses and the e-mail of suspects, so long as the cops have a search warrant or sub-poena.

SL 122: He collects E-mail addresses and invites candor in "all hands call" meetings with troops.

SL 320: […] and the e-mail addresses of such networks as Radio Free Asia and Tibet Web.

ADRESSEN co-occurs with **E-MAIL** in the target language corpus in R1 position of the collocational span. Consider the following example.

TL 14: Außerdem seien die Daten, weil wenige E-Mail-Adressen und gar keine Bank- oder Vertragsdaten enthalten seien […].

The bilingual dictionary *PONS* lists **E-MAIL** ADDRESS as a compound with the translation "E-Mail-Adresse". Thus ADDRESSES and ADRESSEN do not qualify as equivalent lexical collocates of the node **E-MAIL**. The singular forms of ADDRESSES and ADRESSEN, i.e. ADDRESS and ADRESSE were also identified as a significant lexical collocate of **E-MAIL** in the source language corpus. But they did not co-occur frequently enough with **E-MAIL** to make the Top 5.

ENTERTAINMENT

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5	TRANSLATION	Top 5	LC	T
184	7	INDUSTRY (9)	Industrie	WEEKLY (6)	3	130
		SCENE (6)	Szene/ Schauplatz	HOME (6)		
		ONLINE (5)	online	KONZERN (5)		
		SHIFT (5)	Wechsel/ Änderung	-		
		MEDIA (5)	Medien	-		

The node **ENTERTAINMENT** does not feature any identical or equivalent lexical collocates. However, two of the Top 5 target language collocates are Anglicisms. WEEKLY turns up as a frequent target language collocate as a constituent of the magazine title *Entertainment Weekly*. Thus, WEEKLY forms part of a proper name, and is not a collocate of **ENTERTAINMENT**. Below, a corresponding example from the target language corpus is given.

TL 15: In einem Interview mit dem US-Magazin "Entertainment Weekly" erklärt Jolie […].

The target language collocate *HOME* is also an Anglicism. In the target language corpus it is only identified in L1 position. For an example see below.

TL 12: Doch während sich das computergestützte Home-Entertainment mit riesigen Flachbildschirmen […].

In the source language corpus, HOME was not identified as a collocate. There is no lexical entry in the chosen dictionaries for HOME **ENTERTAINMENT**. The collocation HOME **ENTERTAINMENT** seems to be a target language specific one; in other words, it was coined independently of a source language example and is suspicious of being a pseudo-Anglicism.

EVENT

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
202	10	CAMPAIGN	(10)	Kampagne	FILM	(10)	6	120
		BIG	(7)	groß	MANAGER	(8)		
		NEW	(7)	neu	FERNSEHEN	(5)		
		POLITICAL	(5)	politisch	MOVIE	(5)		
		DAY	(5)	Tag	AGENTUR	(5)		

The node **EVENT** does not feature any identical or equivalent lexical collocates in the source and target language corpora. Among the 6 identified lexical target language collocates are 2 Anglicisms: MANAGER and MOVIE. This is exemplified below.

TL 6: In Wolfgang Panzers "Event-Movie" (Buch: Wolfgang Kirchner) fliegen die Kugeln früher, und der großen Gefühle wegen fügte man […].

TL 7: "Manchmal verbirgt sich hinter englischen Begriffen nichts wirklich Neues", sagt Neyer, "Event Manager klingt besser als Veranstaltungsorganisator […].

The most frequent target language collocate of **EVENT** is FILM, which is a translation of the collocate MOVIE. FILM, like MOVIE, is also used frequently in R1 position. One example is provided below.

TL 21: Im besten Falle kostet Sie ein Event-Film kein Geld, sondern Sie erhöhen stattdessen sogar dauerhaft Ihren Umsatz.

Because MOVIE and FILM are both frequent collocates of **EVENT** and they both turn up in the corpus in R1 position, it can be assumed that they are used interchangeably for stylistic variation. In the source language corpus, neither MANAGER nor MOVIE were identified as lexical collocates of **EVENT.**

FAN

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
176	4	BIG (10)	groß	ARTIKEL (16)	4	160
		MAIL (5)	Post	KULTUR (12)		
		HUGE (5)	riesig	GEMEINDE (7)		
		BASE (5)	Basis/ Anhängerschaft	CAMP (7)		
		-	-	-		

The node **FAN** does not share any lexical collocates in the source and target language corpora. One of the 4 identified lexical target language collocates is an Anglicism. CAMP occurs in R1 position in the target language corpus as follows:

TL 39: Es ist brütend heiß unter der schwarzrot-goldenen Zeltplane im deutschen Fan-Camp, und die halbnackte Meute vor dem Kärntner Landeshauptmann brüllt […].

TL 42: Er spendiert den angetrunkenen Deutschen im Fan-Camp auf Wunsch ein Fass Freibier, verabschiedet sich dann in seinen Hubschrauber […].

In the source language corpus, CAMP is not identified as a collocate of **FAN**. Therefore **FAN** CAMP may be a pseudo-Anglicism.

FARM

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
198	14	WIND	(10)	Wind	BODY	(14)	3	116
		ACRE	(9)	Morgen	LAND	(8)		
		✓ ORGANIC	≙ (7)	biologisch/ ökologisch	✓ ÖKO	≙ (5)		
		BILLION	(6)	Milliarde	-			
		ANIMAL	(6)	Tier	-			

The node **FARM** shares one equivalent lexical collocate, namely ORGANIC and ÖKO. **FARM** can also be a verb. In the source language corpus only instantiations in which **FARM** is used as a noun were counted. In the target language corpus, **FARM** is only used as a noun denoting "an area of land and its buildings, used for growing crops and rearing animal" (www.oxforddictionaries.com).

ORGANIC is defined by *Oxford Dictionaries* as "(of food or farming methods) produced or involving production without the use of chemical fertilizers, pesticides, or other artificial chemicals." And it is translated by *PONS* as "aus biologischem Anbau". Thus, equivalence for ORGANIC and ÖKO as collocates of the node **FARM** can be assumed. ORGANIC co-occurs with **FARM** in L2 and L1 position in the source language corpus. In the target language corpus, ÖKO co-occurs with **FARM** in L1 position. Below, examples from both corpora are provided of the co-occurrence of **FARM** and ORGANIC/ÖKO.

SL 100: […] Raoul Adamchak, an organic farmer at the University of California-Davis's certified organic farm and former partner at Full Belly Farm […].

SL 126: […] a 150-acre organic vegetable farm in California.

TL 9: Wenn das Riesenland seine Agrarwirtschaft nach dem Vorbild von Charles' privater Öko-Farm Highgrove ausrichtete […].

A dictionary entry for ORGANIC **FARM** does not exist in the chosen dictionaries. It may be best translated with 'Biobauernhof', but instead ÖKO **FARM** is preferred. 'Biobauernhof' does not occur in the target language corpus. ORGANIC and ÖKO qualify as equivalent lexical collocates of **FARM**.

The top target language collocate of **FARM** is an Anglicism, i.e. BODY. BODY was identified in L1 position of **FARM**. In the chosen mono- and bilingual dictionaries no entry for BODY **FARM** exists. In addition, BODY was not identified as a collocate of **FARM** in

the source language corpus. Judging from the context in which BODY **FARM** occurs in the target language corpus, it refers to a research facility where human decomposition can be studied in a variety of settings. Consider the following examples:

TL 1: Leiter der Body-Farm ist der Forensikprofessor Richard Jantz.

TL 46: […] ein Jahr zuvor nach schwerer Krankheit gestorben - und zu Trainingszwecken für FBI-Agenten auf der Body-Farm eingegraben worden.

Whether BODY **FARM** is a term which is also used in the source language could not clearly be determined. BODY **FARM** is thus suspicious of being a pseudo-Anglicism.

FASHION

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
156	8	✓ WEEK	≡ (13)	Woche	≡ SHOW	6	114
		NEW	(7)	neu	≡ MAGAZINE (8)		
		DESIGNERS	(7)	Designer	BLOG	(6)	
		≡ SHOW	(5)	Schau	GROUP	(6)	
		≡ MAGAZINE	(5)	Magazin	✓ ≡ WEEK	(5)	

According to the quantitative corpus analysis, the node **FASHION** features the exceptional number of three identical lexical collocates in the source and target language corpus. The first identical collocate, WEEK, occurs in both corpora exclusively in R1 position of **FASHION**.

SL 141: […] sent out an e-mail prior to casting for this year's Fashion Week asking designers to be "mindful of diversity."

TL 2: Gleich eine ganze Horde von Fotografen verfolgte Hilson bei ihrem Besuch der Fashion Week in New York.

In the chosen dictionaries, no lexical entry for **FASHION** WEEK exists. Thus, WEEK qualifies as an identical lexical collocate of **FASHION**.

The second identical lexical collocate of **FASHION** is SHOW. In both corpora SHOW co-occurs with **FASHION** exclusively in R1 position. Below, examples from both corpora illustrate this.

SL 104: There's a big difference between a fashion show and the product that a consumer buys.

TL 9: Die Fashion-Show ist gleichzeitig auch Rückentherapie.

According to *PONS*, the translation of **FASHION** SHOW is "Modenschau". SHOW thus does not qualify as a collocate of **FASHION**, but rather forms the compound **FASHION** SHOW with the node.

The third identical lexical collocate of **FASHION**, which was identified by means of the quantitative corpus analysis, is MAGAZINE. In both corpora, MAGAZINE co-occurs with **FASHION** only in R1 position of the collocational span. The examples below illustrate their co-occurrence.

SL 74: […] tech obsessed, glasses-wearing geek, but who's garnered mainstream appeal and a few fashion-magazine covers.

TL 7: Modeanzeigen sind der Kleister des Fashion-Magazine.

FASHION MAGAZINE translates to "Modezeitschrift" and is listed as such in *PONS*. Consequently, MAGAZINE cannot be counted as an identical collocate of **FASHION**.

It should be noted that all Top 5 target language collocates of **FASHION** are Anglicisms. The remaining sixth lexical target language collocate is also an Anglicism, namely AWARD. Remarkably, neither BLOG, nor GROUP or AWARD could be identified as source language collocates of **FASHION**.

FREAK

SOURCE LANGUAGE					TARGET LANGUAGE		
T	LC	Top 5		TRANSLATION	Top 5	LC	T
144	15	CLEAN	(14)	sauber	INTERNET (6)	3	119
		SHOW	(10)	Schau/ Sendung	IMMOBILIEN (6)		
		CONTROL	(8)	Kontrolle	TECHNIK (5)		
		GLORIOUS	(5)	glorreich/ herrlich	-		
		GREEN	(5)	grün/Umwelt-/ ökologisch	-		

The node **FREAK** does not feature any identical or equivalent lexical collocates. The most frequent target language collocate is INTERNET and hence an Anglicism. INTERNET was identified in the target language corpus exclusively in L1 position of the collocational

span. However, INTERNET was not identified as a source language collocate and the combination INTERNET **FREAK** is thus possibly a pseudo-Anglicism.

GANGSTER

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
142	4	LIFESTYLE	(8)	Lebensstil	STADT (8)	3	115
		SMALL-TIME	(5)	unbedeutend/ klein	GHETTO (6)		
		CHASED	(5)	gejagt	FILM (5)		
		STREET	(5)	Straße	-		
		-			-		

The node **GANGSTER** does not feature any identical or equivalent lexical collocates in the source and target language corpora. The target language collocate GHETTO is an Anglicism, but was not identified as a source language corpus collocate.

INTERVIEW

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
398	33	✓ SAID	≙ (63)	sagte	GEBEN (10)	8	195
		EXCLUSIVE	(50)	ausschliesslich/ exklusiv	✓ SAGT ≙ (7)		
		RECENT	(19)	kürzlich	GIBT (7)		
		WEEK	(17)	Woche	MAGAZIN (5)		
		NEWS	(13)	Neuigkeiten	JOURNALISTEN (5)		

The node **INTERVIEW** features one equivalent lexical collocate, i.e. the verb SAID and SAGT. It should be noted that SAYS and SAGTE are also among the significant lexical collocates in the source and target language corpus. In the source language corpus, SAID co-occurs with **INTERVIEW** in L4, L3, and all positions of the span to the right of the node. The examples from the source language corpus below exemplify this.

SL 91: Revuelta said in an interview that she believed the prison director arranged the switch.

SL 107: At the end of the interview the woman said, ‘My goodness, I didn’t know I felt that way.’

SL 180: Davis, in the interview, said he's hopeful that […].

SL 272: In the interview, Dean said Obama has demonstrated that he is ready to be president and that […].

SL 285: […] but he was also motivated by the cause, his lawyers said in a phone interview from Iraq.

SL 378: Even so, in a 2005 interview with NEWSWEEK, Khalilzad said that one thing he had learned during his term in Afghanistan […].

In the target language corpus, SAGT occurs in L4, L3, R1 and R4 position of the collocational span. Examples are provided below.

TL 44: In einem Interview zu dem Film sagt er, seine Frau Jenny Gröllmann habe ohne sein Wissen […].

TL 61: "Wir wachsen sehr, sehr schnell", sagt Joseph beim Interview in einem New Yorker Hotel und strahlt.

TL 132: […] sagt nun Erstaunliches im Interview mit dem US-Magazin "Newsweek": "Ich bin zum ersten Mal optimistisch, was den Irak angeht […].

TL 147: In einem Interview sagt er: "Ich möchte unser Geschäft mit Hypothekenkrediten ausbauen.

SAID and SAGT qualify as equivalent lexical collocates of **INTERVIEW**. This is the first verb which qualifies as either an identical or equivalent collocate of a node. At first it might seem surprising to count SAID and SAGT as equivalent collocates, since they are past and present forms of the verbs ‘say’ and ‘sagen’. So far, the present study did not count morphological variants (such as plural forms) as identical or equivalent collocates. Thus, it should be pointed out that SAID and SAGT can be counted as equivalent collocates in this particular case because SAYS and SAGTE also qualify as frequent lexical collocates of **INTERVIEW**. Neither SAYS, nor SAGTE made the Top 5, therefore they are not shown in the chart above. SAYS co-occurs with **INTERVIEW** 11 times and SAGTE co-occurs with **INTERVIEW** 5 times in the target language corpus.

JOB

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
718	68	FIRST	(28)	erste	SUCHEN	(29)	31	443
		GOOD	(28)	gut	✓ BEKOMMEN	≙ (15)		
		✓ GET	≙ (26)	bekommen	✓ NEUEN	≙ (15)		
		✓ NEW	≙ (18)	neu	VERLOREN	(9)		
		DAY	(16)	Tag	ABBAU	(8)		

The node **JOB** features two equivalent lexical collocates in the source and target language corpora. One of them is a verb and the other one an adjective. GET is identified in the source language corpus left of the node JOB in L4, L3 and most frequently L2 position. For examples, see below.

SL 369: How can you get a work-study job?

SL 445: There will be days when you will wonder why you worked so hard to get this job.

SL 701: They know they can't get the top job, but on the other hand they do know that […].

In one instantiation GET was identified right of the node **JOB** in R2 position. Although node and collocate are stretched across sentence boundaries, they are logically related and form a similar pattern as illustrated in the examples above. SL 781 can therefore also be counted as a collocation of **JOB** and GET (cf. subsection 2.2.1.2 and particularly section 2.4). Consider the following example from the news magazine source language corpus.

SL 781: But McCain wanted one more shot at the top job. To get it, he tacked right during the GOP primaries this year.

In the target language corpus, BEKOMMEN was identified in R3 and most frequently in R2 position of the collocational span of **JOB**. Corresponding examples are given below.

TL 59: […] hat in der Alpenrepublik so großen Erfolg, dass sie einen lukrativen Job als Werbemodell bekommen hat.

TL 406: […] die 16-jährige Jeany hingegen dreht ein Bewerbungsvideo, um einen Job zu bekommen, trotz ihres verrufenen Wohnorts.

The collocates GET and BEKOMMEN qualify as equivalent lexical collocates of the node **JOB**.

The second identified equivalent lexical collocate of **JOB** is NEW and NEUEN. As has been stated before, the corpus analysis and collocation in general are word-form based. However, if an equivalent collocate of a source language adjective collocate is encountered, the different adjective endings in German need to be considered. This is necessary because NEUEN is just as much a (translation) equivalent of NEW as is NEU. In the chart above only the most frequent adjective form is given.

In the source language corpus, NEW was identified in L2 position of the collocational span of **JOB**, but most frequently in L1 positions. Below, examples of the co-occurrence of NEW and **JOB** in the source language corpus are provided.

SL 549: […] on Capitol Hill on the senator's first day in his new job.

SL 733: […] says that she was the most qualified person for the new principal job.

In a couple of instantiations NEW was also identified as a collocate of **JOB** in positions right of the node in the source language corpus. However, in none of these instantiations NEW could be counted as a collocate. Although NEW and **JOB** co-occurred within the same sentence and were identified as node and collocate by means of the quantitative corpus analysis, they were not logically related (cf. section 2.4). Rather NEW related to another constituent in the sentence. Consider the following corpus examples.

SL 524: […] he has positioned himself well for a job in the new administration.

SL 614: My job is to bring new energy and new technologies [to the] strength of the existing organization […].

In the target language corpus, NEUEN co-occurs with JOB exclusively in L1 position. Below a few examples of this are given to illustrate the co-occurrence of NEUEN and **JOB** as collocate and node.

TL 19: Dort hat er seinen neuen Job inzwischen verloren, nachdem seine Vorgeschichte bekannt wurde.

TL 125: So hatte sich Manfred Balz seinen neuen Job wohl nicht vorgestellt.

TL 229: Wie gehen Sie an Ihren neuen Job heran?

NEW and NEUEN qualify as equivalent lexical collocates of **JOB** according to the quantitative and qualitative corpus analysis.

KILLER

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
144	8	SERIAL	(10)	Serien-	PROFESSIONELL	(13)	4	180
		JOB	(8)	Stelle	PERFIDE	(9)		
		LEADING	(8)	führend	OPFER	(6)		
		INSTINCT	(6)	Instinkt	FOTOS	(6)		
		DETERMINED	(5)	entschlossen	-			

The node **KILLER** does not feature any identical or equivalent lexical collocate in the source and target language corpora. **KILLER** is one of the few nodes which is more frequent in the target language corpus than in the source language corpus. In line with earlier results for nodes of this kind, **KILLER** exclusively collocates with target language word-forms in the target language corpus. No Anglicisms were identified as target language collocates.

KNOW-HOW

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
118	5	✓ TECHNICAL	≙ (11)	technisch	✓ TECHNISCHES	≙ (8)	4	104
		PROVIDING	(6)	zur Verfügung stellen	UMWELT	(7)		
		BUSINESS	(5)	Geschäft/ Handel	INTERNATIONALEN	(5)		
		SCIENTIFIC	(5)	Wissenschaftlich	ERWERB	(5)		
		LACK	(5)	Mangel	-			

The node **KNOW-HOW** shares one equivalent lexical collocate in the source and target language news magazine corpora, i.e. TECHNICAL and TECHNISCHES. The identified equivalent collocate is an adjective, thus inflectional variants of the German adjective were also considered because they all are translation equivalents of TECHNICAL. In the chart above, only the most frequent target language adjective form is provided.

TECHNICAL and TECHNISCHES were identified exclusively in L1 position of **KNOW-HOW** in the source and target language corpus respectively. The examples below illustrate

this. In addition, one example of an inflectional target language variant of TECHNISCHES as a collocate of **KNOW-HOW** is also provided in TL 13.

SL 6: For about $80 at any electronics store and some technical know-how, it's possible to tap into […].

TL 13: Die Tinners verpflichteten sich darin, das "technische Know-how" im "Bereich der Vakuumtechnik und des Ventildesigns […].

TL 24: Der Schah habe gehofft, sich so technisches Know-how für den Bau einer Atombombe einkaufen zu können.

The collocates TECHNICAL and TECHNISCHES qualify as equivalent lexical collocates of **KNOW-HOW**.

It should be mentioned in passing that although **KNOW-HOW** is more frequent in the source language corpus than in the target language corpus, it does not feature any Anglicisms as target language collocates.

LIFE

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
1,652	117	REAL	(60)	wirklich/real	SWISS	(7)	5	160
		NEW	(52)	Neu	≡ WAY	(6)		
		PUBLIC	(51)	öffentlich/ Öffentlichkeit	BRIGHT	(6)		
		✓ ≡ AMERICAN	(44)	amerikanisch	✓ ≡ AMERICAN	(5)		
		≡ WAY	(34)	Weg/-weise	SIDE	(5)		

The node **LIFE** shares two identical lexical collocates in the source and target language corpus. In the source language corpus, AMERICAN collocates with **LIFE** in L3, L1 and R3 position of the span. Most frequently, 30 times, AMERICAN co-occurs with **LIFE** in L1 position. Examples of this are given below.

SL 298: […] wanted to protect the American way of life, which requires access to massive amounts of oil.

SL 311: Unfortunately, most of those fans confuse everyday American life with that justice system.

SL 986: With Obama's election, he breathes life into the American Dream by embodying […].

In the target language corpus, AMERICAN co-occurs with **LIFE** exclusively in L3 position. A few examples of this are provided below.

TL 3: […] in der verwegenen Hoffnung, der ganzen Region den American way of life zu eröffnen.

TL 42: Der "American Way of Life", diese besondere Mischung aus vorsätzlicher Sorglosigkeit, lustvoller Verschwendung und einem Schuss Größenwahn.

TL 44: Moskau kann bis heute nicht verstehen, dass der American Way of Life trotz des Irak-Debakels in Osteuropa attraktiver sein soll als sein Modell der "gelenkten Demokratie" […].

In all instantiations in which AMERICAN was identified as a collocate of **LIFE**, collocate and node were constituents of the expression 'American way of life'. It needs to be determined whether AMERICAN qualifies as a collocate of **LIFE**, to be precise whether the expression 'American way of life' qualifies as a collocation, or different type of phraseological unit. The lexical analysis software cannot tell the difference between phraseological units such as idioms and collocations (cf. subsection 2.2.2). This illustrates another reason why qualitative, or manual, corpus analysis is indispensable.

Many of the criteria which have contributed to the formation of the concept of collocation used in this study apply equally to idioms. As has been defined in subsection 2.3.5 above, semantic transparency is a scalar criterion, but it is the criterion to tell apart idioms and collocations. The construction 'American way of life' is semantically transparent. According to the three criteria which were introduced on page 105 f. to define compounds, 'way of life' is qualified as such and defined by *Oxford Dictionaries* as "the typical pattern of behaviour of a person or group." Consequently, AMERICAN WAY OF **LIFE** is the typical pattern of behavior of Americans. It is also possible to call it a 'multi-word lexeme'.

The present study argues accordingly that 'American way of life' qualifies as a complex collocation. This is in line with the concept of collocation used in this study, because collocation as it is understood in this study is the combination of two lexical items which can also be compounds (cf. section 2.4). Thus, AMERICAN qualifies as an identical lexical collocate of **LIFE**. Node and collocate are borrowed as two constituents of the complex collocation (cf. subsection 2.3.1) AMERICAN WAY OF **LIFE**.

The second identical lexical collocate WAY was in most cases identified in the target language corpus as a constituent of the complex collocation AMERICAN WAY OF **LIFE.** However, there are also a few other examples. Still, WAY and **LIFE** are only identified as constituents of WAY OF **LIFE**. Consider the following example.

TL 44: […] jüdischen Vasallenkönige waren gerade dabei, die rückständige Gegend mit einem neuen Way of Life zu beglücken.

The discussion above should have illustrated why WAY cannot be counted as an identical lexical collocate of LIFE. It has been pointed out that WAY of LIFE is a compound, thus WAY and LIFE are constituents of this compound.

LUNCH

SOURCE LANGUAGE				TARGET LANGUAGE		
T	**LC**	**Top 5**	**TRANSLATION**	**Top 5**	**LC**	**T**
121	7	HAVE (9)	haben	THEMEN (7)	4	110
		SCHOOL (7)	Schule	SANDWICH (6)		
		FREE (7)	frei/umsonst	BUSINESS (5)		
		BREAKFAST (6)	Frühstück	COCKTAIL (5)		
		DAY (5)	Tag	-		

The node **LUNCH** does not feature any identical or equivalent lexical collocates in the source and target language corpora. Interestingly, 3 of the 4 identified lexical target language corpus collocates are Anglicisms. It should be noted that in the source language corpus neither SANDWICH, nor BUSINESS, or COCKTAIL were identified as collocates.

MEETING

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5	≡		Top 5	LC	T
497	46	TOWNHALL	(27)	Rathaus	TOWN-HALL	(8)	4	115
		SAID	(14)	sagte	BUSINESS	(8)		
		LEADERS	(12)	Leiter/in	MARATHON	(6)		
		PRESIDENT	(11)	Präsident/in Vorsteher/in	PRAYER	(5)		
		ANNUAL	(11)	jährlich	-			

The node **MEETING** shares one identical lexical collocate in the source and target language corpora. It should be noted that in most instantiations TOWNHALL is spelled without a hyphen in the source language corpus. In the target language corpus, TOWN-HALL is mostly spelled with a hyphen. In both corpora, TOWN HALL also occurs as two words separated by a blank space. *Merriam Webster* defines TOWN HALL as "a public building used for town-government offices and meetings." Only the most frequently occurring spelling variant is given in the chart above.

The identical lexical collocate TOWNHALL occurs in both corpora exclusively in L1 position of the node **MEETING**. The examples below illustrate this.

SL 463: When he takes questions during a town-hall meeting, he has been instructed to approach the audience so they're in the shot.

TL 2: Er hat zu einer seiner Bürgerversammlungen eingeladen, die er Town Hall Meeting nennt […].

In addition, TOWN HALL **MEETING** has a lexical entry in *PONS* and is translated with "Bürgerversammlung". TOWNHALL does thus not qualify as a lexical collocate of **MEETING**, but is one constituent of the compound TOWN HALL **MEETING**.

It should be noted that out of the 4 identified lexical target language collocates of **MEETING**, 2 more are Anglicisms. The target language collocates BUSINESS and PRAYER were not identified as source language corpus collocates. To illustrate their use in the target language corpus, examples are provided below.

TL 4: Je nach Anlass peppt sie ihn mit ein paar Molekülen auf: fürs Business-Meeting mit Hedione, dem Riechstoff für haftenden Erfolg.

TL 5: Ein allabendlicher Gottesdienst soll Berufstätige nach Büroschluss zu einem "prayer meeting" anspornen.

The combinations BUSINESS **MEETING** and PRAYER **MEETING** do not have a lexical entry in the chosen mono- and bilingual dictionaries. As has been pointed out previously, such combinations are referred to as 'Anglo-neologisms' (cf. Steffens 2003: 5). The second constituent provides the general meaning and the first constituent makes this meaning more specific. Usually the combinations are semantically transparent. The imitation of source language structures in the target language leads to the coining of so-called 'Anglo neologisms'. Ultimately they usually prove to be pseudo-Anglicisms.

MODEL

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
331	28	ROLE	(21)	Rolle	≡ NEXT	11	159
		BUSINESS	(17)	Handel/Geschäft	SHOW	(15)	
		NEW	(8)	neu	CAMP	(10)	
		RANKING	(8)	Einstufung/ Rangfolge	AGENTUR	(10)	
		≡ NEXT	(6)	nächste	KÖRPER	(8)	

The node **MODEL** features one identical lexical collocate, i.e. NEXT. The word-form **MODEL** is also used as a verb in the source language corpus. Only instantiations in which **MODEL** was used as a noun in the source language corpus were counted for the corpus analysis. In the target language corpus, **MODEL** only occurs as a noun. In the source language corpus, NEXT co-occurs with **MODEL** exclusively in L2 position. In the target language corpus this was also observed. Consider the following examples from both corpora.

SL 183: […] with a pinch of "America's Next Top Model." The show throws 14 models, both men and women, into a house […].

SL 212: I don't think I'm a mogul, but I have a lot of television shows. There's "America's Next Top Model" and the talk show.

TL 11: Sie habe ihr gesamtes Privatvermögen in "Nigeria's Next Super Model" investiert.

TL 29: […] stach sieben Mitbewerberinnen bei der Show aus, die wie "Germany's Next Top Model" funktionierte.

As the above examples illustrate, NEXT and **MODEL** only co-occur as constituents of television show titles which are proper names. Consequently, NEXT is not counted as an identical lexical collocate of **MODEL**.

There are 2 further Anglicisms among the identified Top 5 lexical target language collocates. The target language collocates SHOW and CAMP were exclusively identified in R1 position of the node **MODEL**. The examples below show this.

TL 4 "Klar, wirkt eine Model-Show da überflüssig", sagt Joan.

TL 6: Sie mietete ein Ferien-Resort in Lagos und nannte es Model-Camp: 5.30 Uhr aufstehen, sich schön machen, dann Morgengebet.

Neither CAMP nor SHOW was identified as a source language corpus collocate of the node **MODEL**. Again, this may indicate the formation of 'Anglo neologisms' and possibly over time 'pseudo-Anglicisms'.

NETWORK

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
279	22	TELEVISION	(14)	Fernsehen	≡ SOCIAL (8)	3	119
		NEWS	(13)	Neuigkeiten	STUNDEN (5)		
		TERRORIST	(8)	Terrorist	KONZEPT (5)		
		HOME	(7)	Haus/Heim	-		
		≡ SOCIAL	(7)	sozial/ gesellschaftlich	-		

Similarly to **MODEL**, the node **NETWORK** can also be used as a verb in the source language. Again, only instantiations in which **NETWORK** is used as a noun are counted for the corpus analysis. The node **NETWORK** features one identical lexical collocate according to the quantitative corpus analysis, namely SOCIAL. In the source language corpus, SOCIAL is identified in L1 and R2 position of **NETWORK**. Below this is exemplified.

SL 131: There's nothing new about Scrabble — which dates to the 1930s — nor was Facebook the first social network.

SL 196: In others, the connection is indirect, as when being disagreeable leaves you with no social network, which can cause stress and thus poor health.

SL 216: Hizbullah runs an impressive network of social services, which provide health care, small loans […].

When looking at the corpus examples SL 131 and SL 196 above, it becomes clear that SOCIAL **NETWORK** is used with two different meanings. *Oxford Dictionaries* defines the two meanings of SOCIAL **NETWORK** as follows. Originally, SOCIAL **NETWORK** refers to a "network of social interactions and personal relationships." More recently, it also refers to "a dedicated website or other application which enables users to communicate with each other by posting information, comments, messages, images, etc." It can be seen from the source language corpus examples that SOCIAL **NETWORK** is used with both of these meanings.

In the target language corpus, SOCIAL exclusively co-occurs with **NETWORK** in L1 position. SOCIAL **NETWORK** is always used in its more recent sense in the target language corpus. Consider the following examples.

TL 36: Wir waren acht und sollten dem Venture-Capital-gestopften Gründer eines neuen Social Network sagen, dass sein Launch-Konzept gut ist.

TL 71: Das *Social Network* soll wegen technischer Mängel das Erstellen umfangreicher Nutzerlisten mit Telefonnummern ermöglicht haben.

TL 115: Anlässlich des Welt-Telekommunikationstages hat sich das Statistische Bundesamt (Destatis) einmal die Social Network-Nutzung in Deutschland angesehen.

The translation of SOCIAL **NETWORK** is either "soziales Netzwerk" or "soziale Kontakte" (www.pons.eu). In German, 'soziales Netzwerk' is also used with both of the meanings provided by *Oxford Dictionaries* above. In the target language corpus each of the translations 'soziales Netzwerk' and 'soziale Kontakte' occurs exactly once. The first translation refers to a newly launched website and the latter to a network of social interactions and personal relationships. Consider the corresponding examples.

TL 1: Sie stellen gerade Ihr soziales Netzwerk Bloomstreet ein. Kann Bertelsmann einfach kein Internet?

TL 1: Diese graubraunen Primaten sind von Natur aus auf soziale Kontakte gepolt.

Basically, this shows that the use of SOCIAL **NETWORK** by far outnumbers that of its target language equivalents. In addition, SOCIAL **NETWORK** is borrowed in the target language only to denote recent web-based services.

It has been mentioned earlier (cf. subsection 5.1.4) that Anglicisms are found especially in those fields which bring or brought technological innovations into everyday life. The frequent use of SOCIAL **NETWORK** is a typical example of this and shows the cultural domination of the United States in this area. The development of social network sites began in 1997 with the launch of SixDegrees.com (cf. http://jcmc.indiana.edu) by a company based in New York City. It can be stated that Fink's first motive for borrowing Anglicisms, which was introduced in subsection 5.1.4 of the present study, does not apply: There is no lack of a German expression to translate SOCIAL **NETWORK** adequately. Moreover the existing target language expression 'soziales Netzwerk' experiences a shift in semantic value and now preferably refers to social network websites.

The quantitatively identified lexical collocate SOCIAL does not qualify as such according to the qualitative corpus analysis because SOCIAL and **NETWORK** form a compound according to the criteria which were introduced on pages 105 f. of the present study.

NEWS

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
629	48	ARTICLES	(54)	Artikel	BREAKING	4	107
		WORLD	(47)	Welt/Erde	≡ STORIES		
		NATIONAL	(43)	national	≡ MEDIA		
		MEDIA	≡ (39)	Medien	CHALLENGE		
		STORIES	≡ (33)	Geschichten	-		

The node **NEWS** features two identical lexical collocates in the source and target language corpora. The first identical collocate, according to the quantitative corpus analysis, is MEDIA. In the source language corpus, MEDIA co-occurs with **NEWS** in L4, R3 and most frequently in R1 position of the collocational span. Examples from the source language corpus are provided below.

SL 244: During the past year, veterans' issues were all over the media — and often the news was grim.

SL 450: The traditional news media, usually near the bottom of popularity contests, ranked higher on the […].

SL 559: They also knew that the news media would give it extensive coverage, which they did.

SL 603: The endorsement is making big news and diverting media attention from the messages that […].

In the target language corpus, MEDIA is only identified in R1 position of **NEWS**. Thus, **NEWS** MEDIA might be a compound. Below examples from the target language corpus are given.

TL 19: Die Krise im Journalismus, so der aktuelle Report "The State of the News Media" des Project for Excellence in Journalism, liege in einer […].

TL 27: […] war von der Newspaper Association of America 2001 sogar zum "News Media Pioneer" ernannt worden.

TL 56: NEWS: Media Markt über die Empörung, die ein TV-Spot mit Olli Dittrich als Italiener […].

None of the chosen mono- or bilingual dictionaries list **NEWS** MEDIA. But the target language examples illustrate that if **NEWS** and MEDIA co-occur, at least one of the constituents forms part of a proper name; in TL 19 this concerns the title of a report, and in TL 27 the name of an award. In TL 56, MEDIA forms part of a company name and **NEWS** is the section headline. Thus, MEDIA does not qualify as an identical lexical collocate of **NEWS**.

The other identical lexical collocate of **NEWS** is STORIES. **NEWS** STORY is translated by *PONS* with either “Bericht” or the Anglicism “Story”. As a result, **NEWS** STORIES qualifies as a compound rather than a collocation.

It should be noted that all 4 identified lexical target language collocates of **NEWS** are Anglicisms. Furthermore, CHALLENGE was not identified as a source language collocate, whereas BREAKING co-occurred with **NEWS** 6 times in the source language corpus. An example is provided below.

SL 381: […] attendees were stunned by breaking news: New York Mayor Mike Bloomberg had just announced […].

TL 42: "Zutiefst beschämt" sei der Papst wegen der pädophilen Priester, meldete der CNN-Sonderkorrespondent, Kategorie: Breaking News.

BREAKING does not qualify as a collocate of **NEWS** because BREAKING **NEWS** is a compound and translated by *PONS* with "Eilmeldung". This translation only occurs three times in the target language corpus. For an example see below.

TL 2: CNN hat eine Eilmeldung geschickt, sie landet auf den Blackberrys der Gäste.

BREAKING **NEWS** and 'Eilmeldung' are used interchangeably in the target language corpus, probably for stylistic variation.

PERFORMANCE

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
335	27	ENHANCING	(14)	steigernd/ verbessernd	DERARTIGE	(10)	4	109
		DEBATE	(13)	Debatte/ Diskussion	KÜNSTLERIN	(8)		
		DRUGS	(13)	Drogen	LIVE	(7)		
		JOB	(12)	Stelle	PERSÖNLICHE	(7)		
		PRESIDENT	(8)	Präsident/ Vorsitzender	-			

The node **PERFORMANCE** does not feature any identical or equivalent lexical collocates in the source and target language corpora. It needs to be mentioned that the node **PERFORMANCE** has two different meanings. *Oxford Dictionaries* defines them as follows. Firstly, **PERFORMANCE** may refer to "an act of presenting a play, concert, or other form of entertainment." See the examples from both corpora below.

SL 25: A fresh look at the Beatles's farewell performance can transport you to the days when musicianship still mattered […].

TL 6: Sie nennt sich Alma Navailles, Performance-Künstlerin und Tänzerin, ihr Programm liegt irgendwo zwischen Flamenco und Salsa.

Secondly, **PERFORMANCE** may denote "a task or operation seen in terms of how successfully it is performed." Consider the following examples.

SL 235: By now, many people are familiar with America's poor academic performance on the international stage.

TL 1: Das ist der Markt und die sich abzeichnende Rezession. Wenn wir die Einzigen wären mit einer derartigen Performance, würde ich Ihnen recht geben.

Both meanings are used in both corpora. All the more, it is surprising that no shared collocates could be identified in the corpora.

Yet, one of the identified lexical target language corpus collocates is the adjective LIVE. *Merriam Webster* defines LIVE as "involving a presentation (as a play or concert) in which both the performers and an audience are physically present." An example from the target language is given below.

TL 8: […] trotzdem bewerben sie sich aus gutem Grund: weil sie sich für Live-Performance begeistern.

LIVE was not identified as a source language corpus collocate.

REALITY

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
358	21	CHECK	≡ (31)	Überprüfung/ Kontrolle	SHOW	≡ (15)	8	132
		NEW	(14)	neu	SOAP	(8)		
		SHOW	≡ (11)	Schau/ Sendung	SERIE	(7)		
		SHOWS	(11)	Schauen/ Sendung	TV	(7)		
		POLITICAL	(8)	politisch	CHECK	≡ (6)		

The node **REALITY** features two identical lexical collocates in both corpora. The collocate CHECK is exclusively identified in R1 position in the source language corpus. Consider the following examples.

SL 60: […] likely to give way to the downbeat of a reality check soon after Inauguration Day.

TL 20: Schwarze dürfen nicht frei sein? Reality check!

TL 21: Frauen dürfen nicht wählen? Reality check!

TL 22: Der Mond ist unerreichbar weit weg? Reality check!

REALITY CHECK is defined by *Oxford Dictionaries* as "an occasion on which one is reminded of the state of things in the real world." *PONS* translates **REALITY** CHECK as "Augenöffner". Consequently, CHECK does not qualify as an identical lexical collocate. However, the compound **REALITY** CHECK is frequently borrowed as one lexical item.

SHOW is the other identical lexical collocate of **REALITY**. In the target language corpus SHOW only occurs as a noun. Therefore all instantiations of SHOW can be counted for the corpus analysis. In the source language corpus, SHOW is identified in R2 and most frequently R1 position of the collocational span of **REALITY**. In the target language corpus, SHOW only co-occurs with **REALITY** in R1 position. Corresponding corpus examples are given below.

SL 118: In a new reality show called "The Secret Life of a Soccer Mom," the network is offering […].

SL 223: […] is now the name of a reality-TV show that bears little resemblance to the book […].

TL 7: Frankreichs Präsident inszeniert Politik als Reality-Show - und will nun das Staatsfernsehen kontrollieren.

The compound **REALITY** SHOW is translated by *PONS* with the Anglicism "Realityshow". SHOW does thus not qualify as a collocate of **REALITY** according to the present study's guidelines.

Among the Top 5 target language corpus collocates is one more Anglicism, i.e. SOAP. In the target language corpus, SOAP co-occurs with reality exclusively in R1 position. SOAP can be used as an informal short form for soap opera according to *Oxford Dictionaries*. In addition, SOAP means "a substance used with water for washing and cleaning". 'Soap opera' is translated by *PONS* as "Seifenoper". Thus, the combination of **REALITY** and SOAP is semantically transparent. However, SOAP is not identified as a collocate in the source language corpus. **REALITY** SOAP seems thus to be a target language-specific combination and can be classified as an 'Anglo neologism'.

SOCIETY

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
459	36	AMERICAN	(60)	amerikanisch	HIGH	(21)	7	129
		CIVIL	(18)	zivil/ bürgerlich	ROYAL	(15)		
		CANCER	(15)	Krebs	RÜCKKEHR	(8)		
		GREAT	(14)	groß/riesig	TREFFEN	(7)		
		HAS	(11)	hat	KONTAKTE	(7)		

The node **SOCIETY** does not feature any identical or equivalent lexical collocates in the source and target language corpora. It should be noted that 2 of the identified Top 5 target language collocates are Anglicism. HIGH and ROYAL both co-occur with **SOCIETY** in L1 position in the target language corpus. HIGH was not identified as a source language corpus collocate, but still preceded **SOCIETY** once. Consider the following examples.

SL 212: […] reinventing herself as a career woman, gradually shedding the persona of a high-society divorcée for that of a serious professional writer.

TL 4: Doch die schwarze High Society ging in ihrem Haus ein und aus.

HIGH **SOCIETY** is defined by *Oxford Dictionaries* as "the aggregate of people who are fashionable, wealthy, and influential, regarded as forming a distinct group in a community." It is translated by *PONS* solely with the Anglicism "High Society".

ROYAL also co-occurs with **SOCIETY** in the source language corpus, but not frequently enough to be counted as a source language collocate. Nevertheless, for illustration purposes examples from the source and target language corpora are provided below.

SL 375: […] they found a strong correlation, they will report in Proceedings of the Royal Society.

TL 16: Die Royal Society in London, eine Art britische Akademie der Wissenschaften, monierte unlängst […].

Even if ROYAL had co-occurred with **SOCIETY** more frequently in the source language corpus, it would not have been counted as an identical lexical collocate. ROYAL **SOCIETY** is defined by *Oxford Dictionaries* as "the oldest and most prestigious scientific society in Britain." It is translated by *PONS* with "die Königlich Britische Akademie der Naturwissenschaften", thus the proper name of an organization.

SOUND

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
307	23	LIKE	(17)	wie	SAGENHAFTE	(11)	9	171
		BITES	(14)	Bissen/Happen	BEISPIELLOSE	(9)		
		MAKE	(11)	machen/ herstellen	NEUARTIGE	(8)		
		FAMILIAR	(8)	vertraut/bekannt	TYPISCHE	(7)		
		POLICY	(6)	Grundsatz	KRÄFTIGER	(6)		

The node **SOUND** does not fea ture any identical or equivalent lexical collocates in the source and target language corpora. In the target language corpus, **SOUND** collocates exclusively with target language word-forms. This is rare and usually observed with nodes which are more frequent in the target language corpus than in the source language corpus. The fact that **SOUND** is a source language adjective, adverb, verb and noun in identical spelling was considered for the quantitative corpus analysis.

STAR

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
279	22	ROCK	(28)	Rockmusik	LONE	(32)	13	256
		MOVIE	(14)	(Kino)film	✓ NEUEN	≙ (10)		
		✓ FORMER	≙ (8)	ehemalig/ einstig	✓ EHEMALIGER EINSTIGER	≙ (8) (5)		
		RISING	(8)	aufstrebend	GEWORDEN	(6)		
		✓ NEW	≙ (8)	neu	DEUTSCHEN	(6)		

The node **STAR** features two equivalent lexical collocates in the source and target language corpora. In the target language corpus, **STAR** only occurs as a noun. Consequently, only instantiations in which **STAR** occurs as noun in the source language corpus were counted for the analysis. The identified equivalent collocates are all adjectives, thus inflectional variants of the German adjectives were also counted because they all are translation equivalents of FORMER and NEW. The first equivalent lexical collocate of **STAR** is FORMER and EHEMALIGER/EINSTIGER. EHEMALIGER and EINSTIGER both have enough instantiations in the target language corpus to be counted individually as significant

lexical collocates of **STAR**. Because EHEMALIGER and EINSTIGER both are frequent translations of FORMER they are included.

In the source language corpus, FORMER co-occurs with **STAR** in L3, L1, R2 and most frequently L2 position of the collocational span. In the target language corpus, EHEMALIGER and EINSTIGER are exclusively identified in L2 position. The examples below illustrate this.

SL 114: If you can shoot a ball into a basket while being tackled by a former star of gay erotica, to you go the spoils.

SL 126: […] the former college-football star was clearly conflicted about what he'd just proposed.

SL 246: Judge Jackie Glass, sentencing former football star O. J. Simpson to at least 15 years in prison […].

SL 278: Way back when (last summer), the Republican establishment was looking at TV star and former Sen. Fred Thompson as the guy who might […].

TL 87: […] davon so wenig verschont wie ein ehemaliger Baseball-Star, der seit acht Jahren im Wachkoma liegt.

TL 103: Tony Curtis, 83, einstiger Hollywood-Star ("Manche mögen's heiß"), rechnet in seinen Erinnerungen […].

In sum, FORMER and EINSTIGER/EHEMALIGER qualify as equivalent lexical collocates of the node **STAR**.

The other equivalent lexical collocate of **STAR** is NEW and NEUEN. In the source language corpus, NEW co-occurs with STAR in L1 and R3 position. This is exemplified below.

SL 179: Any question that Sarah Palin has become the most unexpected new star of the Republican Party […].

SL 248: […] it's probably because he doesn't really exist: he's the star of a new advertising campaign from the California Milk Processor Board.

In the target language corpus, NEUEN co-occurs with **STAR** in L1 and R2 position of the collocational span. The inflectional variant NEUER co-occurs with **STAR** in L2 and L1 position. Corpus examples are provided below.

TL 93: Und Gary, der Literat und Diplomat, sollte einen Star dieser neuen Welle heiraten […].

TL 210: […] mit ihrer Strahlkraft und mit ihrer Biografie zu einem neuen Star auf der Weltbühne wird, zur Figur der Weltgeschichte.

TL 42: […] gilt als neuer Star im Internet.

TL 172: […] Frankreichs neuer Schauspiel-Star Cécile de France.

It can be stated that NEW, NEUEN, and its inflectional variants qualify as equivalent lexical collocates of **STAR**. Although the corpus analysis is word-form based, all inflectional variants of the adjective NEU qualify as equivalent collocates, because they are all translation equivalents of NEW.

Among the Top 5 target language collocates is an additional Anglicism, i.e. LONE. This collocate occurs exclusively in L1 position of **STAR**, forming the name of the American investment company LONE **STAR**.

STORY

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE		
T	LC	Top 5			Top 5	LC	T
1,162	114	TOLD	(41)	erzählte	SENSATIONELL (18)	12	199
		LOVE	(28) ≡	Liebe	SHORT (16)		
		LIFE	(27)	Leben	LOVE (11) ≡		
		SUCCESS	(20)	Erfolg	PATCHWORK (8)		
		DIFFERENT	(15)	anders	WEB (8)		

The node **STORY** shares one identical lexical collocate, namely LOVE. The collocate LOVE is identified in the source language corpus in L1, R2 and R3 position. In the target language corpus, LOVE co-occurs with **STORY** exclusively in L1 position. Corresponding examples from both corpora are provided below.

SL 572: […] a story of forbidden love between an Orthodox Jewish woman and […].

SL 601: His latest work, "The Reserve," is a story of love and class set in upstate New York.

SL 709: She decided to make a feature-length love story.

TL 53: Das Buch ist die rührselig aufbereitete Love-Story eines ehemaligen Mannequins, das im Alter von […].

The compound LOVE **STORY** is translated by *PONS* with "Liebesgeschichte". Therefore LOVE does not qualify as a lexical collocate of **STORY** according to the qualitative corpus analysis.

It should be noted that the node **STORY** features 4 Anglicisms as target language collocates among the Top 5. Out of the remaining 3, i.e. SHORT, PATCHWORK and WEB, only SHORT was identified as a source language collocate. SHORT occurs in both corpora only in L1 position.

SL 513: […] it's easier to make a movie out of a short story than a novel, because then the question becomes how to enrich […].

TL 64: […] ist eigentlich eher ein Mann der kleineren Form, des Gedichts, der Short Story oder der Novelle, bei ihm verbirgt sich die Bedeutung […].

Even if SHORT had qualified as a Top 5 collocate, it could not have been counted as an identical lexical collocate. SHORT **STORY** is a compound and translated by *PONS* as "Kurzgeschichte".

TALK

SOURCE LANGUAGE				TARGET LANGUAGE		
T	LC	Top 5	TRANSLATION	Top 5	LC	T
1,052	108	✓ STRAIGHT ≡ (61)	gerade/offen/ ehrlich	✓ STRAIGHT ≡ (19)	9	198
		SHOW ≡ (47)	Schau/ Sendung	SMALL (11)		
		PEOPLE (45)	Leute/Menschen	EXPRESS ≡ (11)		
		RADIO ≡ (33)	Radio	SHOW ≡ (8)		
		EXPRESS ≡ (29)	ausdrücken/ aussprechen	RADIO ≡ (8)		

The node **TALK** features the exceptional number of four identical lexical collocates. Only instantiations in which **TALK** is used as a noun are counted for the corpus analysis. The first identical lexical collocate is STRAIGHT. STRAIGHT was identified in the source language corpus most frequently in L1 position of the collocational span. Consider the following example.

SL 517: […] McCain kicked off his general election campaign Thursday night with the kind of straight talk that he has been known for throughout his career.

In the target language corpus, STRAIGHT also only co-occurs with **TALK** in L1 position. Below, examples are given.

TL 25: Sie haben eine ehrliche Ansprache bekommen, ich nenne es "straight talk", und so konnte ich meine Wettbewerber zu Boden ringen.

TL 131: Er will den "Straight Talk", Klartext reden, das ist das Markenzeichen seiner Kampagne.

STRAIGHT **TALK** is not a compound. None of the chosen dictionaries list it as an entry. Consequently, STRAIGHT qualifies as an identical lexical collocate of **TALK**. It should be noted that in TL 25 as well as TL 131 above, explanations of the meaning of STRAIGHT TALK are given. This hints at the fact that this is a relatively new combination whose understanding cannot be assumed yet for the average reader of a news magazine.

Remarkably, when the node **TALK** occurs as a verb form in the source language corpus, it also collocates with STRAIGHT. But then STRAIGHT co-occurs with **TALK** in R1 position. It has been mentioned previously (cf. subsection. 2.2.1.2) that Halliday (cf. 1966: 150/151) illustrates that collocations cut across grammatical boundaries and occur in different syntactic constructions. He describes these as “instances of one and the same syntagmatic relation” (1966: 151). STRAIGHT **TALK** and **TALK** STRAIGHT are examples of this.

The second identical lexical collocate according to the quantitative corpus analysis is EXPRESS. In both corpora, EXPRESS only co-occurs with **TALK** in R1 position. See below for examples from both corpora.

SL 463: […] he could sit around with reporters as he did on his Straight Talk Express bus during the primaries.

TL 26: Der "Straight Talk Express" aber bleibt das Kultobjekt der McCainiacs. Drei bis vier dieser blauen Busse […].

‘Straight Talk Express’ is a term which was coined in 2008 during the primaries for the nearing presidential election in the USA. **TALK** and EXPRESS are both constituents of a proper name, in this case the name for a bus. EXPRESS does not qualify as an identical lexical collocate of **TALK**. Accordingly, this also means that in some of the instantiations in which STRAIGHT and **TALK** collocate, they are followed by EXPRESS. However,

this does not influence the result for STRAIGHT as an identical lexical collocate of the node TALK.

The third identical lexical collocate of **TALK** which was identified during the quantitative corpus analysis is SHOW. In the source and target language corpora, SHOW co-occurs with **TALK** exclusively in R1 position. Only instantiations in which SHOW is used as a noun in the source language corpus were counted. This is exemplified below.

SL 303: He recounted the story 15 years later on his talk show.

TL 5: Der Privatsender hat eine Talk-show, die den Titel nicht verdient.

Oxford Dictionaries defines **TALK** SHOW as "a chat show, especially one in which listeners, viewers, or the studio audience are invited to participate in the discussion." **TALK** SHOW is translated by *PONS* only with the Anglicism "Talk show". As a result, SHOW does not qualify as an identical lexical collocate, but rather forms a compound with the node **TALK**. It should be noted that in both corpora the plural form SHOWS also qualified as a significant lexical collocate, but did not make the Top 5. To illustrate this, one example per corpus is given below.

SL 266: […] the point—to spoof the awkward false intimacy of talk shows.

TL 6: […] Wirtschaftssenator mit eigener kleiner Bonusmeilen-Affäre, ansonsten eindeutiger Liebling der Talk-shows - dieser Gregor Gysi hat den Osten so gut überstanden wie den Westen.

The final identical lexical collocate of **TALK** is RADIO. It is claimed that RADIO is an identical rather than an equivalent lexical collocate although it might be pronounced differently in the source and target language. The present study is concerned exclusively with written word-forms, thus it seems legitimate to neglect potential differences in the pronunciation of RADIO and to focus on the identically spelled word-form alone.

In most cases RADIO co-occurred with **TALK** in R1 position in the source language corpus. In a few instantiations, RADIO co-occurred with **TALK** in L1 position of the span. Please see below for corresponding examples.

SL 348: A second argument is that talk radio is not out of step at all […].

SL 432: Last week, radio talk-show king Rush Limbaugh ranked Buckley as a "founding father."

In the target language corpus, RADIO co-occurred with **TALK** also in L1 and R1 position. Below, examples are given.

TL 34: Talk radio ist ein Hörfunkformat mit einem sehr hohen Wortanteil.

TL 71: Im September lief die Sendung Pasternys Plausch – der Radio talk auf […].

Oxford Dictionaries defines **TALK** RADIO as "a type of radio broadcast in which the presenter talks about topical issues and encourages listeners to phone in to give their opinions." RADIO in R1 position of the node **TALK** thus does not qualify as an identical lexical collocate. The instantiations in which RADIO co-occurs with **TALK** in L1 position are not frequent enough (below 5) to count it as an identical collocate in the framework of the present study.

Finally, it should be noted that the target language collocate SMALL also qualified as a significant source language collocate. In both corpora it formed the compound SMALL **TALK** with the node, which is defined by *Oxford Dictionaries* as "polite conversation about unimportant or uncontroversial matters, especially as engaged in on social occasions."

TEST

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
535	56	SCORES	(54)	Punktzahl/ Ergebnis	PRÄSIDENT	(18)	12	274
		GRADES	(22)	Noten	THEORIE	(15)		
		BLOOD	(19)	Blut	FRAGEN	(14)		
		STUDENTS	(17)	Studenten/ Schüler	POSITIV	(14)		
		TAKE	(16)	machen/ sich unterziehen	WICHTIGER	(13)		

The node **TEST** does not feature any identical or equivalent lexical collocates in the source language corpora. It should be noted that the word-form **TEST** is also a verb in English. For the analysis only instantiations of the noun **TEST** were considered. It should be noted that no Anglicisms were identified as lexical target language collocates. This has previously been observed especially with nodes which are more frequent in the target language than in the source language. This criterion is not fulfilled for the node **TEST**. The fact that only target language word-forms were identified as target language lexical collocates may indicate its degree of institutionalization.

TREND

SOURCE LANGUAGE				TRANSLATION	TARGET LANGUAGE			
T	LC	Top 5			Top 5		LC	T
206	16	HOT	(10)	heiss/scharf	MEDIEN	(24)	29	417
		PEOPLE	(6)	Leute/Menschen	EINDEUTIG	(21)		
		NEW	(6)	neu	ÄHNLICH	(18)		
		LONG	(5)	lang	GROß	(17)		
		GROWING	(5)	zunehmend/ Steigend	LANGFRISTIG	(12)		

The node **TREND** does not feature any identical or equivalent lexical collocates in the source and target language corpora. Nevertheless, it should be noted that **TREND** is more frequent in the target language corpus than in the source language corpus. The node is defined by *Oxford Dictionaries* as either "a general direction in which something is developing or changing" or "a fashion." **TREND** would, according to *PONS*, be best translated with the Anglicism "Trend" or the German words "Richtung" or "Tendenz". Neither of the proposed German translation equivalents does quite catch the source language meaning of **TREND** and would have to be further specified. A similar discussion has been carried out for the node **BESTSELLER** above. In addition, **TREND** is not as lengthy as 'Richtung' and 'Tendenz' and most likely preferred for this reason as well.

TREND does not feature any Anglicisms as target language collocates. This is in line with the previous results for nodes which are more frequent in the target language than the source language corpus.

6.3.3 Summary of results and evaluation

From the outset of the present study it was assumed that the use of Anglicisms produces replications of source language collocational structures within the target language, i.e. Anglicisms feature identical or equivalent collocates in the source and target language corpora. In addition it was hypothesized that this assumption is true for specialized and general journalistic texts. The findings of the quantitative analysis of the news magazine corpora suggest that 20 out of the 36 analyzed nodes share identical and/or equivalent lexical collocates. However, the qualitative interpretation of these quantitative results shows that actually only 12 out of the 36 nodes share identical and/or equivalent lexical collocates.

Similar to the corpus analysis of the business magazines, there is one main reason for the deviation in the identical and equivalent lexical collocate count from quantitative to qualitative analysis: The statistical corpus analysis of collocations also identifies constituents of a compound and constituents of proper names as collocates of a given node. A summary of the results of the qualitative analysis of the Top 5 collocates is given in Figure 35 below. Figure 35 is similarly constructed as Figure 33 (cf. subsection 6.3.1).

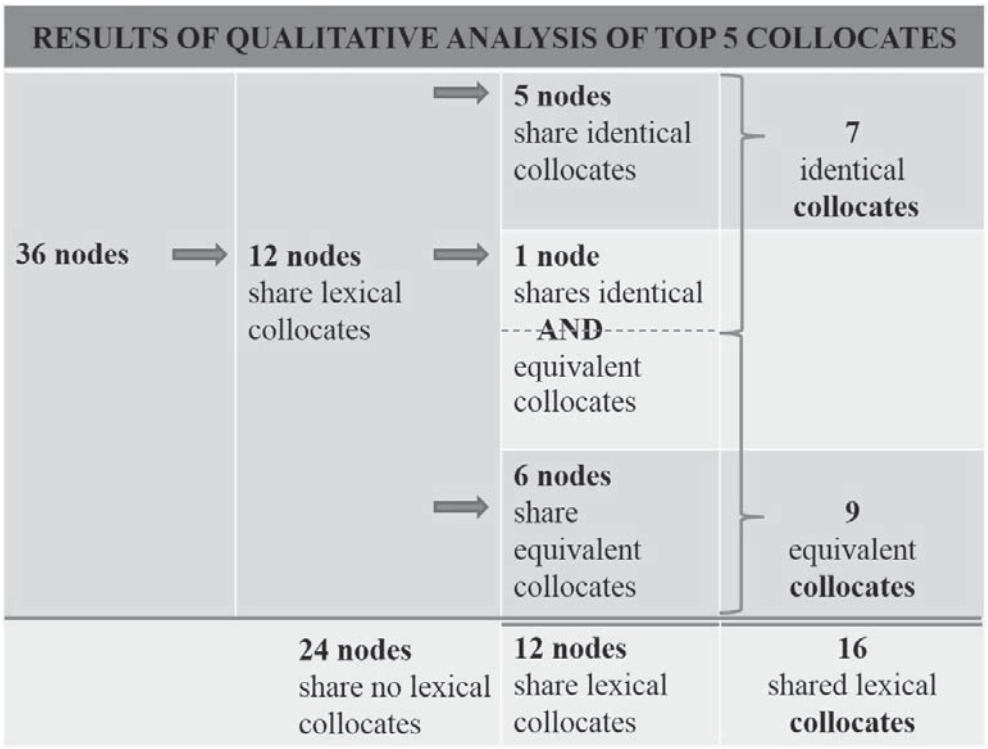

Figure 35 Results of the qualitative analysis of Top 5 news magazine collocates

A more detailed synopsis of the quantitative and qualitative analysis of the news magazine corpora is provided in Figure 36 below. Figure 36 facilitates the comparison between the results from the quantitative and qualitative corpus analysis. The qualitative corpus analysis revealed that out of the originally 20 nodes thought to display similar collocational behavior, 8 do not. These are marked with a small 'x' in figure 36. Identical collocates are marked with the mathematical 'identical to' sign '≡' and equivalent collocates are marked with the mathematical 'corresponds to' sign '≙'.

SYNOPSIS OF QUANTITATIVE AND QUALITATIVE ANALYSIS OF THE NEWS MAGAZINE CORPORA							
			QUANTITATIVE ANALYSIS		QUALITATIVE ANALYSIS		
NODE	FREQUENCY SL	FREQUENCY TL	IDENTICAL COLLOCATES	EQUIVALENT COLLOCATES	IDENTICAL COLLOCATES	EQUIVALENT COLLOCATES	REASON FOR DEVIATION
(1) AIRPORT	111	101	0	0	0	0	-
(2) BESTSELLER x	117	134	0	1	0	0	Compound
(3) BOOM x	191	180	1	0	0	0	Compound
(4) CAMP	165	115	0	0	0	0	-
(5) CLAN	105	217	0	0	0	0	-
(6) COMEBACK	149	165	0	0	0	0	-
(7) COMEDY	113	109	2	0	1 ≡	0	Compound
(8) COMMUNITY	783	127	1	0	1 ≡	0	-
(9) DOTCOM	115	106	2	1	2 ≡	1 ≙	-
(10) ENTERTAINMENT	184	130	0	0	0	0	-
(11) EVENT	202	120	0	0	0	0	-
(12) FAN	176	160	0	0	0	0	-
(13) FARM	198	116	0	1	0	1 ≙	-
(14) FASHION	156	116	3	0	1 ≡	0	Compound
(15) FREAK	144	119	0	0	0	0	-
(16) GANGSTER	142	115	0	0	0	0	-
(17) INTERVIEW	398	195	0	1	0	1 ≙	-
(18) JOB	718	443	0	2	0	2 ≙	-
(19) KILLER	144	180	0	0	0	0	-
(20) KNOW-HOW	118	104	0	1	0	1 ≙	-
(21) LIFE	1,652	160	2	0	1 ≡	0	Compound
(22) LUNCH	121	110	0	0	0	0	-
(23) MAIL	345	339	0	2	0	1 ≙	Compound
(24) MEETING x	497	115	1	0	0	0	Compound
(25) MODEL x	331	159	1	0	0	0	Proper name
(26) NETWORK x	279	119	1	0	0	0	Compound
(27) NEWS x	629	107	2	0	0	0	Compound Proper name
(28) PERFORMANCE	335	109	0	0	0	0	-
(29) REALITY x	358	132	2	0	0	0	Compound
(30) SOCIETY	459	129	0	0	0	0	-
(31) SOUND	307	171	0	0	0	0	-
(32) STAR	279	256	0	2	0	2 ≙	-
(33) STORY x	1,162	199	1	0	0	0	Compound
(34) TALK	1,052	198	4	0	1 ≡	0	Compound Proper name
(35) TEST	535	274	0	0	0	0	-
(36) TREND	206	417	0	0	0	0	-

Figure 36 Synopsis of the results of the quantitative and qualitative analysis of the news magazine corpora

It can be seen that for 8 nodes the quantitatively identified collocates did not ultimately qualify as such, because in fact they are constituents of compounds. Examples would be **FASHION** MAGAZINE or **FASHION** SHOW. In 3 cases, the quantitatively identified collocates formed a proper name with the given node, such as names of a magazine or television show.

It has already been pointed out in subsection 6.2.3 above that the deviation from quantitative and qualitative analysis is caused by the lexical analysis software, which cannot differentiate between constituents of compounds and collocates. In addition, *WordSmith Tools* cannot differentiate between proper names and collocates. As a consequence, the

identification of constituents of compounds as collocates cannot be prevented in the course of the statistical identification of collocations in the present study.

Many news magazine target language collocates are Anglicisms. A total of 155 lexical Top 5 collocates of the 36 nodes were identified in the target language corpus. Out of these 155 lexical target language collocates 62, or 40% are Anglicisms. The other 93 Top 5 target language collocates, or 60%, are German collocates. The fact that numerous Anglicisms were identified as news magazine target language corpus collocates emphasizes that Anglicisms can be borrowed along with identical collocates or as a constituent of a source language compound which is used in identical form in the target language.

In subsection 6.2.3, it was hypothesized that the analysis of the news magazine corpora would reveal more adjectives and verbs as collocates than were identified in the business magazine corpora. This assumption was fulfilled as Figure 37 below illustrates the distribution of the Top 5 collocates according to word classes in the business and news magazine corpora. Figure 37 also gives the share of the different word classes in percentages, to bring out this contrast in all clarity. Note that the fact that the total of Top 5 collocates is identical, i.e. 175, in both types of corpora is pure coincidence.

TOP 5 COLLOCATES ACCORDING TO WORD CLASSES IN BUSINESS AND NEWS MAGAZINE CORPORA								
	BUSINESS MAGAZINE				NEWS MAGAZINE			
	SL[1] CORPUS		TL[2] CORPUS		SL CORPUS		TL CORPUS	
	NUMBER	% OF TOTAL	NUMBER	% OF TOTAL	NUMBER	% OF TOTAL	NUMBER	% OF TOTAL
NOUNS	129	73.9	118	72.4	99	56.6	105	67.7
ADJECTIVES	41	23.5	34	20.9	57	32.6	42	27.1
VERBS	4	2.4	11	6.7	18	10.3	8	5.2
ADVERBS	1	0.6	0	0.0	1	0.5	0	0.0
TOTAL	**175**	**100%**	**163**	**100%**	**175**	**100%**	**155**	**100%**

Figure 37 Top 5 collocates according to word classes in business and news magazine corpora

Figure 37 above shows that the percentage of nouns among the Top 5 collocates is lower in the news magazine corpora than in the business magazine corpora. The percentage of adjectives as Top 5 collocates, on the other hand, is higher. In the business magazine source

[29] In Figure 37, SL stands for 'source language'.
[30] In Figure 37, TL stands for 'target language'.

language corpus, the share of verbs among the Top 5 collocates is lower than in the news magazine source language corpus. The opposite is the case for the business and news magazine target language corpora. It should be mentioned that the difference here is not as distinctive though. It was expected that the amount of Top 5 collocates which are nouns would be lower in the news magazine corpora. This was anticipated because technical terms are mostly nouns and more likely to occur in specialized text than in general news reporting. The qualitative analysis of the news magazine corpora identified 21 nouns as shared, i.e. identical or equivalent collocates, 11 adjectives and 2 verbs.

Generally speaking, a given node's overall frequency in the news magazine target language corpus does not reliably indicate whether it will co-occur with additional Anglicisms or target language word-forms. In the news magazine target language corpus, 4 nodes occurred which are used more frequently than in the source language corpus. The results from the business magazine corpora for nodes of this type were confirmed by the news magazine corpus analysis. The corpus analysis revealed that nodes collocate typically with target language word-forms, instead of additional Anglicisms, if their overall frequency is higher in the target language than in the source language corpus. However, sometimes nodes like **KNOW-HOW**, which occur more frequently in the source language corpus than in the target language corpus, collocate exclusively with target language collocates as well. This illustrates that there are clear tendencies, but also exceptions to the rule. There are also Anglicisms, like **LIFE**, which are less frequent in the target language corpus than in the source language corpus and only feature Anglicisms as collocates.

As a central result of the news magazine corpus analysis, it can be recorded that less technical nodes in more general corpora feature less lexical collocates than more technical nodes in specialized text. For this reason, the minimum frequency of a node in the news magazine target language corpus needed to be set to 100 occurrences. In the business magazine target language corpus, much lower overall frequencies sufficed to return plenty lexical collocates for analysis. This difference between both target language corpora is closely related to the fact that the 'Top 5 methodology' proved more rewarding for the news magazine corpus analysis: The number of lexical collocates which was identified in addition to the Top 5 collocates was very small for most nodes. Therefore, the risk of missing shared collocates was significantly reduced in comparison to the business magazine corpus analysis. It can be further stated that in almost all cases in which additional (in addition to the Top 5 collocates) Anglicisms were identified as target language collocates, they were not identified as source language collocates of the same node.

This can be emphasized again by the fact that the sheer number of Anglicisms which are identified as significant lexical target language collocates does not necessarily indicate similar collocational behavior of a node. The node **AIRPORT** features only three lexical target language collocates, two of which are Anglicisms: NEW and SECURITY. Neither NEW, nor SECURITY are identified as source language collocates. In the same way, the node **LUNCH** features three Anglicisms among the Top 5 target language collocates: SANDWICH, COCKTAIL and BUSINESS. None of them occurs as a collocate in the source language corpus. **STORY** collocates in the target language corpus with four Anglicisms among the Top 5 collocates alone. Two of them are PATCHWORK and WEB, but they do not occur as collocates in the source language corpus. Numerous examples could be added to this, but the quoted three examples should suffice to support the argument put forward.

Yet, it cannot be ultimately dismissed that possibly more identical (and equivalent) collocates of the 36 nodes could have been identified if all lexical collocates without exception, i.e. without restriction to the Top 5, had been analyzed. Consider for example the node **BOOM**. The identified Top 5 target language collocates INTERNET, CRASH and DOTCOM equally qualify as significant lexical collocates in the source language corpus. Then again, they did not occur frequently enough to make the Top 5 source language collocates. However, this example does not alter the principal argument. Because of the generally smaller number of lexical collocates in the news magazine corpora, the risk of missing significant shared lexical collocates is lower than in the business magazine corpora.

Additional results which were already summarized in subsection 6.2.3 are also valid for the news magazine corpus analysis. For instance, numerous Anglicisms are translated as Anglicisms in the chosen bilingual dictionary *PONS*. Examples are ‘Bestseller’ or ‘Reality-show’. Expressions like BUSINESS **MEETING** and PRAYER **MEETING** do not have a lexical entry in the chosen mono- and bilingual dictionaries and cannot be identified in the source language corpus. Such combinations are referred to as ‘Anglo-neologisms’ (cf. Steffens 2003: 5). The second constituent provides the general meaning and the first constituent makes this meaning more specific. Usually the combinations are semantically transparent. ‘Anglo neologisms’ imitate source language structures in the target language. Once they are institutionalized, they ultimately usually prove to be pseudo-Anglicisms. Further terms which are suspicious of being a pseudo Anglicism are HOME **ENTERTAINMENT**, **FAN**-CAMP or **REALITY** SOAP. One example in which a newly coined hybrid Anglicism effectively suppresses the use of target language equivalents is ÖKO-**FARM** versus ‘Biobauernhof’.

It has been pointed out in the summary of results for the business magazine corpus analysis that sometimes source language compounds and their target language equivalents are used interchangeably in the target language corpus. This is a popular strategy for stylistic variation in business magazine and news magazine articles alike. One example from the news magazine target language corpus would be BREAKING **NEWS** and 'Eilmeldung'. In passing, it is pointed out that the news magazine corpus analysis also confirms that for stylistic variation hybrid Anglicisms such as **EVENT**-FILM, instead of **EVENT**-MOVIE, are used.

The news magazine corpus analysis revealed one example of a complex collocation: AMERICAN WAY OF **LIFE**. It could therefore be illustrated that collocation is not limited to the co-occurrence of exactly two words, but rather "a succession of two or more words that may best be learnt as if it were a single word" (Palmer 1938/1968). More importantly, it was shown that also complex collocations are borrowed in the target language.

The investigation of the node **JOB** pointed out firmly the necessity of qualitative analysis to judge the logical relatedness of node and collocate (cf. subsection 2.3.2 and section 2.4). Although NEW and **JOB** co-occurred within the same sentence and were identified as node and collocate by means of the quantitative corpus analysis, they were not logically related. On the other hand, **JOB** and NEW co-occurred in two different sentences, but clearly formed a collocation. Particularly because strong punctuation marks do not generally act as break points for collocation in the present study, these were excellent examples to illustrate the importance of logical relatedness in collocations and the fact that they can cross sentence boundaries.

The borrowing of the identical compound SOCIAL **NETWORK**, also relates back to the theoretical basis laid out in subsection 5.1.4. Anglicisms are used often in fields which bring or brought technological innovations into everyday life. The frequent use of SOCIAL **NETWORK** is a typical example of this and shows the cultural domination of the United States in this area. SOCIAL **NETWORK** is used to denote recent web-based services. It can be argued that 'soziales Netzwerk', which is used at times interchangeably in the target language corpus to refer to such web-based services, qualifies as a loan translation.

During the news magazine corpus analysis, one example which illustrates that collocations cut across grammatical boundaries and occur in different syntactic constructions was identified. This was discussed in subsection 2.2.1.2 with reference to Halliday (1966). The node **TALK** collocates as a noun with STRAIGHT in L1 position. **TALK** also collocates with STRAIGHT as a verb form in the source language corpus. But then STRAIGHT co-occurs with **TALK** in R1 position.

7 Summary and conclusions

The present study is mainly governed by the question whether Anglicisms, utilized by non-native speakers of English, display similar collocational behavior in the target language compared to the Anglicisms source language.

Two hypotheses were formulated to guide the present study. The first is that the use of Anglicisms produces replications of source language collocational structures within the target language, i.e. that Anglicisms feature identical or equivalent lexical collocates in the source and target language corpora. Secondly, it was hypothesized that this is true for general and specialized journalistic texts.

This study has investigated the collocational behavior of Anglicisms in business and news magazines. The approach to the study is empirical, i.e. based on the analysis of authentic language data. With the help of corpora, contrastive linguists can test and quantify intuition-based contrastive statements within a body of empirical data. Contextualists, in particular, argue that language should be studied in authentic instances of use and stress the need for computer-based corpus linguistics. In a wider perspective, the present study is committed to the tradition of British Contextualism.

The present study has analyzed contrastively the collocational behavior of Anglicisms in four corpora of written journalistic texts taken from American business and news magazines and their German counterparts: the 2008 volumes of *BusinessWeek* and *Wirtschaftswoche*, and of *Newsweek* and *Der Spiegel*; all in all more than 9 million words. In general, a corpus aims to be representative in the sense that findings based on its analysis can be generalized for the language as a whole or a specified subset of it. The four corpora compiled for this study, therefore, are considered to be representative of leading business journalism and general news reporting in Germany and the USA in terms of content and language use.

The external criteria for corpus classification ‘origin’, ‘function’, ‘subject matter’ and ‘audience’, served as distinguishing factors between the business and news magazine corpora. The application of these external criteria guided the organization of the language data under investigation into specialized and general interest categories. Based on the external criteria for corpus classification, matching criteria were defined for the American and German business and news magazine corpora.

The construction of comparable corpora is an effective way of enabling contrastive language studies. The corpora achieve their comparability through similar extra-linguistic criteria, i.e. matching criteria (cf. subsections 4.5.3 and 4.6.3). Comparable corpora provide data which make it possible to study how the same concept can be rendered independently in different languages. Structures to be contrasted are described by means of the same theoretical model. In the present study, contrasting is carried out by investigating collocations in all corpora according to the same concept of collocation. More precisely, two types of contrasting were involved: firstly, contrasting collocations in American English versus German, and secondly, contrasting across magazine types, i.e. business magazines versus news magazines. Consequently, corpus linguistic methodology here serves to reveal specific preferences in word combinations and a certain patterning which can be observed across languages.

One central instrument of language comparison and contrast is a common platform of reference called ‘tertium comparationis’. In the present study, there are several levels of such a common platform of reference. The *tertium comparationis* for English and German are the types of text: business and news magazine articles. The second *tertium comparationis* is provided by the sets of analyzed key words which are identical in both languages and whose collocates were contrasted. Thirdly, pragmatic equivalence, i.e. the nearly identical semantic content of equivalent collocates provides a *tertium comparationis*. Finally, the application of identical concepts of ‘collocation’ and ‘Anglicism’ throughout the study and all corpora offers an additional common platform of reference. The present study is corpus-based rather than corpus-driven since it puts the concept of collocation, which has been formulated before large corpora were available, to the test.

The focus of corpus linguistics, and of this study, is to emphasize frequency in language use, i.e. countable entities. However, frequency cannot serve as an absolute criterion for assigning the status of collocation, but at the same time it provides the basis for collocational analysis. Co-occurrence, neighborhood, recurrence and frequency of collocations can be empirically tested in the corpora. It is more difficult to measure a collocation’s degree of semantic transparency and almost impossible to test mutual expectancy. This means that corpus-based analyses must go beyond simple counts of linguistic features and that it is important to include qualitative, i.e. functional interpretations of quantitative patterns.

The corpus analysis required the conversion of the collected corpus data into .txt-files, so that it could be processed by the lexical analysis software *WordSmith*. Lists of key words (word-forms) which are identical in the source and target language corpora were compiled: one list of key words for the business magazine corpora, and another one for the news magazine corpora. The corpus analysis concentrated on these two sets of key words. Each set of key words contains 36 entries. All key words (nodes) are nouns and highly frequent Anglicisms in the business and news magazine target language corpora. The key words were identified via close reading and with the help of *WordSmith* their overall frequency in the corpora was determined. The key words were evenly chosen from all sections of a given magazine in order to avoid overrepresentation of certain sections.

The aim of the corpus analysis was to determine whether an Anglicism is borrowed by itself or if its collocational surroundings are borrowed as well. The first step with any given key word was to identify its co-text in the source language and target language corpora. Identification was achieved by displaying the key word in context, also known as the 'KWIC format' which has been widely used in data processing. Since it is assumed that the key word is surrounded by identical or equivalent collocates in the target language corpora, the next step was to compute the collocates in the identified co-text of a given key word in the source language corpus. The lexical analysis software determined a given key word's collocates within a collocational span of ±4 word-forms appearing to the right and left of it. The analysis was dedicated to lexical collocates, because lexical items are more likely borrowed than grammatical ones (cf. Onysko 2007: 45). The corpus analysis was restricted to the so called 'Top 5' collocates, i.e. the five collocates which co-occurred most frequently with a node in the source and target language corpus. The applied methodology, as a possible form of processing corpus data, achieved interesting results.

Based on solely the quantitative corpus analysis, the initial hypotheses proved valid for the majority of the data, i.e. the majority of nodes featured identical or equivalent collocates in the business and news magazine source and target language corpora. However, after further consideration of the results of the qualitative analysis, which excludes proper names and constituents of compounds as collocates, a considerable deviation from the quantitatively achieved results became evident. For example, according to the quantitative business magazine corpus analysis, 34 of the 36 nodes featured shared collocates. But the qualitative analysis revealed that in fact only 20 of these 34 nodes shared identical or equivalent collocates. Consequently, the analysis of collocations merely on statistical grounds did not

lead to satisfactory results within the present study. Therefore, collocations in large contrastive corpora require not only quantitative analysis, which is highly advantageous for a pre-selection of decisive material, but additional qualitative analysis as well.

For the business magazines, the results from the qualitative corpus analysis revealed that 20 (55.5%) out of the chosen 36 nodes share identical and/or equivalent lexical collocates in the source and target language corpus. Within the news magazine corpora, 12 (33.3%) out of the chosen 36 nodes share identical and/or equivalent lexical collocates. This results in 32 out of 72 nodes (44.4%) which display similar collocational behavior in the American and German business and news magazine corpora. These 32 nodes again shared a total of 21 identical lexical collocates and 31 equivalent lexical collocates. It should be pointed out again that the 32 nodes featured a total of 52 identical and equivalent collocates, because several nodes feature more than one shared collocate in the source and target language corpora.

In the business and news magazine target language corpora alike, many collocates are Anglicisms. Strikingly, approximately 50% of all Top 5 lexical target language collocates are Anglicisms. The presence of Anglicisms as target language collocates is higher in the business magazine corpora (61.3%) than in the news magazine corpora (40%). This illustrates another important result of the present study. In general and specialized magazines alike, Anglicisms are rarely borrowed as isolated items. That is to say, Anglicisms are often borrowed along with identical collocates, but frequently also as a constituent of a source language compound which is used in identical form in the target language. The frequent identification of source language compounds which are used in identical form in the target language corpora is also a valuable result of the corpus analysis. In other words, Anglicisms are not necessarily borrowed as isolated items.

The imitation of source language collocational and compounding structures repeatedly also leads to the formation of so-called ‘Anglo neologisms’ which over time often become ‘pseudo Anglicisms’. Their formation and use takes place in both types of target language corpora. This conclusion was reached because numerous Anglicisms, which were identified as target language collocates, did not qualify as source language collocates. In such cases it can be said that these nodes (Anglicisms) have developed distinctive target language collocations with source language word-forms. As long as such collocations were specific to the target language corpus and not identified in the comparable source language corpus, they needed to be classified as pseudo Anglicisms.

The corpus analysis also revealed that the formation of ‘hybrid compounds’ such as **INSIDER**-HANDEL is probably promoted by the fact that compounding is a productive

source of new terms in German and English. This is connected to an earlier introduced statement by Görlach, namely that Anglicisms become available for use in compounds very soon after their adoption (cf. 2002: 9) in the target language.

In all corpora, nouns account for the lion's share of all quantitatively and qualitatively identified lexical collocates. However, nouns are represented even more strongly among the Top 5 business magazine collocates than among the Top 5 news magazine collocates. It is assumed that this difference can be attributed to the more frequent use of technical terms (which are typically nouns) in specialized text. Further support can be found in the fact that, in contrast to the business magazine nodes, the news magazine nodes generally feature more verbs and adjectives as Top 5 collocates and thus more adjectives and verbs as shared, i.e. identical or equivalent, collocates.

The analysis of the least specialized business magazine nodes **DEAL** and **WORK** already hinted at the fact that fairly general nodes tend to bind a greater number of grammatical collocates, as opposed to entering collocations with lexical words. Later, during the analysis of the news magazine corpora, this assumption was confirmed. As a central result of the news magazine corpus analysis, it can be demonstrated that less technical nodes in more general corpora tend to feature less lexical collocates than more technical nodes in specialized text. For this reason and to arrive at a sufficient amount of analyzable lexical collocates, a minimum frequency had to be introduced for news magazine nodes. In other words, had the minimum frequency of a given key word not been set to 100 for the news magazine corpora, the quantitative corpus analysis would have returned too few lexical collocates to conduct a 'Top 5' analysis. This was different for the analysis of the business magazines. Even key words with comparatively low overall frequencies usually resulted in numerous lexical collocates. It can be concluded from this that the number of lexical and grammatical collocates of a given node is not dependent on how general or specialized the text in which it occurs, but how general or specialized the node is.

Although the corpus analysis exposes clear tendencies, a given node's overall frequency in the business or news magazine target language corpus does not necessarily indicate whether the node will co-occur with additional Anglicisms or target language word-forms. It can be stated though that nodes typically collocate with target language word-forms, instead of additional Anglicisms if their overall frequency is higher in the target language than in the source language corpus. This is true for the business and news magazine corpora. The number of target language word-forms as collocates in the target language corpora is an indicator of a node's integration, i.e. of the institutionalization of the Anglicism in the target language. Frequency of use is one factor which adds to the

institutionalization process of lexical items, and thus also influences the nature of collocates. The corpus analysis did not expose a conclusive relation between the frequency of a given node and its number of identical and equivalent collocates. Accordingly, it can be concluded that high-frequency and low-frequency Anglicisms alike display similar collocational behavior in the two languages.

According to Quirk et al. (cf. 1985: 1522 ff.), the degree of institutionalization of an Anglicism, as the integration of a lexical item, with a particular form and meaning, into the existing stock of words as a generally acceptable and current lexeme in the target language, is somewhat difficult to determine. Frequency is one contributing factor and so are dictionary entries. Matching this, relatively many Anglicisms are translated with Anglicisms in the chosen bilingual dictionary *PONS*. Lexical entries of this kind display the institutionalization of an Anglicism, and in addition it helps to foster the systematic suppression of the use of target language equivalents. The qualitative corpus analysis suggests that the frequent borrowing of identical compounds and collocations leads to giving up not only the establishment of equivalent collocates in the target language, but also effectively suppresses the use of existing equivalent target language collocates and compounds.

Before concluding, a few additional remarks may prove helpful for future studies. The applied Top 5 methodology has proved more successful for the analysis of the news magazine corpora compared to the business magazine corpora. This is the case because the overall frequency of the chosen nodes is on average lower in the news magazine corpora, than in the business magazine corpora. Their lower frequency results in fewer lexical collocates. Hence, the risk of missing additional shared collocates outside of the Top 5 collocates decreases in comparison to the business magazine analysis. As a consequence it is recommended that future analyses consider a slight modification of the chosen methodology. If Top 5 analysis is carried out, the minimum and maximum frequency of the chosen nodes should be set to stay within suitable limits, proportional to corpus size.

In conclusion, identical and equivalent lexical collocates were identified for nearly half of all 72 analyzed nodes in the business and news magazine corpora. That is to say, similar collocational behavior of the nodes in the source and target language corpora is evident in specialized and general corpora alike. The corpus evidence is more conclusive for the specialized corpora. Whether this is related to the higher frequency of Anglicisms in specialized text remains an open question for further research. The corpus analysis provides sufficient evidence that it is particularly the most frequently occurring source language collocational and compounding structures which are adopted in the target language. This also raises questions about the impact collocations have on the structure of

vocabulary and thus language acquisition, i.e. whether it is useful to expand the units of vocabulary learning to the level of compounds or collocations. However, it should be pointed out that collocations, like any other linguistic phenomenon, depend on the developments of the extra-linguistic world and its conceptual categorization. Items that occur together physically, but also concepts from certain areas of expertise, are likely to be mentioned together. This may account for part of the identified equivalent collocates in particular. Nevertheless, choices in language are not influenced merely by external factors or organizing features, thus the achieved results not less meaningful.

The hypotheses, which the study set out to answer, could be confirmed, i.e. similar collocational behavior of Anglicisms occurs across general-specialized magazine boundaries. Many quantitatively identified collocates proved to be constituents of compounds. Noun-noun combinations are notoriously difficult to classify and to distinguish from other syntactic groups, i.e. from collocations which lack institutionalization, for example in the form of lexical entries in most dictionaries.

While the present study takes a synchronic perspective, collocations are subject to diachronic change. Recurrence and frequency determine the fixedness of collocations and may ultimately be the factors which promote the institutionalization of some collocations as compounds.

Naturally, all conclusions drawn are primarily valid for the analyzed corpora. However, the corpora aim to be representative of German and American business and news journalism and were compiled accordingly. Thus, the achieved results can be generalized to some extent for American and German business and news magazine articles as a whole. In future, it would be interesting to investigate the collocational behavior of Anglicisms in comparable corpora from additional domains. It is very likely that the use of Anglicisms will continue to increase in German and the corpus analysis has also illustrated the extensive borrowing and use of source language compound nouns in the target language corpora. Building on the results of the present study then, a comparison of the collocational behavior of source language compounds such as PRIVATE **EQUITY** or INVESTMENT **BANKING**, DOTCOM **BOOM** or **REALITY** SHOW in the target language, might lead to further insights on the subject matter.

As a result of the globalization of society, multilingual studies have increased. Research in this area demonstrates the effective cross-cultural exchange of ideas. In this respect, corpus analysis is useful to document and explain these cross-cultural exchanges.

References

Abrahamson, David. 1995. *The American Magazine. Research Perspectives and Prospects.* Ames, IA: Iowa State University Press.

Adams, Valerie. 1973. *An Introduction to Modern English Word-Formation*. London: Longman.

Alatis, James E. 1968. *Contrastive Linguistics and its Pedagogical Implications.* Washington, D.C.: Georgetown University Press.

Bahns, Jens. 1993. "Lexical Collocations: a Contrastive View." *ELT journal: An International Journal for Teachers of English to Speakers of Other Languages* 47, 56-63.

Bahns, Jens. 1996. *Kollokationen als lexikographisches Problem: Eine Analyse allgemeiner und spezieller Lernerwörterbücher des Englischen*. Tübingen: Niemeyer.

Baker, Mona. 1992. *In Other Words. A Coursebook on Translation*. London: Routledge.

Baldwin, Timothy, Colin Bannardz, Takaaki Tanaka and Dominic Widdows. 2003. "An Empirical Model of Multiword Expression Decomposability." *Proceedings of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, 89-96.

Bartsch, Sabine. 2004. *Structural and Functional Properties of Collocations in English*. Tübingen: Narr.

Bassnett, Susan. 1991. "Translating for the Theatre: Textual Complexities". *Essays in Poetics* 15, 71-84.

Bassnett, Susan. 2002. *Translation Studies.* New York, NY: Routledge.

Bell, Allan. 1991. *The Language of News Media.* Oxford: Blackwell.

Benson, Morton, Evelyn Benson and Robert F. Ilson. 1986. *Lexicographic Description of English*. Amsterdam, Philadelphia: John Benjamins.

Betz, Werner. 1936. *Der Einfluss des Lateinischen auf den alt-hochdeutschen Sprachschatz. 1. Der Abrogans.* Heidelberg: Carl Winter Universitätsverlag.

Betz, Werner. 1959. "Lehnwörter und Lehnprägungen im Vor- und Frühdeutschen." *Deutsche Wortgeschichte* (Vol. 2). Eds. Friedrich Maurer and Friedrich Stroh. Berlin, New York, NY: Mouton de Gruyter, 127-147.

Biber, Douglas. 1985. "Investigating Macroscopic Textual Variation through Multi-Feature/Multidimensional Analysis. *Journal of Linguistics* 23, 337-360.

Biber, Douglas. 1986. "Spoken and Written Textual Dimensions in English: Resolving the Contradictory Findings". *Language* 62, 384-414.

Biber, Douglas. 1988. *Variations across Speech and Writing.* Cambridge: Cambridge University Press.

Biber, Douglas. 1995. *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.

Biber, Douglas. 2009. "A Corpus-Driven Approach to Formulaic Language in English: Multi-Word Patterns in Speech and Writing." *International Journal of Corpus Linguistics* 14, 275-311.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Biber, Douglas, Susan Conrad and Randi Reppen. 1998**.** *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, Douglas, Susan Conrad and Viviana Cortes. 2004. "*If you look at...: Lexical Bundles in University Teaching and Textbooks.*" *Applied Linguistics* 25, 371-405.

Bölsche, Jochen. "Obituary John Seymour Chaloner." *Der Spiegel*, 23 February 2007. Web. Retrieved 29 August 2010.

Bohmann, Stephanie. 1996. *Englische Elemente im Gegenwartsdeutsch der Werbebranche*. Marburg: Tectum-Verlag.

Bublitz, Wolfram. 1996. "Semantic Prosody and Cohesive Company: 'Somewhat Predictable'." *Leuvense Bijdragen* 85, 1-32.

Buck, Timothy. 1974. "'Selfmade-Englisch': Semantic Peculiarities of English Loan-Material in Contemporary German." *Forum for Modern Language Studies* 10, 130-146.

Burger, Harald. 2007. *Phraseologie. Eine Einführung am Beispiel des Deutschen*. Berlin: Schmidt.

Busse, Ulrich. 2008. "Anglizismen im Deutschen. Entwicklung, Zahlen, Einstellungen." *Sprachkontakt und Mehrsprachigkeit. Zur Anglizismendiskussion in Deutschland, Österreich, der Schweiz und Italien.* Ed. Sandro M. Moraldo. Heidelberg: Carl Winter Universitätsverlag, 37-68.

Busse, Ulrich and Manfred Görlach. 2002. "German." *English in Europe*. Ed. Ulrich Busse and Manfred Görlach. Oxford: Oxford University Press, 13-36.

Bybee, Joan L. 2006. "From Usage to Grammar: The Mind's Response to Repetition." *Language* 82, 711-733.

Cameron, Lynne and Graham Low. 1999. "Survey Article: Metaphor." *Language Teaching* 32, 77-96.

Caro Cedillo, Anna. 2004. *Fachsprachliche Kollokationen.* Tübingen: Gunter Narr.

Carroll, Susanne E. 1992. "On Cognates." *Second Language Research* 8, 93–119.

Carstairs-McCarthy, Andrew. 2006. *An Introduction to English Morphology: Words and their Structure*. Edinburgh: Edinburgh University Press.

Carstensen, Broder. 1965. *Englische Einflüsse auf die deutsche Sprache nach 1945*. Heidelberg: Carl Winter Universitätsverlag.

Carstensen, Broder. 1971. *Spiegel-Wörter, Spiegel-Worte: zur Sprache eines deutschen Nachrichtenmagazins*. München: Hueber.

Carstensen, Broder. 1975. "Amerikanische Einflüsse auf die deutsche Sprache." *Jahrbuch für Amerikastudien* (8/1963). 3rd edition. Heidelberg: Carl Winter Universitätsverlag, 11-34.

Carstensen, Broder and Ulrich Busse. 2001. *Anglizismenwörterbuch. Der Einfluss des Englischen auf den deutschen Wortschatz nach 1945*. Berlin, New York, NY: De Gruyter.

Carter, Ronald and Michael McCarthy. 1988. *Vocabulary and Language Teaching*. Harlow: Longman.

Catford, John C. 1965. *A Linguistic Theory of Translation*. Oxford: Oxford University Press.

Chesterman, Andrew. 1998. *Contrastive Functional Analysis*. Amsterdam, Philadelphia: John Benjamins.

Church, Kenneth and Patrick Hanks. 1990. "Word Association Norms, Mutual Information & Lexicography." *Computational Linguistics* 16, 22-29.

Clear, Jeremy. 1993. "From Firth Principles: Computational Tools for the Study of Collocation." *Text and Technology: In Honour of John Sinclair*. Eds. Mona Baker, Gill Francis, and Elena Tognini-Bonelli. Amsterdam: John Benjamins, 271-292.

Comrie, Bernard. 1976. "The Syntax of Causative Constructions: Cross-Language Similarities and Divergences." *Syntax and Semantics*. Ed. Masayoshi Shibatani. New York, NY: Academic Press, 261-312.

Coseriu, Eugenio. 1970. "Über Leistung und Grenzen der kontrastiven Grammatik." *Probleme der kontrastiven Grammatik*. Ed. Hugo Moser. Düsseldorf: Schwann, 9-30.

Coseriu, Eugenio. 1981. "Kontrastive Linguistik und Übersetzung: ihr Verhältnis zueinander." *Kontrastive Linguistik und Übersetzungswissenschaft*. Eds. Wolfgang Kühlwein, Gisela Thome, Wolfram Wilss. München: Wilhelm Fink.

Cowie, Anthony P. 1978. "The Place of Illustrative Material and Collocations in the Design of a Learner's Dictionary." *In Honour of A.S. Hornby*. Ed. Peter Strevens. Oxford: Oxford University Press, 127-139.

Cowie, Anthony P. 1992. "Multiword Lexical Units and Communicative Language Teaching." *Vocabulary and Applied Linguistics.* Eds. Pierre J.L. Arnaud and Henri Béjoint. London: Macmillan, 1-12.

Cowie, Anthony P. 2001. *Phraseology: Theory, Analysis and Applications.* Oxford: Oxford University Press.

Cruse, Alan D. 1986. *Lexical Semantics.* Cambridge: Cambridge University Press.

Crystal, David. 2004. *English as a Global Language.* Cambridge: Cambridge University Press.

Crystal, David and Derek Davy. 1969. *Investigating English Style*. London: Longman.

Deveny, Kathleen. “Reinventing Newsweek.” *The Daily Beast*, May 18, 2009. Web. Retrieved 29 May 2009.

Di Pietro, Robert J. 1971. *Language Structures in Contrast*. Rowley, MA: Newbury House.

Dirven, René. 1974. *A Redefinition of Contrastive Linguistics*. Trier: Linguistic Agency University Trier.

Dobrovok’skij, Dimitrij. 1997. *Idiome im mentalen Lexikon. Ziele und Methoden der kognitivbasierten Phraseologieforschung*. Trier: Wissenschaftlicher Verlag Trier.

Donalies, Elke. 1992. “Hippies Hopping und Toughe Trendies. Über (neu)modische, noch nicht kodifizierte Anglizismen in deutschsprachigen Female-Yuppie-Zeitschriften.” *Deutsche Sprache: Zeitschrift für Theorie, Praxis, Dokumentation*, 97-110.

Drews, Jörg. 1999. “Auf dem Weg zum Denglisch. Wieviel Angloamerikanisch verträgt die deutsche Sprache?” *Sprache in Not? Zur Lage des heutigen Deutsch*. Ed. Christian Meier. Göttingen: Wallstein, 15-31.

Duckworth, David. 1977. “Zur terminologischen und systematischen Grundlage der Forschung auf dem Gebiet der englisch-deutschen Interferenz.” *Sprachliche Interferenz*. Eds. Herbert Kolb and Hartmut Lauffer. Tübingen: Niemeyer, 36-56.

Engels, Barbara. 1976. *Gebrauchsanstieg der lexikalischen und semantischen Amerikanismen in zwei Jahrgängen der ‘Welt’ (1954 und 1964)*. Frankfurt am Main: Lang.

Erman, Britt and Beatrice Warren. 2000. “The Idiom Principle and the Open-Choice Principle.” *Text* 20, 29-62.

Esser, Jürgen. 1980. “Contrastive Analysis at the Crossroads of Linguistics and Foreign Language Teaching.” *International Review of Applied Linguistics in Language Teaching* 18, 181-91.

Esser, Jürgen. 1999. “Collocation, Colligation, Semantic Preference and Semantic Prosody: New Developments in the Study of Syntagmatic Word Relations.” *Words, Lexemes, Concepts – Approaches to the Lexicon: Studies in Honour of Leonard Lipka*. Eds. Wolfgang Falkner and Hans-Jörg Schmid. Tübingen: Gunter Narr, 155-166.

Esser, Jürgen. 2009. *Introduction to English Text-Linguistics*. Frankfurt am Main: Lang.

Fellbaum, Christiane. 2007. *Idioms and Collocations*. *Corpus-based Linguistic and Lexicographic Studies*. London: Continuum.

Fink, Hermann. 1970. *Amerikanismen im Wortschatz der deutschen Tagespresse dargestellt am Beispiel dreier überregionaler Zeitungen (Süddeutsche Zeitung, Frankfurter Allgemeine Zeitung, Die Welt)*. München: Hueber.

Fillmore, Charles J., Paul Kay and Mary Catherine O’Connor. 1988. “Regularity and Idiomaticity in Grammatical Constructions: The Case of ‘Let alone’.” *Language* 64, 501-538.

Firth, John R. 1935. “The Technique of Semantics”. *Transactions of the Philological Society* 34, 36-73.

Firth, John R. [1951] 1957a. “Modes of Meaning.” *Papers in Linguistics 1934-1951*. Ed. John R. Firth. London: Oxford University Press, 180-215.

Firth, John R. 1957b. “A Synopsis of Linguistic Theory, 1930-1955.” *Studies in Linguistic Analysis*. London: Blackwell, 1-31.

Firth, John R. 1957c. “Ethnographic Analysis and Language with Reference to Malinowski’s Views”. *Man and Culture: An Evaluation of the Work of Bronisław Malinowski.* London: Routledge & Kegan Paul, 93-118.

Firth, John R. 1966. *The Tongues of Men & Speech*. Oxford: Oxford University Press.

Firth, John R. 1968. *Selected Papers of J.R. Firth, 1952-59*. Ed. Frank Palmer. London: Longman.

Fisiak, Jacek. 1980. *Theoretical Issues in Contrastive Linguistics*. Amsterdam: John Benjamins.

Fisiak, Jacek. 1981. *Contrastive Linguistics and the Language Teacher*. Oxford, New York, NY: Pergamon Press.

Ford, James L.C. 1969. *Magazines for Millions. The Story of Specialized Publications*. Carbondale, IL: Southern Illinois University Press.

Fries, Charles Carpenter. 1940. “The Grammatical Structure of Present-Day American English with Especial Reference to Social Differences or Class Dialects.” *American English Grammar*. New York, NY: Appleton-Century-Crofts.

Fries, Charles Carpenter. 1945. *Teaching and Learning English as a Second Language*. Ann Arbor, MI: University of Michigan Press.

Fries, Charles Carpenter. 1952. *The Structure of English. An Introduction to the Construction of English Sentences*. New York, NY: Harcourt, Brace and World.

Gast, Volker. (forthcoming). “Contrastive Analysis: Theories and Methods.” *Dictionaries of Linguistics and Communication Science: Linguistic Theory and Methodology*. Eds. Bernd Kortmann and Johannes Kabatek. Berlin: De Gruyter. [The article is already available online at http://www.personal.uni-jena.de/~mu65qev/papdf/contr_ling_meth.pdf].

Gitsaki, Christina. 1996. *The Development of ESL Collocational Knowledge*. Diss. The University of Queensland: Centre for Language Teaching and Research Institution.

Gläser, Rosemarie. 1990. *Fachtextsorten im Englischen*. Tübingen: Gunter Narr.

Gläser, Rosemarie. 1998. “Language for Specific Purposes.” *Concise Encyclopedia of Pragmatics.* Ed. Jacob L. Mey. Amsterdam: Elsevier, 469-74.

Görlach, Manfred. 2002. *English in Europe.* Oxford: Oxford University Press.

Granger, Sylviane. 2003. “The Corpus Approach: A Common Way Forward for Contrastive Linguistics and Translation Studies?” *Corpus-based Approaches to Contrastive Linguistics and Translation Studies.* Eds. Sylviane Granger, Jacques Lerot and Stephanie Petch-Tyson. Amsterdam: Rodopi, 17-31.

Greenbaum, Sidney. 1970. *Verb-Intensifier Collocations in English: An Experimental Approach.* Den Haag: Mouton de Gruyter.

Greenbaum, Sidney. 1988. *Good English and the Grammarian.* London, New York, NY: Longman.

Gross, Alan G., Joseph H. Harmon and Michael Reidy. 2002. *Communicating Science. The Scientific Article from the 17th Century to the Present.* Oxford: Oxford University Press.

Gunnarsson, Britt-Louise. 1987. "Facktexten och den sociala kontexten. En analysmodell." *Facktext.* Ed. Britt-Louise Gunnarsson. Malmö: Liber, 72-103.

Haarmann, Harald. 2002. "Englisch, Network Society und europäische Identität: eine sprachökologische Standortbestimmung." *Deutsch-Englisch-Europäisch: Impulse für eine neue Sprachpolitik.* Ed. Rudolf Hoberg. Mannheim: Duden, 152-170.

Halliday, M.A.K. 1966. "Lexis as a Linguistic Level." *In Memory of J.R. Firth.* Eds. Charles E. Bazell, John C. Catford, M.A.K. Halliday and Robert H. Robins. London: Longman, 148-162.

Halliday, M.A.K. 1978. *Language as a Social Semiotic: The Social Interpretation of Language and Meaning.* London: Edward Arnold.

Halliday, M.A.K. 1984. "Language as Code and Language as Behaviour: A Systemic-Functional Interpretation of the Nature and Ontogenesis of Dialogue." *The Semiotics of Culture and Language Vol.1: Language as Social Semiotic.* Eds. Robin Fawcett, M.A.K. Halliday, Sydney M. Lamb and Adam Makkai. London: Frances Pinter, 3-35.

Halliday, M.A.K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.

Halliday, M.A.K. 1988. "On the Language of Physical Science." *Registers of Written English: Situational Factors and Linguistic Features*. Ed. Frances Pinter. London: Mohsen Ghadessy, 162-178.

Halliday, M.A.K. 1989. "Some Grammatical Problems in Scientific English." *Australian Review of Applied Linguistics Series S* 6, 13-37.

Halliday, M.A.K. 2004. *Lexicology and Corpus Linguistics: An Introduction.* London: Continuum.

Halliday, M.A.K., Angus McIntosh and Peter Strevens. 1964. *The Linguistic Sciences and Language Teaching*. Eds. M.A.K. Halliday, Angus McIntosh and Peter Strevens. London: Longman.

Halliday, M.A.K. and Ruqaiya Hasan. 1976. *Cohesion in English.* London: Longman.

Handl, Susanne. 2005. *Collocation – Convenience Food for the Learner*. Diss. Ludwig-Maximilians-Universität München.

Haramundanis, Katherine. 1998. *The Art of Technical Documentation*. Boston: Digital Press.

Hausmann, Franz. 1979. "Un dictionnaire des collocations est-il possible?" *Travaux de Linguistique et de Littérature* 17, 187-195.

Hausmann, Franz. 1984. “Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen.” *Praxis des neusprachlichen Unterrichts* 31, 395-406.

Hausmann, Franz. 1985. “Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels.” *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 1984.* Eds. Henning Bergenholtz and Joachim Mugdan. Tübingen: Niemeyer, 118-129.

Hausmann, Franz. 1989. “Le dictionnaire de collocations.” *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie 1*. Eds. Franz J. Hausmann, Otto Reichmann, Herbert Wiegand, Ladislav Zgusta. Berlin: De Gruyter, 1010-1019.

Hausmann, Franz. 2004. “Was sind eigentlich Kollokationen?” *Wortverbindungen – mehr oder weniger fest.* Ed. Kathrin Steyer. Berlin: De Gruyter, 309-334.

Hawkins, John. 1986. *A Comparative Typology of English and German. Unifying the Contrasts.* London: Croom Helm.

Hedderich, Norbert. 2003. “Language Change in Business German.” *Global Business Languages: Discovering New Economies* (Vol. 8). Eds. Christiane E. Keck and Allen G. Wood. West Lafayette, IN: Purdue Research Foundation, 47-55.

Herbst, Thomas. 1996. “What are Collocations: Sandy Beaches or False Teeth.” *English Studies* 4, 379-393.

Hertzler, Joyce O. 1965. *A Sociology of Language*. New York: Random House.

Hill, Jimmie. 2000. “Revising Priorities: From Grammatical Failure to Collocational Success.” *Teaching Collocation*. Ed. Michael Lewis. Hove, England: Language Teaching Publications, 47-67.

Hockett, Charles F. 1956. “Sammelrezension von vier Werken zur Anthropologie.” *Language* 32, 460-469.

Hoey, Michael. 1991. *Patterns of Lexis in Text.* Oxford: Oxford University Press.

Hoey, Michael. 2005. *Lexical Priming.* London: Routledge.

Holmes, James. 1970. *The Nature of Translation Essays on the Theory and Practice of Literary Translation*. The Hague: Mouton.

Holmes, James. [1972] 1988. “The Name and Nature of Translation Studies.” *Translated! Papers on Literary Translation and Translation Studies*. Ed. James Holmes. Amsterdam, Atlanta, GA: Rodopi.

Holmes, James. 1991. *Translation Studies: The State of the Art. Proceedings of the 1st James S. Holmes Symposium on Translation Studies*. Amsterdam, Atlanta, GA: Rodopi.

Hornby, Albert S. 2011. “Preface.” *Oxford Advanced Learner’s Dictionary*. 8th edition. Oxford: Oxford University Press, vii-viii.

Howatt, Anthony P.R. 1984. *A History of English Language Teaching*. Oxford: Oxford University Press.

Huckin, Thomas A. and Leslie A. Olsen. 1991. *Technical Writing and Professional Communication of Nonnative Speakers of English*. 2nd edition. New York, NY: McGraw Hill.

Hunston, Susan. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Ivir, Vladimir. 1969. "Contrasting via Translation." *The Yugoslav Serbo-Croatian - English Contrastive Project* (*YSCECP) Studies* 1, 13-25.

James, Carl. 1980. *Contrastive Analysis.* London: Longman.

Johansson, Stig and Hilde Hasselgård. 1999. "Corpora and Cross-linguistic Research in the Nordic Countries." *Linguistique contrastive et traduction* 34, 145-162.

Johansson, Stig and Knut Hofland. 1994. "Towards an English-Norwegian Parallel Corpus." *Creating and Using English Language Corpora.* Eds. Udo Fries, Gunnel Tottie and Petra Schneider. Amsterdam, Atlanta, GA: Rodopi, 25-37.

Jones, Susan and John M. Sinclair. 1974. "English Lexical Collocations." *Cahiers de Lexicologie* 24, 15-61.

Just, Dieter. 1967. *Der Spiegel. Arbeitsweise – Inhalt – Wirkung.* Hannover: Verlag für Literatur und Zeitgeschehen.

Kelley, Robert T. 1997. "Chaos out of Order: The Writerly Discourse of Semipopular Scientific Texts." *The Literature of Science. Perspectives on Popular Scientific Writing.* Ed. Murdo W. McRae. Athens, GA: Athens University Press, 132-151.

Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics.* London, New York, NY: Longman.

Kjellmer, Göran. 1984. "Some Thoughts on Collocational Distinctiveness." *Corpus Linguistics. Recent Developments in the Use of Computer Corpora in English Language Research (*Vol. 1). Eds. Jan Aarts, Willem Meijs. Amsterdam: Rodopi, 163-172.

Kjellmer, Göran. 1987. "Aspects of English Collocations." *Corpus Linguistics and Beyond. Proceedings of the 7th International Conference on English Language Research on Computerized Corpora*. Ed. Willem Meijs. Amsterdam: Rodopi, 130-140.

Kjellmer, Göran. 1990. "Patterns of Collocability." *Theory and Practice in Corpus Linguistics.* Eds. Jan Aarts and Willem Meijs. Amsterdam, Atlanta, GA: Rodopi, 163-178.

Kjellmer, Göran. 1991. "A Mint of Phrases." *English Corpus Linguistics.* Eds. Karin Aijmer, Bengt Altenberg. London: Longman, 111-127.

Klotz, Michael. 2000. *Grammatik und Lexik: Studien zur Syntagmatik englischer Verben.* Tübingen: Stauffenburg.

König, Ekkehart. 1991. *The Meaning of Focus Particles. A Comparative Perspective.* London: Routledge.

König, Ekkehart and Volker Gast. 2009. *Understanding English-German Contrasts.* Eds. Ekkehart König and Volker Gast. Berlin: Erich Schmidt.

Kohn, Kurt. 1992. "Bemerkungen zur Kollokationsproblematik." *Texte, Sätze, Wörter und Moneme. Festschrift für Klaus Heger zum 65. Geburtstag.* Ed. Susanne Anschütz. Heidelberg: Orient-Verlag, 369-387.

Konecny, Christine. 2010. *Kollokationen. Versuch einer semantisch-begrifflichen Annäherung und Klassifizierung anhand italienischer Beispiele*. München: Meidenbauer.

Koller, Werner. 1978. "Äquivalenz in kontrastiver Linguistik und Übersetzungswissenschaft." *Theory and Practice of Translation.* Eds. Lillebil Grähs, Gustav Korlén and Bertil Malmberg. Bern, Frankfurt am Main: Lang, 69-82.

Koller, Werner. 1979. *Einführung in die Übersetzungswissenschaft*. Heidelberg: Quelle & Meyer.

Krenn, Brigitte. 2000. *The Usual Suspects. Data-oriented Models for Identification and Representation of Lexical Collocations*. Saarbrücken: German Research Center for Artificial Intelligence.

Kress, Gunther. 1988. *Communication and Culture. An Introduction.* Kensington: New South Wales University Press.

Krishnamurthy, Ramesh. 2000. "Collocation: from Silly Ass to Lexical Sets." *Words in Context: A Tribute to John Sinclair on his Retirement.* Eds. Chris Heffer, Helen Sauntson and Gwyneth Fox. Birmingham: Birmingham University Press, 31-47.

Krzeszowski, Thomasz P. 1990. *Contrasting Languages: The Scope of Contrastive Linguistics.* Berlin, New York, NY: De Gruyter.

Kussmaul, Paul. 1995. *Training the Translator.* Amsterdam: John Benjamins.

Lado, Robert. 1957. *Linguistics Across Cultures: Applied Linguistics for Language Teachers.* Ann Arbor, MI: University of Michigan Press.

Laffling, John. 1992. "On Constructing a Transfer Dictionary for Man and Machine." *Target* 4, 17-31.

Lakoff, Robin. 1972. "Language in Context." *Language* 48, 907-927.

Leech, Geoffrey. 1992. "Corpora and Theories of Linguistic Performance." *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Ed. Jan Svartvik. Berlin: De Gruyter, 105-122.

Lewis, Michael. 2000. *Teaching Collocations*. Hove: Language Teaching Publications.

Lewis, Michael and Jimmie Hill. 1998. *What is Collocation?* Hove: Language Teaching Publications.

Lehr, Andrea. 1996. "Kollokationen und maschinenlesbare Korpora. Ein operationelles Analysemodell zum Aufbau lexikalischer Netze." *Germanistische Linguistik* 168, 7-95.

Lipka, Leonhard. 1990. *An Outline of English Lexicology, Lexical Structure, Word Semantics, and Word-Formation*. Tübingen: Niemeyer.

Lipka, Leonhard. 2002. *English Lexicology: Lexical Structure, Word Semantics and Word-Formation.* Tübingen: Gunter Narr.

Liston, Jerry L. 1970. "Formal and Semantic Considerations in Contrastive Analysis." *The Yugoslav Serbo-Croatian – English Contrastive Project B* (*YSCECP/B) Studies* 2, 27-49.

Longman Dictionary of Applied Lingistics. LDAL. 1985. Essex: Longman.

Lück, Hartmut. 1963. "Zeitungsdeutsch und Umgangssprache; Untersuchungen zur Sprache des Spiegel." *Muttersprache* 73, 327-337.

Lyons, John. 2004. *Language and Linguistics. An Introduction.* Cambridge: Cambridge University Press.

Mair, Christian and Manfred Markus. 1992. *New Departures in Contrastive Linguistics/Neue Ansätze in der kontrastiven Linguistik. Proceedings of a Conference Held at the Leopold-Franzens-University of Innsbruck, Austria, 10-12 May 1991 2.* (Vol. 2). Eds. Christian Mair and Manfred Markus. Innsbruck: Institut für Anglistik.

Malblanc, Alfred. 1963. *Stylistique compare du francais et de l'allemand.* Paris: Didier.

Malinowski, Bronisław. 1923. "The Problem of Meaning in Primitive Languages." *The Meaning of Meaning.* Eds. Charles K. Ogden and Ivor A. Richards. London: Routledge & Kegan Paul, 296-336.

Manning, Christopher D. and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing.* Cambridge, MA: MIT Press.

Martelli, Aurelia. 2007. *Lexical Collocations in Learner English: A Corpus-based Approach*. Alessandria: Edizioni dell'Orso.

Martin, Joan R. 1984. "Language, Register and Genre." *Language Studies: Children Writing Reader.* Ed. Frances Christie. Geelong, VIC: Deakin University Press, 21-30.

Martin, Joan R. 1993a. "Life as a Noun: Arresting the Universe in Science and Humanities." *Writing Science: Literacy and Discursive Power*. Eds. M.A.K. Halliday and Joan R. Martin. Pittsburgh: Pittsburgh University Press, 221-267.

Martin, Joan R. 1993b. "Technicality and Abstraction: Language for the Creation of Specialized Texts." *Writing Science: Literacy and Discursive Power*. Eds. M.A.K. Halliday and Joan R. Martin. Pittsburgh: Pittsburgh University Press, 203-220.

Mason, Oliver. 1999. "Parameters of Collocation: The Word in the Centre of Gravity." *Corpora Galore: Analyses and Techniques in Describing English*. Ed. John M. Kirk. Amsterdam: Rodopi, 267-280.

McCarthy, Diana, Bill Keller and John Carroll. 2003. “Detecting a Continuum of Compositionality in Phrasal Verbs.” *MWE '03 Proceedings of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*. (Vol. 18). Eds. Sidney Siegel and John N. Castellan. Stroudsburg, PA: Association for Computational Linguistics, 73-80.

McEnery, Tony and Andrew Wilson. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

McKeown, Kathleen R. and Dragomir R. Radev. 2000. “Collocations.” *A Handbook of Natural Language Processing*. Eds. Robert Dale, Hermann Moisl and Harold Somers. New York, NY: Marcel Dekker, 507-523.

McRae, Murdo W. 1997. “Science in Culture.” *The Literature of Science. Perspectives on Popular Scientific Writing.* Ed. Murdo W. McRae. Athens, GA: Athens University Press, 1-17.

Mel’čuk, Igor. 1998. “Collocations and Lexical Functions.” *Phraseology. Theory, Analysis, and Applications*. Ed. Anthony P. Cowie. Oxford: Claredon Press, 23-53.

Mel’čuk, Igor. 2003. “Levels of Dependency in Linguistic Description: Concepts and Problems.” *Dependency and Valency. An International Handbook of Contemporary Research.* (Vol. 1). Eds. Vilmos Ágel, Ludwig Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Herringer, Henning Lobin. Berlin, New York, NY: De Gruyter, 188-229.

Michaels, James W. 1981. “Covering Economics for Magazines and Periodicals.” *Reporting on Business and Economy*. Eds. Louis M. Kohlmeier, Jon G. Udell and Laird B. Anderson. New Jersey, NJ: Prentice-Hall, 308-328.

Miller, Lance A. 1991. “Computers for Composition: A Stage Model Approach to Helping.” *Plain Language: Principles and Practice.* Ed. Erwin R. Steinberg. Detroit, MI: Wayne State University, 93-123.

Möhn, Dieter and Roland Pelka. 1984. *Fachsprachen. Eine Einführung.* Tübingen.

Mohamad Ali, Afida. 2005. *A Comparative Study of Business Journalism Style in Malaysian Business and Management Today. A Functional Textual Approach.* Diss. University of Lancaster.

Moon, Rosamund. 1998. *Fixed Expressions and Idioms in English. A Corpus-based Approach*. Oxford: Oxford University Press.

Muhr, Rudolf. 2004. “Anglizismen als Problem der Linguistik und Sprachpflege in Österreich und Deutschland zu Beginn des 21. Jahrhunderts.” *Eurospeak: Der Einfluss des Englischen auf europäische Sprachen zur Jahrtausendwende.* Eds. Rudolf Muhr and Bernhard Kettemann. Frankfurt am Main: Lang, 9-54.

Muller, Charles. 1963. “Le Mot, unité de texte et unité de lexique en statistique lexicologique.” *Travaux de linguistique et de literature* 1, 155-73.

Munday, Jeremy. 2001. *Introducing Translation Studies. Theories and Applications*. London, New York, NY: Routledge.

Munday, Jeremy. 2009. *The Routledge Companion to Translation Studies.* London, New York, NY: Routledge.

Näslund, Harry. 1987. "Fachjournalistik–en hybrid?" *Språk-vårdssamfundets skrifter* 18, 125-144.

Nelson, M. 2000. *A Corpus-based Study of Business English and Business English Teaching Materials*. Diss. University of Manchester.

Nickel, Gerhard. 1971. *Papers in Contrastive Linguistics.* Cambridge: Cambridge University Press.

Nida, Eugene A. 1964. *Toward a Science of Translating.* Leiden: Brill.

Nida, Eugene A. and Charles R. Taber. 1969. *The Theory and Practice of Translation.* Leiden: Brill.

Onysko, Alexander. 2007. *Anglicisms in German. Borrowing, Lexical Productivity, and Written Code Switching.* Berlin, New York, NY: De Gruyter.

Palmer, Frank R. 1938. (Reprint 1968). *A Grammar of English Words. One Thousand English Words and their Pronunciation, Together with Information Concerning the Several Meanings of each Word, its Inflections and Derivatives, and the Collocations and Phrases into which it Enters*. London: Longman.

Partington, Alan. 1998. *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam: John Benjamins.

Partington, Alan. 2004. "Utterly Content in each other's Company: Semantic Prosody and Semantic Preference." *International Journal of Corpus Linguistics* 9, 131-56.

Pérez-Peña, Richard. "The Popular Newsweekly Becomes a Lonely Category." *The New York Times*, 16 January 2009. Web. Retrieved 17 January 2009.

Peters, Carol, Eugenio Picchi and Lisa Biagini. 2000. "Parallel and Comparable Bilingual Corpora in Language Teaching and Learning." *Multilingual Corpora in Teaching and Research.* Eds: Simon P. Botley, Anthony M. McEnery and Andrew Wilson. Amsterdam, Atlanta, GA: Rodopi, 73-85.

Peters, Jeremy. "Audio Pionieer Buys Newsweek." *The New York Times*, 2 August 2010. Web. Retrieved 7 September 2010.

Pfeiffer, William S. 1998. *Pocket Guide to Technical Writing*. Upper Saddle River, NJ: Prentice Hall.

Pickett, Douglas. 1986. *Defining Business English*. Talk given at IATEFL Conference, Brighton, UK. April 1986.

Pickett, Douglas. 1989. "The Sleeping Giant: Investigations in Business English." *Language International* 1, 5-11.

Porzig, Walter. [1934] 1973. “Wesenhafte Bedeutungsbeziehungen.” *Wortfeldforschung. Zur Geschichte und Theorie des sprachlichen Feldes*. Ed. Lothar Schmidt. Darmstadt: Wissenschaftliche Buchgesellschaft, 78-103.

Posteguillo, Santiago and Juan Carlos Palmer. 1997. “Layers of Specificity in Business English Discourse.” *Applied Languages: Theory and Practice in ESP*. Eds. Jordi Piqué and David J. Viera. Valencia: Universidad de Valencia, 107-118.

Quirk, Randolph, Sidney Greenbaum, Jan Svartvik and Geoffrey Leech. 1985. *A Comprehensive Grammar of the English language.* London: Longman.

Rein, Kurt. 1983. *Einführung in die kontrastive Linguistik.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Reiss, Katharina. 1995. *Grundfragen der Übersetzungswissenschaft: Wiener Vorlesungen von Katharina Reiss.* Eds. Mary Snell-Hornby and Mira Kadric. Wien: Universitätsverlag Wien.

Robins, Robert H. 1961. “John Rupert Firth: Obituary.” *Language* 37, 191-200.

Robins, Robert H. [1971] 2004. “Malinowski, Firth and the ‘Context of Situation’.” *Social Anthropology and Language.* London: Routledge.

Russell, Nicholas. 2010. *Communicating Science. Professional, Popular, Literary.* Cambridge: Cambridge University Press.

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger. 2002. “Multiword Expressions: A Pain in the Neck for NLP.” *Proceedings of the 3rd International Conference on Computational Linguistics and Intelligent Text Processing* (CICLING). London: Springer, 1-15.

Sajavaara, Kari. 1981. “Expanding the Contrastive Analysis Framework.” *Kontrastive Linguistik und Übersetzungswissenschaft.* Eds. Wolfgang Kühlwein, Gisela Thome and Wolfram Wilss. München: Wilhelm Fink, 53-63.

Schenk, André. 1994. *Idioms and Collocations in Compositional Grammars.* Diss. University of Utrecht.

Schenk, André and Patrick Rössler. 1997. *Wirtschaftsberichterstattung in Zeitschriften. Literaturbericht und Inhaltsanalyse.* München: Fischer.

Schmid, Hans-Jörg. 2003. “Collocation: Hard to Pin Down, but Bloody Useful.” *Zeitschrift für Anglistik und Amerikanistik* 51, 235-258.

Seretan, Violeta. 2011. *Syntax-based Collocation Extraction*. Berlin: Springer.

Schröder, Hartmut. 1991. *Subject-oriented Texts. Languages for Special Purposes and Text Theory*. Berlin, New York, NY: De Gruyter.

Scott, Michael. 2008. *WordSmith Tools Version 5*. Liverpool: Lexical Analysis Software Ltd.

Sinclair, John M. 1966. “Beginning the Study of Lexis.” *In Memory of J.R. Firth*. Eds. Charles E. Bazell, John C. Catford, M.A.K. Halliday and Robert H. Robins. London: Longman, 410-431.

Sinclair, John M. 1987. “Collocation: A Progress Report.” *Language Topics. Essays in Honour of Michael Halliday*. (Vol. 2). Eds. Ross Steele and Terry Threadgold. Amsterdam, Philadelphia: John Benjamins, 319-331.

Sinclair, John M. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, John M. 1998. “The Lexical Item.” *Contrastive Lexical Semantics.* Ed. Edda Weigand. Amsterdam, Philadelphia: John Benjamins, 1-24.

Sinclair, John M. 2009. *Collins COBUILD Advanced Dictionary of English*. Boston, MA: Heinle Cengage Learning.

Sinclair, John M., Susan Jones and Robert Daley. 2004. *English Collocation Studies: The OSTI Report.* Ed. Ramesh Krishnamurthy. Birmingham: Continuum.

Snell-Hornby, Mary. 1988. *Translation Studies. An Integrated Approach*. Amsterdam: John Benjamins.

Stave, Joachim. 1960. “Über den Spiegel und seine Sprache.” *Muttersprache* 70, 226-237.

Steffens, Doris. 2003. “Nicht nur Anglizismen...neue Wörter und Wendungen in unserem Wortschatz.” *IDS Sprachforum* 4, 1-9.

Steinbügl, Birgit. 2005. *Deutsch-englische Kollokationen.* Tübingen: Niemeyer.

Steiner, Erich. 1983. *Die Entwicklung des britischen Kontextualismus.* Heidelberg: Groos.

Stubbs, Michael. 1993. “British Traditions in Text Analysis: From Firth to Sinclair.” *Text and Technology. In Honour of John Sinclair*. Eds. Mona Baker, Gillian Francis and Elena Tognini-Bonelli. Philadelphia, Amsterdam: John Benjamins, 1-33.

Stubbs, Michael. 1995. “Corpus Evidence for Norms of Lexical Collocation.” *Principle and Practice in Applied Linguistics. Studies in Honour of H.G. Widdowson.* Eds. Guy Cook and Barbara Seidlhofer. Oxford: Oxford University Press, 245-256.

Stubbs, Michael. 1996. *Text and Corpus Analysis. Computer-assisted Studies of Language and Culture*. Oxford: Blackwell.

Sunderman, Gretchen and Ana I. Schwartz. 2008. “Using Cognates to Investigate Cross-language Competition in Second Language Processing.” *TESOL Quarterly* 42, 19–33.

Svartvik, Jan, ed. 1992. “Directions in Corpus Linguistics. Proceedings of the Nobel Symposium 82.” *Trends in Linguistics. Studies and Monographs* (Vol. 65). Berlin, New York, NY: De Gruyter, 79-97.

Swales, John M. 1990. *Genre Analysis. English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Swales, John M. 2000. “Language for Specific Purposes.” *Annual Review of Applied Linguistics* 20, 59-76.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work.* Amsterdam: John Benjamins.

Toury, Gideôn. 1980. *In Search of a Theory of Translation*. Tel Aviv: Porter Institute for Poetics and Semiotics.

Ungerer, Friedrich and Hans-Jörg Schmid. 2006. *An Introduction to Cognitive Linguistics*. London: Longman.

Vachek, Josef, ed. 1964. *A Prague School Reader in Linguistics.* Bloomington, IN: Indiana University Press.

van der Wouden, Ton. 1997. *Negative Contexts: Collocation, Polarity and Multiple Negation.* London: Routledge.

van Roey, Jaques. 1990. *French-English Contrastive Lexicology: An Introduction*. Leuven: Peeters.

Vatavuk, William M. 1992. *Marketing Yourself with Technical Writing: A Guide for Today's Professionals*. Boca Raton, FL: Lewis Publishers.

Ventaloa, Eija. 1984. “The Dynamics of Genre.” *Nottingham Linguistic Circular* 13, 103-123.

Vinay, Jean-Paul and Jean Darbelnet. 1958. *Stylistique comparée du francais et de l’anglais: Méthode de traduction.* Paris: Didier.

Violi, Patrizia. 2000. “Prototypicality, Typicality, and Context.” *Meaning and Cognition*. Ed. Liliana Albertazzi. Amsterdam: John Benjamin, 103-122.

Walker, David G. 1996. *The Art of Technical Writing.* New York, NY: Kroshka Books.

Weinreich, Uriel. 1979. *Languages in Contact.* 9th edition. The Hague, Paris: De Gruyter.

Wetzler, Dagmar. 2006. *Mit Hyperspeed ins Internet. Zur Funktion und zum Verständnis von Anglizismen in der Sprache der Werbung der Deutschen Telekom*. Frankfurt am Main: Lang.

Whitley, Richard. 1985. “Knowledge Producers and Knowledge Acquirers: Popularisation as a Relation between Scientific Fields and Their Publics.” *Expository Science: Forms and Functions of Popularisation.* Ed. Terry Shinn and Richard Whitley. Dordrecht: D. Reidel, 1-30.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations.* Oxford: Blackwell.

Wolseley, Roland E. 1985. *Understanding Magazines.* Ames, IA: State University Press.

Wood, Mary McGee. 1986. *A Definition of Idiom.* Bloomington, IN: Indiana University Linguistics Club.

Young, David J. 1985. “Some Applications of Systemic Grammar to TEFL or ‘Whatever Became of Register Analysis?’.” *Systemic Perspecives on Discourse Vol. 2: Selected Applied Papers from the 9th International Systemic Workshop.* Eds. James D. Benson and William S. Greaves. Norwood, NJ: Ablex, 282-294.

Yule, George. 2003. *The Study of Language.* 2nd edition. Cambridge: Cambridge University Press.

Zindler, Horst. 1959. *Anglizismen in der deutschen Pressesprache nach 1945.* Diss. Christian-Albrechts-Universität Kiel.

Dictionaries

www.businessdictionary.com

www.duden.de

www.merriam-webster.com

www.oxforddictionaries.com

www.pons.eu

Internet sources

http://de.statista.com/statistik/daten/studie/30671/die-meistzitierten-printmedien-deutschlands

http://stateofthemedia.org/ 2011/magazines-essay/data-page-4/

www.businessweek.com

www.economist.com

www.factiva.com

www.huffingtonpost.com

www.iqm.de/medien [iq media marketing GmbH]

www.lexically.net/wordsmith/version5/index.html. [WordSmith Tools]

www.mcgraw-hill.com.

www.mediakit.businessweek.com

www.mediakit.newsweekdailybeast.com/combined.html.

www.newsweek.com

www.nyt.com [New York Times]

www.observer.com

www.spiegel.de

www.spiegel-qc.de/deutsch/partner__preise/der_spiegel/ index.php.

www.thedailybeast.com/newsweek.html

www.thenation.com

www.wiwo.de [WirtschaftsWoche]

www.wiwo.de/service/digitale-wirtschaftsquellen