A Construction Grammar approach to the analysis of translation shifts: a corpus-based study

Von der Philosophischen Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen zur Erlangung des akademischen Grades einer Doktorin der Philosophie genehmigte Dissertation

vorgelegt von

Tatiana Serbina

Berichter: Universitätsprofessorin Dr. Stella Neumann
Universitätsprofessor Dr. Gert de Sutter
Tag der mündlichen Prüfung: 30.04.2015

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.
Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Dr. Stella Neumann for all the advice and support during these years. I would also like to thank Prof. Dr. Gert de Sutter for being the second reviewer of this thesis and Prof. Irene Mittelberg, Ph.D. for chairing the PhD committee.

The empirical part of the thesis has benefited from the Perl script for data extraction written by Michael Gubesch and from statistical analyses that would not be possible without valuable statistics courses given by Prof. Dr. Stefan Evert, Prof. Dr. Gert de Sutter and Bodo Winter. I would also like to thank the participants of the pilot translation experiment as well as Arndt Heilmann, Nadja Heil, Dr. Paula Niemietz and Jennifer Fest for helping me prepare and carry out the experiment. I appreciate the feedback from anonymous reviewers and editors of the article based on parts of the thesis. I would also like to thank Prof. Dr. Paul Georg Meyer who started my interest in linguistics.

Thank you Bek and Jenny for proof-reading the thesis and for sharing the office with me. Thank you Nicole, Julia D., Julia V., Antje, Elma, Aljoscha, Timo and all other friends and colleagues for listening to my presentations and asking critical questions.

Last but not least, I would like to thank Jan for his support and patience. Thank you for being there for me.
Table of Contents

List of Tables ..............................................................................................................................4

List of Figures ...............................................................................................................................6

List of abbreviations ......................................................................................................................7

1. Introduction ...............................................................................................................................9

2. Translation Studies ....................................................................................................................12
   2.1 Translation Unit ...................................................................................................................16
   2.2 Equivalence .......................................................................................................................21
   2.3 Translation shifts ...............................................................................................................23
   2.4 Naturalness .......................................................................................................................29
   2.5 Translation properties .....................................................................................................31
   2.6 Factors influencing translations .....................................................................................43
   2.7 Conclusion .......................................................................................................................46

3. Construction Grammar ............................................................................................................48
   3.1 Characteristics of Construction Grammar ......................................................................48
      3.1.1 Descriptive and explanatory adequacy of constructionist approaches ............53
      3.1.2 Language as a network of constructions ...............................................................58
      3.1.3 Corpus to cognition principle ................................................................................62
      3.1.4 Argument structure constructions ......................................................................67
   3.2 Reasons for selecting the Construction Grammar framework ........................................76
   3.3 CxG as a framework for multilingual analysis .................................................................78
      3.3.1 CxG and contrastive studies ..................................................................................79
      3.3.2 CxG and translation studies ....................................................................................81
   3.4 Conclusion .......................................................................................................................86
4. Methodology .................................................................................................................................88
  4.1 Corpora in translation studies.................................................................................................88
  4.2 Identification of constructions for analysis.................................................................................91
    4.2.1 Extraction of construction shifts..........................................................................................91
    4.2.2 Special cases of annotation and alignment............................................................................97
    4.2.3 Selected constructions.......................................................................................................103
  4.3 Types of analyses.....................................................................................................................104

5. Corpus-based analysis of construction shifts .............................................................................110
  5.1 Subject Verb Direct Object ......................................................................................................111
    5.1.1 Subj_inanimate V_volition Dobj........................................................................................137
    5.1.2 Subj_animate V_non-volition Dobj....................................................................................159
    5.1.3 Subj_human V_volition Dobj ............................................................................................169
    5.1.4 Subj_non-ref V Dobj........................................................................................................184
    5.1.5 Subj V Clause_non-fin.......................................................................................................189
    5.1.6 Translation shifts from [Subj V Dobj]................................................................................191
  5.2 Subject Verb Complement ......................................................................................................194
    5.2.1 Demonstrative wh-cleft construction .................................................................................204
    5.2.2 Further cleft, cleft-like and similar patterns.........................................................................213
    5.2.3 Translation shifts from [Subj V Compl]...............................................................................221
  5.3 Subject Verb Prepositional Object ............................................................................................223
  5.4 Shifts between the analyzed argument structure constructions..............................................230
  5.5 Conclusion................................................................................................................................233

6. Cognitive processing during the process of translation .............................................................237
  6.1 Research questions and hypotheses..........................................................................................238
  6.2 Experiment design.....................................................................................................................240
  6.3 Results and discussion of the pilot experiment .........................................................................243
  6.4 Conclusion................................................................................................................................250
7. General discussion and evaluation ................................................................. 252

8. Conclusion ........................................................................................................ 262

List of References ................................................................................................. 269

Appendix I: Covarying collexeme analysis for the English construction (C) [NP\textsubscript{research} V\textsubscript{show} clause/NP] ................................................................. 291

Appendix II: Stimuli used in the translation experiment in the order presented to the participants .................................................................................................................. 292

Appendix III: Questions asked during the retrospective interview and the corresponding stimuli with highlighted regions ............................................................................. 294
List of Tables

Table 1: Examples of constructions (adapted from Goldberg 2003,220, Boas 2014, 54)...................................... 58
Table 2: Schematic argument structure constructions ............................................................................................. 68
Table 3: [Subj V Dob] and other clause-level constructions with and without shifts ................................................. 111
Table 4: [Subj V Dob] - Contingency table for construction shift & translation direction, POPSCI ......................... 113
Table 5: [Subj V Dob] - Contingency table for construction shift & translation direction, ESSAY............................ 113
Table 6: [Subj V Dob] - Contingency table for construction & language ................................................................. 113
Table 7: Senses of the agentive verbs (adapted from Meyer 1997, 100) ............................................................... 120
Table 8: Parameters of transitivity (adapted from Hopper and Thompson 1980, 252-253) ......................................... 131
Table 9: [NP research V show clause/NP] - Significantly attracted collexemes for NP research & V show ............. 141
Table 10: [NP research V show clause/NP] - Contingency table for construction & language .............................. 143
Table 11: [NP research V show clause/NP] in BNC and DWDS ............................................................................. 145
Table 12: [NP research V show clause/NP] - Contingency table for construction & construction shift ............... 150
Table 13: [NP research V show Clause/NP] - Subject variation for the translation direction E-G ......................... 154
Table 14: [NP research V show Clause/NP] - Contingency table for construction & translation status, German ................................................................................................................................. 155
Table 15: [NP research V show Clause/NP] - Contingency table for construction & translation status, English 156
Table 16: [NP country/government V feel/think clause/NP] - Significantly attracted collexemes for NP country/government & V feel/think .............................................................................................................. 162
Table 17: [NP human V speak clause/NP] - Significantly attracted collexemes for V speak ..................................... 175
Table 18: [NP human V speak clause/NP] - Contingency table for construction & language ................................. 176
Table 19: [NP human V speak clause/NP] - Contingency table for construction & language, BNC and DWDS . . . 179
Table 20: [NP human V speak clause/NP] - Contingency table for construction & translation status, English .. 184
Table 21: [NP human V speak clause/NP] - Contingency table for construction & translation status, German 184
Table 22: [There/Es is/gibt NP] - Contingency table for construction & language, BNC and DWDS ............. 187
Table 23: [Subj V Dob] - Summary of the fixed effects on the variable of construction shift .............................. 193
Table 24: [Subj V Compl] and other clause-level constructions with and without shifts ..................................... 196
Table 25: [Subj V Compl] - Contingency table for construction shift & translation direction, ESSAY .......... 196
Table 26: [Subj V Compl] - Contingency table for construction shift & translation direction, POPSCI .......... 196
Table 27: [Subj V Compl] - Contingency table for construction & language ......................................................... 196
Table 28: Shifts from [Subj V Compl] in the translation directions E-G and G-E .................................................. 197
Table 29: Shifts into [Subj V Compl] in the translation directions E-G and G-E .................................................... 197
Table 30: [Pro dem-sing be (Pro wh...)-clause] - Significantly attracted collexemes for Pro dem-sing and Pro wh........ 207
Table 31: [Pro dem-sing be (Pro wh...)-clause] - Contingency table for construction & translation status, English ................................................................................................................................. 211
Table 32: [Subj V Compl] - Summary of the fixed effects on the variable of construction shift ............................ 222
Table 33: [Subj V Probj] and other clause-level constructions with and without shifts .................................... 225
Table 34: [Subj V Probj] - Contingency table for construction shift & translation direction, ESSAY ................. 225
Table 35: [Subj V Probj] - Contingency table for construction shift & translation direction, POPSCI ............... 226
Table 36: [Subj V Probj] - Contingency table for construction & language .......................................................... 226
Table 37: Shifts from [Subj V Probj] in the translation directions E-G and G-E ........................................... 228
Table 38: Shifts into [Subj V Probj] in the translation directions E-G and G-E ........................................... 228
Table 39: [NP_{research} VP_{show clause}/NP] & [NP_{human} VP_{show clause}/NP] used as stimuli .......................... 242
List of Figures

Figure 1: Continuum of translation properties (Hansen-Schirra and Steiner 2012, 272) .............. 42
Figure 2: Concepts of translation studies .................................................................................. 47
Figure 3: The symbolic structure of a construction (adapted from Croft and Cruse 2004, 258) ....... 49
Figure 4: English ditransitive construction .............................................................................. 75
Figure 5: German ditransitive construction ............................................................................. 75
Figure 6: Distribution of the transitive patterns in English and German originals.................... 135
Figure 7: Transitive patterns & the languages English and German (mosaic plot) .................... 135
Figure 8: Construction shifts from the transitive patterns in the translation direction E-G .......... 136
Figure 9: [NP_research V_show clause/NP] in EO and GO ................................................... 143
Figure 10: NP_study/Ergebnis and V_show/zeigen in BNC and DWDS ...................................... 146
Figure 11: [NP_study V_show clause/NP] - Distribution of construction variants in BNC .......... 149
Figure 12: [NP_Ergebnis Vzeigen clause/NP] - Distribution of construction variants in DWDS ...... 149
Figure 13: [NP_country/government V_feel/think NP/clause] - English and German originals & translations, SPEECH 164
Figure 14: [NP_country/government V_feel/think clause/NP] - With and without construction shifts in E-G .... 166
Figure 15: [NP_human V_speak Clause/NP] - With and without construction shifts in E-G and G-E .......... 176
Figure 16: [NP_human V_speak clause/NP] - Frequency distribution in BNC and DWDS ............ 178
Figure 17: [NP_human V_speak Clause/NP] - Distribution of construction variants in BNC ................. 180
Figure 18: [NP_human V_speak Clause/NP] - Distribution of construction variants in DWDS .......... 180
Figure 19: [There/Es is/gibt NP] – Frequency distribution, random samples from BNC and DWDS .... 187
Figure 20: Copular constructions in English and German originals and translations ....................... 201
Figure 21: Copular patterns & the languages English and German (mosaic plot) ....................... 202
Figure 22: Construction shifts from the copular patterns in the translation direction E-G .......... 203
Figure 23: Copular patterns & construction shifts in E-G (mosaic plot) ..................................... 203
Figure 24: [Pro_germ-sing be (Pro_wh...)clause] - Distribution across registers in BNC ................. 208
Figure 25: Inventory of argument structure constructions analyzed in the current data set .......... 234
Figure 26: Boxplots for the effect of stimulus type on fixation count (left) and fixation duration (right) 244
Figure 27: Boxplot for the effect of stimulus type on average time per keystroke ....................... 246
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>Adj</td>
<td>Adjective</td>
</tr>
<tr>
<td>adv</td>
<td>Adverbial</td>
</tr>
<tr>
<td>agt</td>
<td>Agent</td>
</tr>
<tr>
<td>AOI</td>
<td>Area of Interest</td>
</tr>
<tr>
<td>AU</td>
<td>Alignment Unit</td>
</tr>
<tr>
<td>AP</td>
<td>Adjective Phrase</td>
</tr>
<tr>
<td>BNC</td>
<td>British National Corpus</td>
</tr>
<tr>
<td>CxG</td>
<td>Construction Grammar</td>
</tr>
<tr>
<td>Clause&lt;sub&gt;Non-fin&lt;/sub&gt;</td>
<td>Non-finite clause</td>
</tr>
<tr>
<td>compl</td>
<td>Complement</td>
</tr>
<tr>
<td>cqp</td>
<td>Corpus Query Processor</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative</td>
</tr>
<tr>
<td>dobj</td>
<td>Direct Object</td>
</tr>
<tr>
<td>DWDS</td>
<td>das Digitale Wörterbuch der deutschen Sprache</td>
</tr>
<tr>
<td>E-G</td>
<td>English-German</td>
</tr>
<tr>
<td>EO</td>
<td>English originals</td>
</tr>
<tr>
<td>ER</td>
<td>English reference corpus</td>
</tr>
<tr>
<td>ESSAY</td>
<td>Political essays</td>
</tr>
<tr>
<td>ETrans</td>
<td>English translations</td>
</tr>
<tr>
<td>Europarl</td>
<td>European Parliament Proceedings Parallel Corpus</td>
</tr>
<tr>
<td>FICTION</td>
<td>Fictional texts</td>
</tr>
<tr>
<td>fin</td>
<td>Finite verb</td>
</tr>
<tr>
<td>G-E</td>
<td>German-English</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>GO</td>
<td>German originals</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>GR</td>
<td>German reference corpus</td>
</tr>
<tr>
<td>GTrans</td>
<td>German translations</td>
</tr>
<tr>
<td>INSTR</td>
<td>Instruction manuals</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative</td>
</tr>
<tr>
<td>NP</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>P</td>
<td>Preposition</td>
</tr>
<tr>
<td>pat</td>
<td>Patient</td>
</tr>
<tr>
<td>POPSCI</td>
<td>Popular-scientific texts</td>
</tr>
<tr>
<td>PP</td>
<td>Prepositional Phrase</td>
</tr>
<tr>
<td>pred</td>
<td>Predicator</td>
</tr>
<tr>
<td>Pro</td>
<td>Pronoun</td>
</tr>
<tr>
<td>probj</td>
<td>Prepositional Object</td>
</tr>
<tr>
<td>Pro&lt;sub&gt;dem-sing&lt;/sub&gt;</td>
<td>Demonstrative pronoun, singular</td>
</tr>
<tr>
<td>rec</td>
<td>Recipient</td>
</tr>
<tr>
<td>SBJV</td>
<td>Subjunctive</td>
</tr>
<tr>
<td>SFL</td>
<td>Systemic Functional Linguistics</td>
</tr>
<tr>
<td>SHARE</td>
<td>Letters to shareholders</td>
</tr>
<tr>
<td>SL</td>
<td>Source Language</td>
</tr>
<tr>
<td>SPEECH</td>
<td>Prepared speeches</td>
</tr>
<tr>
<td>ST</td>
<td>Source Text</td>
</tr>
<tr>
<td>subj</td>
<td>Subject</td>
</tr>
<tr>
<td>TL</td>
<td>Target Language</td>
</tr>
<tr>
<td>TOU</td>
<td>Tourism leaflets</td>
</tr>
<tr>
<td>TT</td>
<td>Target Text</td>
</tr>
<tr>
<td>TU</td>
<td>Translation Unit</td>
</tr>
<tr>
<td>V</td>
<td>Verb</td>
</tr>
<tr>
<td>VP</td>
<td>Verb Phrase</td>
</tr>
<tr>
<td>WEB</td>
<td>Websites</td>
</tr>
</tbody>
</table>
1. Introduction

The globalized world has a growing need for high-quality translations to ensure successful communication within multilingual and multicultural communities. For instance, the European Union lists 24 official languages as well as over 60 regional and minority languages (European Union 2014). Since the organization guarantees that all citizens of the EU can communicate with the European institutions using their native language, it requires services of a considerable number of translators and interpreters. Similarly, in the recent press release from October 2014, the German Association of Interpreters and Translators (BDÜ 2014) reports an increasing demand for language services in the economy: well-trained experienced translators play a crucial role to sustain the effective work flow particularly in the export sector. These are just two examples that stress the value of professional translations for our society and help explain the estimated 10% of growth in the language industry sector, which encompasses translation (Rinsche and Portera-Zanotti 2009, 44).

As already indicated above, for intercultural communication to be successful, it is essential for translations to meet certain standards. These could be ensured through such measures as constant improvement of training materials and development of software supporting the process of translation. These applications are based on the evolving understanding of the translation process and product, achieved through e.g. empirical investigation of previously published translations and their comparison to the originals. An analysis of these groups of texts, collected in a digital form according to certain criteria, allows researchers to identify possible factors that frequently result in differences between source and target texts, i.e. translation shifts (e.g. van Leuven-Zwart 1989, 1990; Čulo et al. 2008; Cyrus 2009). Within the descriptive translation studies, translation shifts are recognized as a natural type of relation between source and target texts. The empirical findings are used to identify conditions under which the translation shifts are likely to occur. In some contexts the shifts can be not only fully acceptable, but even desirable, leading to more natural target texts that better conform to the norms of the target language.

The present study examines translation shifts from the theoretical perspective of Construction Grammar (Goldberg 1995, 2006), where linguistic structures of various sizes
and levels of abstractness are interpreted as form-meaning pairings, i.e. constructions. In other words, constructions, be it specific lexical items or abstract argument structures, are claimed to be combinations of certain formal and functional features. They could encode information on several levels of linguistic analysis, namely phonological, morphological, syntactic on the formal pole and semantic, pragmatic and/or discourse-functional properties on the functional side (e.g. Croft and Cruse 2004, 258). It is suggested that new insights into the nature of translations can be gained by examining translation phenomena within the framework of Construction Grammar. The selected constructionist approach presents one way of examining translations: I argue that it provides us with a different perspective on the data and could lead to analyses of linguistic units that are not in the focus of other approaches but may also be relevant for the investigated phenomena. The study shows that research can also benefit from a combination of several approaches: the tenets of constructionist theories are complemented by certain theoretical concepts from Systemic Functional Linguistics.

Previous research combining translation studies and constructionist approaches is limited to a few studies, which are either mainly theoretical (Tabakowska 1993, 2013; Halverson 2003, 2007, 2010, 2013; Szymańska 2011) or deal with case studies related to individual constructions (Ebeling 1998; Rojo and Valenzuela 2013). These studies indicate the potential for enriching our understanding of translation by adopting a constructionist perspective on the questions asked within translation research. This thesis goes beyond the previous investigations as it examines a range of complex English and German constructions on various levels of abstraction with the aim to generalize beyond individual cases and assess the role of complex patterns in translations. Argument structure constructions are selected based on the frequency of their shifts in translations and the known contrastive differences on the level of grammatical functions.

The analysis of translation shifts through the perspective of constructions, i.e. construction shifts, is the main focus of the present study. One of the more specific aims is to examine the potential factors leading to translation shifts. The investigated variables include, but are not limited to, source and target language constructions, translation directions, and register differences. To analyze these factors it is essential to describe the features of the constructions within both languages considering the formal structures of the patterns, their functions, association to particular registers and frequency distributions. Moreover, the
effects of presence or lack of construction shifts on the target texts are investigated both in
terms of abstract translation properties, which distinguish translated and non-translated
texts (Baker 1996; Teich 2003), and more specific features such as differences in the
information structure.

The thesis is structured in the following way. Chapters 2 and 3 describe the general
theoretical background of the study. First the main concepts of translation studies are
introduced in Chapter 2, followed by a discussion of constructionist approaches to language
in Chapter 3. After an outline of the main tenets of this linguistic framework, we address
different motivations for the combination of Construction Grammar and translation studies
and report on previous construction-based approaches both to contrastive and translation
studies. Chapter 4 introduces the methodology of this study describing the corpus used for
the main analyses, the selection of constructions examined in the thesis and the main types
of corpus-based investigations. More specific methodological details related to corpus-based
analyses and the findings of the research are discussed in Chapter 5, which is structured in
terms of the selected constructions. In Chapter 6 the corpus-based results are then combined
with data from a pilot translation experiment using eye-tracking and keystroke logging
methods to record eye movements and keystrokes during the process of translation. This
part provides a future perspective on the possible analyses of the related research questions
connected to the cognitive processing during translation of constructions and the size of the
text production chunks, i.e. translation units (e.g. Alves and Couto Vale 2009; Carl and Kay
2011). Chapter 7 evaluates the analysis of translation phenomena within the linguistic theory
of Construction Grammar and discusses the encountered methodological challenges. A more
general conclusion and an outlook are presented in Chapter 8.
2. Translation Studies

Translation studies as an independent discipline is considered to have developed in the 1970s as a result of the pragmatic turn in linguistics, during which the areas of pragmatics and textlinguistics gained in importance (Snell-Hornby 2006, 35-47). During the so-called "cultural turn" of the 1980s, the research concentrated on the target language and the target culture to analyze the functions of translated texts (Snell-Hornby 2006, 47-67), whereas those research questions that require comparisons of target texts to their source texts moved to the background. This started to change in the 1990s with the rise of what is referred to as the "empirical turn" in translation studies (Snell-Hornby 2006, 123-128). The compilation of parallel corpora consisting of originals and their corresponding translations has contributed to renewed interest in the questions of equivalence and translation shifts (described in Chapters 2.2 and 2.3). As a result, the area of descriptive studies has advanced in particular (Cyrus 2009, 89). This development has continued in recent years and the results are visible in the vast number of publications reporting both on corpus-based and experimental research. The relevant findings are summarized in this and the following chapters.

The discipline of translation studies owes its name, as well as the so-called “research map”, to Holmes (1988, first presented in 1972 as a conference paper), who divided translation studies into two main branches, namely “pure” and “applied“. The “pure” branch is further subdivided into “theoretical” and “descriptive” studies (see also Toury 2012). Thus, already, in this early work on translation studies a descriptive approach, adopted in the present study, has been identified as a separate research strand. Descriptive translation studies are concerned with empirical analyses of authentic translation data. Chesterman (2004, 46) stresses the role of corpora in the increased use of quantitative methods applied to test scientific hypotheses. He considers the quantitative aspects to be the main contribution of descriptive studies, even though it could be added that empirical translation research often represents a qualitative analysis of selected examples. While Chesterman mentions only corpus-based research in this context, I also consider translation experiments among descriptive translation studies. Experimental research of the translation process uses psycholinguistic methods and contributes to the interdisciplinarity of empirical translation studies, a feature which is named among the advantages of the descriptive approach.
Chesterman (2004, 42-46) also argues that, similar to the prescriptive and the pejorative approaches\textsuperscript{1}, the descriptive strand of research could have an impact on translation quality. He acknowledges that translations both into non-native and native languages can benefit from quantitative findings. The former can sound more natural, if the translators use the structures that occur frequently in the texts written by native speakers, whereas the latter might be further improved, if the translators increase their awareness of the linguistic features that typically distinguish translated from non-translated language, i.e. translation properties (Chesterman 2004, 46). For instance, once the translators know what structures are frequently overrepresented in the translations, they may decide to pay special attention to the use of these features in their own texts consulting corpora to inform themselves on the frequency distributions in the non-translated language. It can be added that translation strategies frequently applied by translators, which can be observed in the parallel corpora, can be applied in the translation training.

Chesterman (2004) also observes that descriptive translation studies could still be improved by further refining the hypotheses and the operationalization of abstract variables, making the analyses more representative and clarifying the terminology used in research (see the discussion below). Furthermore, it is claimed that the empirical approach is not free from overgeneralizations. For instance, it is disputable whether the suggested translation properties are truly universal, as is sometimes claimed within the discipline (see Chapter 2.5). The present study accounts for these shortcomings by carefully considering the operationalization of variables and various factors that could influence these, as well as the level of generalizations appropriate for the individual analyses. Before the discussion can proceed to various phenomena examined within translation studies, it is important to take into account how the main object of study, namely translation, has been delineated in the previous body of literature.

There have been a number of attempts to define the notion of translation. For instance, Catford (1965, 20) defined it as “the replacement of textual material in one language (SL) by equivalent textual material in another language (TL)”. This definition presupposes the knowledge of what constitutes equivalence. While the notion of equivalence is discussed in

\textsuperscript{1} Prescriptive studies attempt to define the standards, which are argued to enhance the translations’ quality, whereas the aim of pejorative descriptions is to reveal flaws of existing translations (Chesterman 2004).
detail in Chapter 2.2, here it is sufficient to state that this definition of translation is rather narrow and requires a translation to stay close to the source text. Catford extends his definition by introducing the difference between total and restricted translation, the distinction being based on the correspondences either both on the grammar and lexis, or only on one of these linguistic levels (Catford 1965, 22). Still only the relation of the translation to the corresponding original rather than to non-translated texts written in the target language is in focus. However, as is shown in various discussions in the course of the present study, the latter perspective is equally important. According to Koller (1995, 196), translation is “the result of a text-processing activity, by means of which a source-language text is transposed into a target-language text”. The definition combines both the product- and the process-based perspectives with the focus on the product of the translating activity. Moreover, this description of what constitutes translation does not mention the relation of equivalence: a transposition into a target-language text could imply a close correspondence to the source text and/or to the target language norms. Thus, the second definition discussed here is certainly a broader one encompassing different aspects of the translation phenomenon.

Using these two definitions as examples, it is already visible how different the definitions of the same notions could be. Thus, the question arises of whether it is necessary to agree on one single definition of translation. It has been argued that definitions represent hypotheses characterized by different levels of generalization. According to this view, as long as common terminology is used and scholars are able to understand each other, it might not be necessary to refine definitions making them very rigid. The idea is rather to have a definition that can be modified to fit one’s needs and improved by new findings (Chesterman et al. 2003, 198-199). This argument has been further developed by stressing the need for “acceptance of the multifaceted character of the object of study and thus of the possibility of differing positions being equally adequate, but reflecting different aspects of the object of study” (Chesterman et al. 2003, 207, see also Hermans 2013, 75). In other words, the above definitions could exist side by side as they highlight different perspectives within the area of translation studies. At the same time, it should be stressed that difficulties can arise when the same phenomenon is referred to through different terminology (Chesterman et al. 2003; Chesterman 2005; Marco 2007).
An important reason for defining translations is to distinguish them from texts and processes that do not belong to this category (Hermans 2013, 75). For instance, does a summary of a text in the same language qualify as a translation and should it be analyzed by translation studies? Or should translations by professional translators as well as translations by students of translation or by a general group of bilinguals be treated equally (Chesterman 2004, 42; Halverson 1998, 500-501)? With regard to the first of these questions Halverson (1999, 15) suggests to recognize only interlingual translations, which have been referred to as “translations proper” by Jakobson (1959), as prototypical ones. Schreiber (2004, 273) also notes that, depending on the register, the set of text editing operations included into the notion of translation may differ. Translations of texts for specialized purposes, in contrast to translations of literary texts, may require additional adaptation of the original material. Moreover, translations by professional translators are claimed to be more central to the category of translations because “it seems that professionals acquire their status by virtue of their ability to produce texts which fit the concept of translation which is current in their time and place, at least” (Halverson 1999, 22). Thus, the notion of translation can also be defined using prototype theory: although a large number of texts could be regarded as translations, in every culture some of these are more prototypical representatives of the category than others (Halverson 1998, 1999; Hermans 2013, 81-82).

The present study adopts a rather broad definition of translation. First of all, it combines both product- and process-based perspectives on translation. In the product-based part a number of published interlingual translations for the language pair English-German are in the center of analysis, whereas the process-based part deals with interlingual translations from English into German produced by students of English linguistics during a pilot translation experiment. Moreover, the analyses involve comparisons both with the corresponding source texts and comparable target texts (see Chapters 5 and 6).

The remainder of this chapter introduces the major research questions in the area of translation studies. It begins by introducing the notion of translation unit (Chapter 2.1), which is essential for discussing various types of relations existing between analyzed sets of texts. These are relations of similarity and difference between originals and the corresponding translations, i.e. equivalence (Chapter 2.2) and translation shifts (Chapter 2.3), as well as relation of similarity between translated and non-translated texts within the
same language, i.e. naturalness (Chapter 2.4). These relations result in a number of translation properties discussed in Chapter 2.5. In Chapter 2.6 we will have a closer look at factors that can have an effect on translations. Finally, Chapter 2.7 provides a summary by showing how these translation phenomena are interrelated. The concepts introduced here are relevant to the empirical analyses presented in Chapter 5.

2.1 Translation Unit

The concept of translation unit (TU) could be defined from the product- or process-based perspective (Carl and Kay 2011, 953; Malmkjær 1998, 286). For the present study both definition types are relevant because both perspectives on translation are taken into account. However, as is shown in the discussion, it is not always easy to draw a clear line between the product- and process-oriented studies.

One of the earlier definitions of the TU is provided in a prescriptive study that suggests a number of translation procedures for the language pair English-French. Thus, the definition was developed not to describe a translation process or product, but rather to give instructions for the translation process. In this prescriptive manner the TU is defined as “the smallest segment of an utterance whose signs are linked in such a way that they should not be translated individually” (Vinay and Darbelnet 1995, 21). Even though the definition can be applied during the translation process, it can at the same time be conceptualized as a descriptive category for an analysis of a translation product (Carl and Kay 2011, 953). To differentiate a product-based TU from a TU viewed from the process perspective, Carl and Kay (2011, 953) suggest referring to the former as an alignment unit (AU). According to the authors, AUs are characterized as the smallest non-compositional segments of ST and TT that could be matched with each other. While no concrete criteria for the non-compositionality are provided, an example indicates that within their approach AUs are created mainly on the word level, either between individual words or compound nouns. For instance, the English compound Hospital nurse was aligned with the Danish noun Sygeplejeren ('nurse’), because the more explicit characterization of the nurse, realized through the token Hospital in the original, had been omitted in the translation (Carl and Kay 2011, 961).

In the present study the notion of an AU is extended beyond the so-called “translation atoms” (Carl and Kay 2011, 965) to include larger segments that can be aligned with each other (Alves et al. 2010, 126-127), for instance on the level of grammatical functions or
constructions, i.e. on the levels that are compared. AUs are then understood as units of comparison between the original and its translation (cf. Malmkjær 1998). Summarizing previous research in this area, Malmkjær (1998, 286) suggests that a clause is an appropriate unit of analysis. This position is supported through the following arguments: clause is seen as a level at which events are typically described, at which propositions are expressed and at which languages tend to share more common features than lower levels of analysis, as well as a level which could be processed as a whole (Malmkjær 1998, 286-287). The first two of these reasons establish a connection between a linguistic unit and the external world, whereas the third argument refers to a possibility of creating a semantic link between clauses in two languages (SL and TL): since individual words and phrases are more likely to undergo various shifts, the semantic correspondences on these levels are not always straightforward. In contrast, clauses are more prone to refer to the same semantic propositions (Malmkjær 1998, 286).

These functional characteristics of a clause mentioned as the first three factors are product-based reasons for the use of a clause in descriptions of originals and translations. The last reason mentioned above is more process-oriented: a clause is taken as a basis of comparison because translators potentially work with units of this size. Several studies on writing and translation (Schilperoord 2001, 83; Immonen 2006, 332; Immonen and Mäkisalo 2010, 57-58) suggest that pause length increases with rising complexity of grammatical units. For instance, pauses between phrases are longer than between words, and pauses between clauses are, in turn, longer than between phrases. Since, as described below, pauses are interpreted as boundaries between units of processing, this finding suggests that the identified size of the translation units depends on the pause threshold set by researchers. Immonen and Mäkisalo (2010, 58) show that the average pause length between clauses during production of translations is approximately six seconds, compared to four second pauses between phrases. Therefore, if researchers identify translation units based on five second pauses (cf. Alves and Couto Vale 2011), the units in the center of attention are likely to be clauses. Even though it is argued that in general a clause is suitable to compare ST and TT, particular studies may have goals, which are more compatible with smaller chunks of text. For instance, lexicographers compare words and phrases rather than larger units. At the same time, a whole text could be considered as a unit (Malmkjær 1998, 287-288).
When analyzed from the process-based perspective, TUs can be defined as “ST segments, of whatever extension or nature, which attract the translator’s focus of attention at a given time” and are typically identified as units between two pauses in the translation process (see discussion below on the pause length) (Alves and Couto Vale 2009, 254). Alves and Couto Vale (2009, 254) suggest that the ST units can be linked to the translation of these text segments and identified in intermediate and final versions of the unfolding translations. Other studies stress that TUs are cognitive units (Dragsted 2005, 49-50; Carl and Kay 2011, 953, 955).

TUs are examined with the help of experimental methods. For instance, the method of keystroke logging, i.e. recording keystrokes, mouse movements and pauses between these events by specialized software, such as Translog (Jakobsen and Schou 1999), can be employed to establish the focus of translator’s attention based on the observable process of typing a translation. In the research by Alves and Couto Vale (2009, 2011), intermediate versions of an unfolding TT are first split into micro (translation) units occurring in the translation process. In the second step those micro TUs that contribute to the final translation of the same ST part are grouped together to form macro units. Thus, all revisions affecting individual TUs are taken into account. Further classification of macro units leads to an identification of translators’ profiles.

Number and size of TUs can also be compared taking into account different independent variables, such as translator’s expertise (professional vs. student translators) and text’s difficulty. The investigation by Dragsted (2005) has shown that while professional translators work with longer and fewer TUs than students, this difference decreases when the ST is classified as having a higher level of complexity. It should be noted that her experiment involved an inverse translation task: the native language of the participants functioned as the SL. The complexity was determined by the amount of specialized terminology included in the texts. Thus, Dragsted’s study has considered the lexical level of text complexity and its effect on the working memory reflected in the keystroke logging data, whereas the qualitative investigation reported in Alves et al. (2010) exploits the Systemic Functional Linguistics (SFL) concept of grammatical metaphor to explore grammatical complexity and, among other things, its interaction with TUs. The authors discuss in detail an example of one macro unit corresponding to a ST chunk, which involves the verb
widersprechen ‘contradict’. As is shown in (1), the translation of this verb was attempted several times. The corresponding micro units express the same information through different parts of speech: the initial translation corresponds to the verb contradict, which is directly modified into the noun contradiction. In the context of the study by Alves et al. (2010), nominalizations are considered grammatically complex variants, because a semantic process is presented in a condensed way, obscuring the agent of the action (Halliday and Matthiessen 1999, 230-232). Therefore, the first of these micro units preserves the level of complexity, whereas in the next step the passage is changed towards a more complex version.

(1)  **GO:** [widersprechen] _verb_ \(\rightarrow\) **ETRANS:** [contradict] _verb_ \(\rightarrow\) [contradiction] _noun_ \(\rightarrow\) [...]  
(Alves et al. 2010)

While the studies by Alves and Couto Vale (2009, 2011) and Dragsted (2005) use only the method of keystroke logging to collect the data, the one by Alves et al. (2010), though limited to two participants, combines the data from keystroke logging and eye tracking to investigate translation shifts (see also Carl and Kay 2011). This combination of methods allows the researchers to observe not only the writing process of the intermediate versions of the TUs, but also reading of the corresponding ST unit. Therefore, it provides additional evidence for translator’s attention and allows for the interpretation of the pauses encountered in the keystroke logging data. Using these types of behavioral information, the analysis may then draw a distinction between fixation and production units and establish how often these follow each other or overlap. Carl and Kay (2011, 969) suggest that the former type of reading and writing behavior, operationalized as a number of ”alternating fixation units”, is characteristic of students of translation, whereas the latter – ”divided fixation units” involving parallel processes of fixating the ST and typing parts of the TT – are more typical of professional translators.

An important question considered in process-based studies, such as those described above, is the identification of the pause threshold necessary for separating individual TUs. Studies often experiment with different pause thresholds to see the effect of pause length on the size of the TU (Alves and Couto Vale 2009; Carl and Kay 2011). It has been noted that when a pause is too short, the translator may still be working on the same TU, the pause being a consequence of slow typing speed. In contrast, when the pause is too long, the resulting segments might correspond to several units of attention. A threshold of one second has been
suggested as an optimal one as the segments identified on its basis correspond to meaningful units (Carl and Kay 2011, 969). The longer pauses are typically associated with more meaningful TUs. Based on their studies, Carl and Kay (2011, 973) conclude that “translators process maximal segments which seem to correspond to capacity and experience (professional and students) of the translator rather than a minimum unit or a 'translation atom’”.

Another approach that allows for the establishment of an appropriate pause threshold involves adjusting the pause length to every individual subject. This approach appears more plausible, since the speed of text production and, therefore, the pause threshold that helps identify tokens belonging to individual TUs differs depending on the experiment participant. This could be based, for instance, on the subject’s typing speed, which can be established through a copy test during which the participants are asked to re-type a short text. Furthermore, Dragsted (2005) suggests considering the overall text production speed of subjects during the translation task to determine individual TUs. As the calculation is performed using the main process texts under analysis rather than a separate monolingual copy test as a baseline for comparison, this measure is influenced not only by the typing speed, but also by the translator’s expertise and other factors having an effect on speed. The method is problematic if the purpose of the study is to investigate one of these factors, e.g. translator’s expertise (cf. Dragsted 2005). However, if the study does not aim at comparing different groups of participants, this approach might fit better than the use of a copy test, as the pause thresholds derived from the latter could be too low to allow a meaningful interpretation of the resulting micro units. In other words, if an average pause between the typing of any two characters is taken as a baseline, the resulting segments are likely to encompass only individual keystroke events.

The analysis of the pilot experiment reported in Chapter 6 considers the question of whether clause-level constructions, i.e., form-meaning pairings as described in Chapter 3, could be classified as TUs from a product- and/or process-based perspective. From the former perspective, constructions on this level may well be seen as TUs. As mentioned earlier, the clause has been recognized as a suitable unit of comparison and, therefore, alignment links can be created between clauses identified in STs and TTs. From the perspective of process-oriented research, it should be tested whether translators operate with whole clauses or use
individual words and phrases as TUs during their translation process. Here it is essential to consider that the size of TUs differs depending on the level of training and experience. Thus, professional translators are able to process longer stretches of text than students of translation (Carl and Kay 2011). Among other factors that could have an effect on the length of TUs are the degree of familiarity with domain-knowledge, the degree of linguistic complexity as well as the external conditions (Alves and Couto Vale 2009; Alves et al. 2010).

2.2 Equivalence

The issue of equivalence was briefly touched upon in the previous sub-chapter. The concept can be defined as “the relationship between a source text (ST) and a target text (TT) that allows the TT to be considered as a translation of the ST in the first place” (Kenny 1998b, 77). However, as Kenny admits, the definition is circular as translation and equivalence are concepts that mutually define each other (Halverson 1997, 227; Kenny 1998b, 77; Pym 2010, 37). There has been a considerable amount of disagreement within translation studies on the definition and the existence of equivalence (Kenny 1998b, 77). Whereas some scholars define the notion of translation through equivalent relations between ST and TT (e.g. Catford 1965, 22), others have characterized equivalence between any languages as an “illusion”, which misrepresents the concerns within the discipline (Snell-Hornby 1995, 22). It can be argued that Pym (2010) tries to establish a common ground between these two opposite positions: he suggests that though there are indeed no natural correspondences between linguistic systems, an “ideal” equivalence is still persisting as a defining feature of translations. This claim should be interpreted from the perspective of the TT readers. Pym claims that the translation is accepted by the target audience only as long as the illusion of equivalence is sustained. While from this social point of view translations are interconnected with an ideal of equivalence, the perceived correspondences between the original and the translation are claimed to be created during the process of translation by the actual translators.

This argumentation shows that the opposition between equivalence as a defining feature of translations vs. its non-existence is not the only way of characterizing the notion. Equivalence can be viewed from the source text or the target text perspective, as a feature of langue or parole, as well as being a prescriptive or descriptive category (Leal 2012). Catford (1965), who takes a ST viewpoint on the equivalence relation, distinguishes between formal correspondence and textual equivalence. The former category is concerned with equivalence
between any two languages, in our case English and German, on the level of language system or langue, whereas the latter accounts for equivalence relations established between texts on the level of language use or parole.

Since this study analyzes authentic texts, it is the textual equivalence (and difference) that is considered in the empirical parts (Chapters 5-6). At the same time, I follow previous studies (e.g. Hansen-Schirra, Neumann and Steiner 2012) considering contrastive differences among the factors that play a role in the process of translation. Therefore, these differences should also be taken into account as an explanatory source in the analysis of the translation product.

It should be mentioned that the usage-based constructionist approaches to language, which are used as the main theoretical framework of this study, do not subscribe to a strict distinction between langue and parole. From this perspective usage events play an important role in the way languages are stored and processed (Croft and Cruse 2004, 292). Contrastive differences between any two languages could be interpreted as results of different inter-language frequency distributions of linguistic features.

This approach is also compatible with observations that every word or a more complex linguistic structure can be translated in a number of acceptable ways (Pym 2010, 70), that translation shifts are quite natural and common in every pair of texts, and that all structures that might be perceived as equivalent do in fact differ to some extent (Snell-Hornby 1995, 16-17) both within the same language and between different languages. Equivalence is established between authentic source and target texts. For instance, it is suggested that in a "translation (or adaptation)" of the book Alice in Wonderland into the Australian language Pitjantjatjara, equivalence relation is established between the lexical items rabbit and kangaroo. This example is used to show that "equivalence is not something permanent and guaranteed": thus, in other texts this particular equivalence will most likely not be present (Yallop 2001, 239). At the same time, there are often strong tendencies towards ways of translating certain linguistic structures. Particularly if two structures in a language pair have similar form and function, they are more likely to be used as equivalent in translations. In other cases, although the structures may be characterized by a translation shift according to the TU chosen as the basis for comparison, certain translation solutions may be highly automatized and could be seen as equivalent if examined from a different research perspective. Therefore, it depends on the TU and further details of methodology, which in
turn are determined by the particular research questions, whether linguistic patterns in the ST and the TT are identified as equivalent or not.

Moreover, in the present study the concept of equivalence is used as a descriptive rather than prescriptive category. From this perspective, research often distinguishes between two poles of equivalence, one established locally on the levels of words, phrases or clauses, and one established on a more global level. Whereas the former equivalence pole is characterized by closeness to the source text both in terms of form and corresponding functions of the individual forms, the latter one pays more attention to the naturalness of the text in the target language. As a consequence, the translator may employ structures that deviate from the original so that the TTs function in the target environment in a similar way as the STs in their original contexts. Depending on the equivalence pole, the selected translation strategy can be characterized as literal or free. Similar distinctions include the notions of formal vs. dynamic equivalence (Nida 1964) or overt vs. covert translation (House 1997). These possibilities should be understood as two extremes with different degrees of variation in between.

Another classification has been introduced by Koller (2011, 1995), who distinguishes between denotative, connotative, text-normative, pragmatic and formal-aesthetic equivalence based on the properties which are considered during a process of equivalence creation. Thus, discourse on equivalence requires not only the description of the units, which are being compared, but also the degree of similarity or difference as well as a specification of qualities which should be similar for an equivalence relation to be present (Halverson 1997, 209-210). It should also be noted that, at least from the translator’s perspective, these three elements of equivalence are closely interconnected: the size of the units between which the equivalence is established could have an effect on the other aspects of equivalence. For instance, if the relation is created between individual ST and TT words, then the translator is more interested in establishing local formal and functional correspondences rather than preserving more global pragmatic functions. As indicated above, the researcher’s perspective could yield different analyses of the same text pairs depending on the particular object of study and, therefore, it is essential to describe the methodology in a transparent manner.

2.3 Translation shifts

The originals and the corresponding translations are characterized not only through their similarities, or equivalence established by translators on one level or another, but also
through differences, or translation shifts. Among earlier investigations in this area one should name work by Vinay and Darbelnet (1995), who have identified a number of oblique translation procedures resulting in translation shifts, even though the term as such was introduced only later. Similarly to other notions in the area of translation studies, the term "translation shift" and the related concepts of "technique", "procedure" and "strategy" are not used systematically by all researchers. Thus, while some might use the terms interchangeably, others draw distinctions (cf. e.g. Chesterman 2005). Their acceptance also depends on research communities, "translation shift" being wide-spread in the English research literature (Marco 2007). The present study follows this tradition and refers to all types of differences between the original and the product of translation, or changes in the process of translation as translation shifts.

The study by Vinay and Darbelnet identifies four different types of shifts – namely, "transposition" (change of word class), "modulation" (change in perspective), "equivalence" (change in stylistic and structural means for expressing the same situation) and "adaptation" (change of situation) (Vinay and Darbelnet 1995, 36-39). Example (2) contains a number of transpositions: changes from verbs in the English original to nouns in the corresponding German translations. For instance, the verb protect is translated as the noun Schutz (‘protection’). These translation shifts are in line with the finding for the register of letters to shareholders, according to which English-German translations are characterized by shifts from verbs to other parts of speech (Čulo et al. 2008, 50).

Example (3) can be used to illustrate modulation. The subject of the original sentence the resulting database is integrated into the adverbial mit Hilfe der so angelegten Datenbank ‘with the help of the database created in this way’ in the aligned translation. This shift in the mapping of semantic information onto grammatical functions leads to a shift in perspective by changing the focus from the database to die Meinungsforscher ‘the pollsters’, i.e. the subject of the translated sentence.

(2) **EO:** They helped [protect]verb labor rights, [advance]verb the status of women and [promote]verb sustainable development. (EO_ESSAY_027 s12)
The didactic focus of the work of Vinay and Darbelnet distinguishes their study from the more theoretical description of translation shifts by Catford (1965), who, as indicated in the preceding chapter, concentrates on the level of language system, defining shifts as “departures from formal correspondence” (Catford 1965, 73). Marco (2007) observes that "[s]uch a definition implies that formal correspondence is a kind of default translation”. In other words, it presumes that pre-established relations of equivalence exist between languages and can be identified (Cyrus 2009, 90). This view could be interpreted as treating any deviations from the ideal as unwelcome, though sometimes necessary features of translations (Bakker, Koster and van Leuven-Zwart 1998, 227). Further conceptual work on translation shifts has also stressed the importance of the translation being “faithful” to the source text (Popovič 1970, 80).

A later study of shifts (van Leuven-Zwart 1989, 1990) is more descriptively oriented and empirical (Cyrus 2009, 95). It operates with so called “transemes” and “architransemes”. The former could be interpreted as translation units in a product-based perspective. These are comparable ST and TT stretches of text comprising either a predicate and its arguments or adverbials. The latter – architransemes – are assumed to capture common features of both transemes. They are defined in general terms and serve as a basis for comparison of both corresponding transemes. Establishing connections of both transemes to the common architranseme and, therefore, also a relationship between the two transemes enables researcher to classify the cases of translation shifts into three broad groups. These are “modulations” (two transemes are in a hyponymic relationship), “modifications” (two

---

2 The complete glosses are provided only for those examples where the understanding of the sentence is required for the discussion. In other cases, the glosses for the relevant parts have been included within the text.
transemes are in a relationship of contrast) or “mutations” (no relationship can be identified or one of the corresponding transemes is missing). The categories of modulation and modification can be illustrated using example (4). First of all, the example is characterized by a shift from general to specific: the noun phrase this July is translated by the prepositional phrase am 1. Juli; compared to the original this adverbial of time provides more detailed information on the time of the event.

(4) **EO:** [This July] the Cassini-Huygens spacecraft is expected to go into orbit around the solar system’s second-largest planet. (EO_POPSCI_001 s6)


Applying the descriptive framework developed by van Leuven-Zwart (1989, 1990) the relation of the English transeme to the corresponding German transeme is hyponymic as the exact date, July 1st, could be classified as one instantiation of the month July in general. Therefore, the subtype of modulation, namely specification, is present in this example. Secondly, the present tense with a future reference of the source sentence is exchanged by the past tense in the translation. Here the relationship between grammatical categories is that of contrast and, therefore, it is an example of modification. The second shift indicates that the article was originally written before the described event took place, but translated only afterwards. Therefore, the translator knew more details on this mission than the writer of the original and decided to include this information into the target text.

(5) **EO:** As one commentator put it, "Children feel the pressure ... to be sure they don't slide back. Everything's about going forward... Falling back is the American nightmare." [So if your perch is high, you have much further to fall than if your perch is low]. The amount of choice we now have in most aspects of our lives contributes to high expectations. (EO_POPSCI_005 s110-114)

**GTrans:** Ein Kommentator meinte, die Kinder stünden unter dem Druck, nur ja nicht zurückzufallen; alles drehe sich darum, voranzukommen; Zurückfallen sei der amerikanische Albtraum. Die große Auswahl, die uns heute in den meisten Lebensbereichen geboten wird, erhöht das Anspruchsniveau.

A passage in (5) includes an example of an empty link on the level of sentence (Čulo et al. 2012, 108-110): the forth sentence in the English original does not have any correspondence in the German translation. Since this sentence contains two transemes, there are two cases of mutation. Furthermore, the German translation merges the first three English sentences
into one, at the same time turning direct into indirect speech. These could also be classified as translation shifts, if we take a sentence as a level of analysis. At the same time, the number of clauses is preserved.

The discussion of these examples has shown that the cases could be classified without an initial description of an architranseme but through a direct comparison of the transemes in the two languages. While it may be argued that an identification of similarities is a prerequisite to the description of differences, it is not entirely clear how and on what level of abstraction these similarities should be listed.

(6) /His wife saw him [...] (van Leuven-Zwart 1989, 158)

ATR: wife + to see

In example (6) the description of an architranseme uses the lexical item *wife*. Even though the language pair under analysis (Spanish-English) has similar definitions of the concept “wife”, applying the same framework to languages influenced by more divergent cultures may require a description in even more general terms. The use of English for establishing this common ground may be questioned as well. Should the description always be provided in the same language? And could the choice of language have an effect on the analysis of differences between the pairs of text? Researchers are thus confronted with a typical problem of talking about language using language. To make the comparison of transemes more objective, monolingual dictionaries in both languages have been suggested as a source for descriptions of architransemes (van Leuven-Zwart 1989, 158). Another critical aspect with respect to transemes and architransemes is the possibility of identifying several types of shifts within one and the same pair of transemes. It is argued that the length of transemes is responsible for this problem: since a predicate and all of its arguments constitute the same transeme, it is sometimes possible to establish a relation of contrast between some of its elements and a relation of mutation between others (Cyrus 2009, 97). Taking these limitations into account, the application of the model could provide a detailed classification of all shifts present in the text. The most persistent types of shifts are then believed to influence the macrostructural level of the text, i.e. the level of story (van Leuven-Zwart 1989, 171).

As mentioned above, more recent studies employ corpora to investigate translation shifts. Empirical investigations into the nature of translation shifts (Munday 1998; Cyrus 2006; Čulo et al. 2008; Alves et al. 2010) have both benefited from and contributed to the compilation of
parallel corpora annotated and aligned on one or several levels. For instance, the studies by Čulo et al. (2008) and Čulo et al. (2012) use the multi-level annotation and alignment of the CroCo corpus (see Chapter 4) to analyze transpositions and modulations in Vinay and Darbelnet’s sense resulting from, e.g., shifts in grammatical functions in the thematic position, or to identify shifts heuristically. The focus of these studies is not on introducing yet another classification of translation shifts, but rather on providing quantitative findings with respect to different types of shifts. The present study employs a similar methodology to identify the shifts heuristically based on the so called crossing lines between words and grammatical functions. Crossing lines are defined as cases where “the alignment line on one level crosses that on another level” (Čulo et al. 2012, 92): translation pairs contain crossing lines between words and grammatical functions when aligned words are integrated into different grammatical functions in originals and the corresponding translations. A previous investigation of crossing lines has shown that their distribution is register-dependent and may also be explained through a number of contrastive differences (Čulo et al. 2012). Taking these results into account, the present work goes beyond them by classifying shifts in relation to constructions and analyzing reasons for and effects of these ST-TT differences in a more detailed manner. Interpretation of linguistic features identified in originals and translations may benefit from theoretical insights of a general linguistic theory provided that the research is based on suitable operationalization of the theoretical concepts in terms of observable indicators (Neumann 2013, 4). Explanatory power of a given linguistic theory may help identify additional reasons for various translation phenomena, such as translation shifts.

Translation shifts can also be analyzed from a process-based perspective accounting for multiple changes of the same ST unit in the translation process data. As discussed in Chapter 2.1, micro units identified in the translation process might contain different translation versions of the same ST element: in the example discussed above the verb widersprechen was first translated by another verb, namely contradict, which was changed into the noun contradiction and revised further at a later stage. Examining these shifts it is possible to observe whether the linguistic structures of the original are changed automatically or involve additional cognitive effort reflected through long pauses and multiple micro units (Alves et al. 2010). A systematic identification and analysis of translation shifts in the keystroke logging data requires a corpus consisting of source, target and process texts (keystroke logs)
(cf. Alves and Couto Vale 2009, 2011). Compilation of such a keystroke logging corpus requires not only alignment of behavioral information (individual keys and mouse movements) to the linguistic level (words, phrases, etc.), but also enrichment of these units with further linguistic information. For instance, POS tagging would allow querying for transpositions present in the process data and analyzing these quantitatively (see Serbina, Niemietz and Neumann 2015).

2.4 Naturalness

While equivalence and translation shifts are relations of similarity and difference that hold between translations and their corresponding source texts, translations can also be compared to comparable target texts. Also in this type of comparison one could distinguish between similarities and differences, this time between translated and non-translated target texts. The former constitute the relation of naturalness, also referred to as the relation of acceptability or textual fit (Chesterman 2004, 6), whereas the latter could be analyzed in terms of so-called translation drifts, i.e. differences between translated and non-translated texts within one language (Chesterman 2007, 58). However, as Chesterman states in a later publication (2010, 39), “we do not yet seem to have a set of drift types that could be compared to Catford’s shifts. And, as far as I know, the relation of naturalness or fit has not given rise to the kind of analytical debate that we find concerning equivalence”.

But it should be mentioned that translation properties that present an interaction between different types of translation shifts and drifts are intensively investigated (see Chapter 2.5). In fact, the starting point of this strand of research has been comparison of translations to comparable target texts. As long as no source text-related variables are taken into account, the studies on translation properties may be seen as parallel to the ones on equivalence and shifts: whereas the latter are oriented towards the source texts, the former are often oriented towards the target texts. At the same time, a number of studies have shown that source texts certainly do have an influence on the properties that characterize translations, for instance the property of shining through of the ST features (Teich 2003). This translation property could be seen as a consequence of the equivalence relations and the level on which these have been established, whereas the normalization towards the TT features (Baker 1996) could be considered as a result of translation shifts. Shining through and normalization, along with
other translation properties discussed in more detail in Chapter 2.5, are abstract features, which are related both to the ST and the TT.

For instance, in (7) we can identify a crossing line between words and grammatical functions: the same information is expressed through the direct object in the original and through the subject in the translation. On a more abstract level this translation shift results in normalization towards a structure typical of the target language. Because the word order is more fixed in English than in German, the English translation could not keep the direct object in the sentence-initial position, as it is done in the German original. Rather than changing the information structure, the translator decided to change the mapping of semantic content onto grammatical functions. It is important to stress that the shift could be interpreted as contributing to the normalization property only as long as both the SL and TL originals are taken into account. Normalization presupposes that the German original has a different structure that is changed in the translation, and at the same time it also implies that something is changed towards a norm in the target originals.

(7) **GO:** [Diese durch das IIC weltweit vermittelte Botschaft]_dobj_ haben immer mehr ausländische Unternehmen verstanden. (GO_ESSAY_020 s58)

**ETrans:** [This message, which the IIC has sent around the world]_subj_ is being understood by more and more foreign businesses.

If seen from the perspective of the target texts, equivalence on a level that leads to a translation property of shining through corresponds to certain translation drifts on this linguistic level. Similarly, translations shifts that contribute to normalization correlate to the relation of naturalness established between target texts. Due to this interaction between equivalence and translation drifts as well as translation shifts and naturalness, it is possible that no separate analysis of drifts and naturalness is required. Description of, for instance, translation shifts based on a certain TU as a basis for comparison could be at the same time a description of naturalness, unless the introduced shifts do not lead to the text being more similar to the target texts. This could be illustrated using example (8), in which the adjective *pleased* is shifted to the reflexive verb *freuen uns*: “[i]n terms of ‘markedness’, the original construction is typical of English, just as the translated construction is typical of German” (Čulo et al. 2012, 99). In other words, a translation shift has a direct effect on the naturalness of the translation.
While naturalness has been described as a rather subjective concept (Szymańska 2011, 41), applying the theoretical framework of CxG Szymańska suggests that it can be accounted for by register-specific frequencies of certain constructions and the lexical items they attract, i.e. their collexemes (see Chapter 4 on the method of collostructional analysis) (Szymańska 2011, 131). We will return to this discussion in Chapter 3.2.

### 2.5 Translation properties

The analysis of how translations are similar to or different from the corresponding STs as well as comparable originals in the same language could help us establish whether translations as texts in their own right could be characterized by a number of typical features. These features that set translations apart from the originals are typically referred to as translation universals. In this thesis, the less generalizing term of translation properties is used (see discussion below on the universality debate). The major translation properties suggested for analysis include simplification, explicitation, leveling out, (over)normalization towards the TL (Baker 1996) and shining through of the SL (Teich 2003). Systematic empirical analysis of these properties requires their operationalization, i.e. definition of the abstract properties in terms of concrete linguistic features that can be quantified. Chesterman (2010, 44) observes that the same features have been operationalized in such various ways that the results of the studies are not always comparable. Thus, it is essential to take into account the details of every research design considering the investigated features and characteristics of the corpus used for analysis. In what follows each of the main translation properties is introduced with the focus on the findings of quantitative corpus-based studies.

Simplification has been defined as “the tendency to simplify the language used in translation” (Baker 1996, 181). Its operationalization often considers rather global measures, sometimes referred to as shallow statistics (Steiner 2012). These include lexical density, i.e. the proportion of content vs. grammatical words (e.g. Laviosa 2002), (standardized) type/token ratio (e.g. Grabowski 2013), proportion of high frequency vs. low frequency words (e.g. Chesterman 2010, 44).
Laviosa 2002), repetition of the most frequent words in the corpus (e.g. Laviosa 2002, Grabowski 2013) and proportion of lemmas in the list of the most frequent words (list head) (e.g. Laviosa 2002). All of these measures are employed to investigate simplification mainly on the lexical level, i.e. to establish whether translated texts are characterized by lower number of content words and narrower range of vocabulary comprising more frequent words. While analyses of English translated and non-translated language summarized in Laviosa (2002) indicate lexical simplification in the translation product, the results discussed in Grabowski (2013) on translated vs. non-translated Polish contradict the simplification hypothesis on this level of analysis. The observed differences are either not statistically significant or point in the opposite direction. These contrasting results could be explained through different target languages investigated in the two studies, as well as different source languages involved. Thus, both studies use a comparable corpus of originals and translations in the same language, as propagated in Baker (1993, 1996), but the number of source languages corresponding to the translations included in the corpora varies. Whereas the corpus of Polish contains only translations from English, the English Comparable Corpus (e.g. Laviosa 2002) exhibits a range of source languages.

Apart from lexical simplification, also syntactic/stylistic measures of simplification have been looked into. These include a rather shallow measure of mean sentence length (e.g. Laviosa 2002; Grabowski 2013) and a more concrete syntactic feature, namely the frequency of non-finite and finite clauses (Puurtinen 2003). Also with respect to these features the results are not clear-cut: whereas, in accordance with the simplification hypothesis, the mean sentence length is lower in the Polish literary and English newspaper translations, it is higher in the English translations of literary texts. Moreover, compared to the comparable originals, the frequency of non-finite clauses is higher in Finnish translations of children’s literature. Based on her prior empirical studies with children, Puurtinen claims that the processing of these clause types requires more cognitive effort. Therefore, high frequency of non-finite clauses in translations of children’s literature is interpreted as a contradiction to the translation property of simplification.

Thus, a review of this selection of studies shows that the features selected for analysis of this translation property may depend not only on the translation status of the texts (originals vs. translations), but also on such parameters as source and target languages as well as register.
For instance, a further study (Kruger and van Rooy 2012) of translated and non-translated language, this time English used in South Africa with mainly (but not only) Afrikaans as the source language, investigates two independent variables, translation status and register, and their effect on the standardized type/token ratio as well as mean word length as measures of simplification. Whereas no significant differences have been identified with respect to the latter dependent variable, the type/token ratio is significantly higher in non-translated language indicating less lexical diversity in translations. At the same time, this variable varies depending on register (Kruger and van Rooy 2012, 58-59). It should be mentioned that the type/token ratio depends on the corpus size (deBoer 2014, 140): since the analyzed translated and non-translated corpora along with the sub-corpora corresponding to different registers have different sizes (Kruger and van Rooy 2012, 46), the comparisons of type/token ratios are not reliable. While the independent variables taken into account by Kruger and van Rooy can be studied using comparable corpora, it is admitted by the authors that the inclusion of the source texts corresponding to the translations would allow further investigating the reasons for certain distributions of the investigated features (cf. Laviosa 1998, 9; Puurtinen 2003, 147). An important step would be to establish whether there actually was a change in complexity from the source to the target text. Further analyses of the translation property of simplification are required, which would explore additional concrete indicators of this abstract feature as compared to shallow measures such as word counts. Teich (2003, 22-23) observes that such linguistic features as grammatical or semantic functions provide a better basis for bridging the gap between the abstract translation properties and the cognitive processes involved in the translation. The operationalization of simplification should also address the issue of a possible interdependence between simplification and explicitation. For instance, Baker (1996, 180) observes that substitution of clause types is an example of both translation properties, since a finite clause often specifies information left implicit in a non-finite construction. This increase in the level of explicitness is one of the reasons for a claim that the finite constructions are also simpler (Puurtinen 2003, 142-143). However, it might well be that certain features have an effect on the ease of processing without changing the level of explicitness. Identification and multivariate investigation of such features that go beyond the measures of shallow statistics is an important challenge in this area of research.
Among the suggested translation properties, explicitation, generally defined as “an overall tendency to spell things out rather than leave them implicit in translation” (Baker 1996, 180), has received considerable attention. However, also with respect to this property a variety of ways in which it is operationalized hinders a direct comparison of the findings. Different specific definitions of the property have consequences for the methodology used and the formulated generalizations (Becher 2010a, 2010c). In the following paragraphs, different directions of research and debates in this area are reviewed.

Firstly, a distinction between explicitness and explicitation has been suggested. Explicitness is understood as a property of texts within one language. The texts could either realize certain syntactic, semantic and textlinguistic relations overtly, i.e. explicitly, or use such linguistic structures where a certain amount of information is implied. In contrast, discussions of shifts from less explicit text passages in the originals to more explicit realization of the same information in the corresponding translations often refer to the process of explicitation (Baumgarten, Meyer and Özçetin 2008; Becher 2010a; Hansen-Schirra, Neumann and Steiner 2007; Steiner 2012). Thus, in these studies explicitation is conceptualized as a translation shift, whereas the resulting property of the translated text could be referred to as "explicitness". The notion of explicitness is not specific to translation properties. For instance, the explicitness-implicitness continuum has been identified as one of the dimensions of cross-cultural variation between English and German (House 2006, 252), whereas in register studies of English a distinction is made between situation-dependent and elaborated reference (Biber 1995). House (2006, 251) suggests that German texts are more likely to be explicit than English ones and that explicitness is more typical of written registers (for a summary of different approaches to explicitness see e.g. Steiner 2005; Baumgarten, Meyer and Özçetin 2008). These insights can certainly be used when identifying possible reasons for explicitation.

The idea of explicitation as a feature or a process inherent to translations goes back to the explicitation hypothesis formulated by Blum-Kulka (1986). Also in some recent studies the translation property of explicitation is mainly attributed to the nature of the translation process (e.g. Baker 1996; Olohan and Baker 2000; Olohan 2002), without accounting for

---

3 The present study follows the terminological conventions according to which all translation properties are conceptualized as shifts. For instance, we refer to the translation property of explicitation.
other possible sources of explanation (Becher 2010a). By contrast, other investigations distinguish between different types of explicitation depending on the factors that lead to this process: according to a classification by Klaudy (1998, 82-83), explicitation could be obligatory, optional, pragmatic and translation-inherent. Both obligatory and optional types could be related to contrastive differences: whereas obligatory shifts in the level of explicitness are believed to occur due to syntactic and semantic differences, optional shifts are linked to textlinguistic and stylistic variation between the two languages. Pragmatic explicitation is explained through cultural differences: for instance, certain proper names may be familiar only in one of the two cultures and require a more explicit referent in the translation. Finally, according to Klaudy (1998), translation-inherent explicitations are those that can be explained by the translation process. Studies that consider different factors playing a role in the process of explicitation do not account for all instances of explicitation found in a corpus as process-related (Baumgarten, Meyer and Özçetin 2008; Becher 2010c; Hansen-Schirra, Neumann and Steiner 2007). Since the acceptability of linguistic structures in different languages is a matter of degree, one might argue whether the distinction between obligatory and optional shifts is appropriate. For instance, even if an equivalent translation might result in a very marked option in the target language, the shift is not necessarily obligatory, since in these cases the use of the corresponding structure is not natural but still possible. At the same time, some differences are more fine-grained: for instance, comparable structures can be typical of both languages but used more or less frequently depending on the register. Here it might be misleading to apply the obligatory-optional distinction: while in this case the shift would be probably considered to be optional because both structures are accepted in the two languages, observing register-specific frequency distributions increases the naturalness of the produced translation.

The original explicitation hypothesis has mainly been operationalized through “more complete/less economical surface realisation in translation” and “more explicit relations between conceptual propositions in text” (Kruger and van Rooy 2012, 42). Thus, the first sub-type of explicitation has been studied in terms of optional structures such as the complementizer that (Olohan and Baker 2000; Olohan 2002): the insertion of that has been shown to be more frequent in English translated from a range of languages than in the original English texts. This finding is interpreted in favor of the explicitation hypothesis
(Olohan and Baker 2000). However, as mentioned above, the study has considered only the variable of translation status as a potential explanation for this phenomenon.

The second type of explicitation features relates particularly to cohesion in originals and translations. This operationalization, introduced in the original study on explicitation (Blum-Kulka 1986), is still frequently found in recent work in this area (e.g. Becher 2011; Hansen-Schirra, Neumann and Steiner 2007). For instance, in Becher (2011) explicitation is studied through analyses of the cohesive device “conjunction” (Halliday and Hasan 1976), also referred to as connectives (conjunctions, linking adverbials or particles), in the corpus containing business texts: the results indicate that there is more explicitation in translations from English to German than from German to English. These findings are interpreted taking into account the level of explicitness in the two languages under analysis. First of all, Becher suggests that German business texts require more cohesive explicitness than English texts belonging to this register. Secondly, he identifies the general lexicogrammatical preferences of German as a reason for using additional connectives (Becher 2011). Another study (Hansen-Schirra, Neumann and Steiner 2007) investigates explicitness and explicitation based on the main types of cohesive devices as described by Halliday and Hasan (1976). Hansen-Schirra et al. explain examples of explicitation and implicitation mainly through contrastive differences (between English and German) and register characteristics. At the same time, the authors suggest that higher lexical density (as a potential indicator of lexical cohesion, in particular of repetition) in the register of fictional texts for the translation direction English-German could be attributed to the translation process. Moreover, their study shows how the analysis of explicitness and explicitation can be linked to lexicogrammatical features associated with the three metafunctions distinguished in SFL (Hansen-Schirra, Neumann and Steiner 2007, 250).

Related research within this theoretical framework analyzes the discussed phenomena through the notion of grammatical metaphor: nominalizations, associated with grammatical metaphor, are more condensed and thus often include more implicit realizations of information than their clausal counterparts. The process of explicitation is identified in the translation process when the translator selects more congruent, i.e. less metaphorical, unmarked structures than those that could be found in the corresponding originals. When the level of grammatical complexity is reduced during the translation process, for instance
by translating a nominalization through a clause, the resulting structures can be more explicit (Steiner 2001; Hansen 2003; Hansen-Schirra, Neumann and Steiner 2012).

Moreover, a quantitative study by Kruger and van Rooy (2012) combines the features related to omission (the use of complementizer that and contracted forms) and cohesion (frequency of linking adverbials) and investigates an effect of two independent variables, namely translation status and register. With respect to the omission of the complementizer that, the results of this multivariate study show a significant effect of both variables as well as an interaction effect between them. The distribution of other variables reveals some tendencies, but the only other significant difference is identified between register and the use of linking adverbials. In other words, explicitation in translations is reflected in particular through the use of the optional that, and varies depending on the register. Thus, this study demonstrates that explicitness is not a one-dimensional phenomenon, but rather depends on a number of variables.

Examining explicitation from the translation process perspective, Englund Dimitrova (1995) observes that this translation shift is also influenced by the level of experience in translation: in a translation experiment using think-aloud protocols, (language and translation) students and professional translators inserted certain types of cohesive conjunctions during different phases of translation and did so with different amounts of certainty. The study suggests that certain types of explicitation, referred to as “norm-governed”, become automatized, i.e. entrenched, and are applied systematically depending either on experience or individual preferences. In other cases, explicitation is employed as a problem-solving strategy (Englund Dimitrova 1995). This number of factors having an effect on explicitation certainly challenges the claim that it is a universal translation-inherent process. At the same time, this could be the result of different and potentially too broad definitions of the analyzed phenomenon adopted in various studies. Thus, similarly to studies of other translation shifts and properties, the research in this area should carefully consider the required methodology and its implications for the generalizations that can be formulated.

Normalization, defined as “a tendency to exaggerate features of the target language and to conform to its typical patterns” (Baker 1996, 183), is a parallel concept to the law of growing standardization (Toury 2012, 303-310). According to this hypothesis, translators (over-)use the norms of the TL and are more conservative than the writers of the original texts: this is
done to avoid creating an unnatural text by staying too close to the original (Baker 2007). The property can be operationalized through the (over-)use of typical grammatical structures, collocations and punctuation in the translations (Laviosa-Braithwaite 1998). Thus, it has been shown that an analysis of normalization could be performed by assessing how creative individual collocations are in the originals and investigating whether this level of originality is kept in the corresponding translations (Kenny 1998a, 517-519).

Moreover, one of the studies on normalization has considered idioms: these prefabricated units are particularly interesting from the perspective of normalization/conservatism because they are typical of natural language use, but at the same time are often creatively changed in the originals, and individual instances can be, among other things, rather rare and/or informal. Based on several examples of idioms, the study by Baker has shown that if an expression is strongly associated both with idiomatic and literal meaning, translators tend to use them only in the literal sense (Baker 2007). While this investigation represents only the first step in the analysis of normalization and cannot be interpreted as strong evidence for or against the underlying hypothesis, this operationalization can be further applied to study normalization by comparing the proportions of idiomatic language (fixed or creatively modified) in translations and the comparable originals.

In Hansen (2003) this translation property is examined through an operationalization of linguistic features typical of the register of fictional texts and a comparison of their frequencies in translated English (from several source languages) and English originals. The results reveal that the typical features are normalized, i.e. occur more often in translations, whereas the untypical features are anti-normalized, i.e. also occur more often in translations. The phenomenon of anti-normalization (or translation drifts) might be at least to some extent explained through the tendency towards "shining through" of the source language (see below), but, since the corpus used in Hansen’s study does not contain the corresponding originals, this source of explanation cannot be used. Thus, as mentioned by Hansen, her study considers different variables that could have an effect on translations for different objects of investigation: the analysis of translation properties concentrates only on the register characteristics as an explanatory factor (Hansen 2003, 158-159).

In contrast, the studies by Kruger and van Rooy (2012) and de Sutter, Delaere and Plevoets (2012) consider several independent variables and their effect on the use of coinages,
loanwords and lexical bundles as well as neutral/formal language. As mentioned above, the study by Kruger and van Rooy (2012) investigates the features across six registers in translated and non-translated English: with respect to coinages and loanwords, only the register of popular scientific writing is characterized by an interaction with the translation status indicating a tendency towards normalization. For the use of lexical bundles (or trigrams) the variable of register but not the translation status or their interaction has a significant effect (Kruger and van Rooy 2012, 56-58). The study by de Sutter, Delaere and Plevoets (2012) also considers the variables of register and translation status, but distinguishes between three levels for the second one, namely Dutch translations from French, Dutch translations from English and Dutch originals. Therefore, the source language is also taken into account. De Sutter, Delaere and Plevoets show that the investigated registers have different proportions of formal linguistic items and that there are differences between translated and non-translated language. It is important to mention that there is also an effect of the source language: thus, the conservatism hypothesis could be confirmed for the translations from French, but not from English (de Sutter, Delaere and Plevoets 2012). This particular finding underlines the need for a more differentiated investigation of this translation property and translated language in general.

It has also been suggested that features of the source language may "shine through" in translations. The property of shining through as the opposite of normalization is said to characterize translations which are "oriented more towards the source language" (Teich 2003, 145). It can be related to the law of interference (Toury 2012, 310-315) and the phenomenon of transfer (Mauranen 2004, Toury 2012). In the present study the notion of shining through is used to refer to quantitative differences with respect to certain concrete linguistic features: in other words, when the feature under analysis is present in both languages, but its frequency increases in translations due to influence of the source texts (Teich 2003, 155-159). In contrast, interference occurs when a feature does not exist in the TL but is introduced in translations due to literal translations. It should be noted that the difference between the two concepts is gradual (Mauranen 2004, 80). For instance, an ungrammatical feature can slowly become more accepted through its repeated use either in translations or other language contact situations. Hansen-Schirra and Steiner (2012, 274-275) suggest that semantic features of the originals are more likely to
shine through in translations than syntactic ones, which tend to be normalized and are, therefore, less prone to be the subject of diachronic change due to language contact. Mauranen (2004, 66-67) notes that a certain degree of SL influence is natural for translations, as these texts are characterized by interaction between different languages and involve bilingual processing. In this respect the research questions posed in translation studies are similar to the ones in the area of (second) language acquisition, even though the former typically deal with influence of the L2 onto the L1 and the latter typically consider the opposite direction (an effect of L1 onto L2).

The two properties of normalization and shining through do not necessarily contradict each other since they can co-occur affecting different linguistic features (Teich 2003, 207). This finding is based on Teich’s analysis of grammatical features characterized by contrastive differences and typical features of the popular-scientific register for the language pair English-German. The use of a parallel bidirectional corpus allows for a direct comparison between the use of these features in originals and the corresponding translations to study the effect of the SL on the TL in both translation directions. Furthermore, a quantitative analysis is combined with qualitative discussions of selected examples, in which normalization or shining through can be observed (Teich 2003). It is also possible that the same linguistic feature is normalized in some cases and shines through in others: as a result, the translations exhibit properties of both the source and the target languages (Neumann 2013, 273).

The study by Mauranen (2004), the aim of which is to show that the property of shining through⁴ may be universal rather than attributed to specific language pairs, uses a combination of comparable corpora. Similar to the study by de Sutter, Delaere and Plevoets (2012), it analyses a corpus of (in this case Finnish) originals and Finnish translations distinguishing between translations from Russian, from English and from a mixture of different language. Applying the method of word list ranking the author shows that all translated texts differ from the originals, whereas translations from individual languages can be distinguished from each other. "This implies that transfer is one of the causes behind the special features of translated language" (Mauranen 2004, 78). However, it is essential to

⁴ The author uses the term "transfer", but, for reasons of terminological consistency within the present study, the term "shining through" will be used to refer to cases in which the frequency of a certain linguistic feature existing in the TL is changed due to influence of the SL (Teich 2003).
stress that it is not universal since the effect of shining through differs depending on the specific language pairs. The study can be further improved by considering the frequencies of the analyzed lexical items in the English and Russian originals to analyze whether their use in the translations can be explained through contrastive differences. Moreover, extension of this analysis to other registers and different linguistic features will probably show even more variation, as is the case with other translation properties.

The property of levelling-out has been defined as “the tendency of translated text to gravitate towards the centre of a continuum” (Baker 1996, 184) or, in other words, for the translations to show less variation with respect to the use of different linguistic features than the originals. According to Laviosa (2002, 72), this property could be tested by calculating variation for two continua, namely for translated texts and non-translated texts. It is assumed that the realizations of linguistic features will be clustered closer together in the former group of texts. An example of how this hypothesis could be operationalized is provided by Grabowski (2013): his study compares the standard deviation of the standardized type/token ratio and of sentence length in Polish originals and translations. Moreover, Principal Component Analysis and Cluster Analysis are used to study whether translations and originals form two distinct groups. His results (with the exception of the type/token ratio) have confirmed the levelling out hypothesis, but it is noted that quantitative studies should be supplemented by more fine-grained analyses (Grabowski 2013, 275).

It has been suggested that the properties of normalization and shining through form a continuum (see Figure 1): judging by the use of some concrete linguistic features, the translations could be closer to one of the extremes, namely the SL or the TL norms, thus exhibiting shining through or normalization respectively. However, the adoption of the TL norms might be exaggerated resulting in over-normalization. When the features used in the translations are situated in the center of the continuum, the translated texts are considered
to reflect the property of levelling-out (Hansen-Schirra and Steiner 2012, 272).

<table>
<thead>
<tr>
<th>SL norms</th>
<th>TL norms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybridization, levelling-out, etc.</td>
<td>Normalization Over-normalization</td>
</tr>
<tr>
<td>Shining through</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1: Continuum of translation properties (Hansen-Schirra and Steiner 2012, 272)**

As mentioned above these features of translated language have been introduced as translation universals (Baker 1996). This claim has been criticized as being a rather broad generalization: the review presented in this chapter shows that these properties are influenced by a variety of factors, such as register, specific languages involved, translation direction and translator’s experience. Further suggested variables include other translator-dependent properties, time period during which translations were produced, cultural features, etc. (Chesterman 2010, 44-45). To this list of typological and socio-cultural factors, one could also add a variety of cognitive aspects having an influence on the translation process and ultimately on the translation product (Toury 2004, 15, see also Halverson 2003, 2010, 2013). Thus Chesterman (2004, 43) suggests that patterns should be searched in translations on a lower level of generalization, taking into account these possible factors. Furthermore, Toury (2004, 20) observes that the postulation of translation universals is ambiguous: on the one hand, it could mean that only these features, such as explicitation or simplification, and no (or at least less) cases without explicitation or even instances of implicitation occur in translations, while on the other hand, it could simply be a statement that explicitation is present in every translation. While the former claim is too broad as translations are often characterized by both explicitations and impicitations, the latter has been considered to be self-evident. It is the first interpretation suggested by Toury that is taken as the basis of various operationalizations of the properties. Moreover, it has been argued that only those instances of e.g. explicitation that are introduced into translations due to the translation process itself rather than other factors could qualify as translation universals (Hansen-Schirra and Steiner 2012, 274).
Considering these points of criticism, it should be mentioned that though initial research of these phenomena has been triggered by a broad generalization concerning their universality, the operationalizations, methods and classifications of the translation properties continue to be refined, as shown by the literature overview presented in this chapter. In particular, the multivariate methods seem promising as these increase the explanatory potential of the features found in the translated texts.

2.6 Factors influencing translations

In the preceding sub-chapters it has been established that translations may be similar to or different from both the source texts and comparable texts in the target language. A combination of these similarities and differences results in translations being texts with their own characteristic features. Discussions of multivariate analyses of translation properties have already indicated a number of factors having an effect on translation. This sub-chapter presents these and other reasons in a more systematic manner by considering three major sources of explanation, namely contrastive differences for a given language pair, register characteristics in both languages and the translation process itself (Steiner 2001; Hansen-Schirra and Steiner 2012).

The first of these factors, namely linguistic features of the two languages, have been assumed to play a role in the establishment of equivalence relations (Koller 1995, 196) and have been considered in a number of empirical studies in relation with the resulting translation properties (e.g. Hansen 2003; Teich 2003; Hansen-Schirra, Neumann and Steiner 2007; Becher 2011). Thus, the properties of shining through and normalization could be studied by investigating a number of contrastive differences to see whether the translations will be closer to the SL or the TL end of the continuum (Teich 2003). Such differences could also lead to explicitation, as is shown in Becher (2011) or Hansen-Schirra, Neumann and Steiner (2007). The studies could either start from the contrastive differences that were previously established in the literature (for the language pair English-German see Hawkins 1986; Teich 2003; König and Gast 2009; Königs 2011) formulating hypotheses based on these differences, or they can operationalize translation shifts and/or the resulting translation properties in a different way using the contrastive accounts to explain the quantitative findings.

Furthermore, the translations may differ from originals due to differences limited to a particular register. Except for situations when a register exists only in one of the two
languages, which is rather unlikely when dealing with such close languages as English and German (Hansen-Schirra and Steiner 2012, 256), both the ST and the TT are likely to belong to the same register. However, the register could be characterized through specific linguistic features, different for the respective languages. A study concentrating on register-induced features (Neumann 2013) has considered three main variables, namely language, register and translation status, i.e. originals or translations. Various constellations of these variables result in different types of comparison, such as intra-language comparison of registers to establish typical features of registers for each of the languages; cross-linguistic comparison of registers to determine whether the originals in the two languages belonging to the same register are similar or different from each other with respect to the analyzed linguistic features; as well as intra-language and cross-linguistic comparison of texts depending on their translation status. Here it is again important to control for the variable of register, while comparing originals and translations (Neumann 2013, 102-105). Different comparison types are required to establish appropriate baselines for comparison: it is observed that since the same linguistic feature may play a different role in different languages, a cross-linguistic analysis relates to the values obtained in the intra-language comparison. Moreover, in this study the variation due to translation status is investigated against the background of the extensive analyses of the register features in the originals (Neumann 2013). However, the majority of empirical research in translation studies has considered only the types of comparison involving translations (comparing translations to the corresponding originals or non-translated texts in the same language). Thus, the studies are often performed using parallel texts belonging to one register, for instance register of popular-scientific texts (Teich 2003), register of literary texts (Hansen 2003) or register of letters to shareholders (Becher 2011). It is important to keep in mind that the results of these analyses could be limited to the particular register and should not be generalized to all originals and translations. Moreover, intra-language investigations are often conducted comparing translations and comparable originals within the same language, for one or several registers, typically to study translation properties (Kruger and van Rooy 2012, de Sutter, Delaere and Plevoets 2012).

The third factor listed above is the translation process. This source of explanation is rather heterogeneous and can be in turn sub-divided into a number of variables playing a role during translation and having an effect on the final product. One of the aspects named in the
literature (Steiner 2001) is understanding on the part of the translator: the translation process is modeled as unpacking the information present in the ST to a more explicit version and then repacking it either completely or only to some extent in the corresponding TT. Thus, previous studies suggested that whether the TT passage, compared to the ST, is more explicit, shows the same level of explicitness or is even more implicit depends on the translator-related variables, namely his/her level of proficiency or domain experience, certain external factors such as time pressure, and linguistic peculiarities, for example the level of similarity between the languages involved or the degree of lexicalization of a ST unit (Steiner 2001, Hansen 2003, Hansen-Schirra and Steiner 2012). In other words, understanding is interconnected with other process-related variables.

It has been claimed that contrastive differences existing between a language pair result in obligatory translation shifts, the changes due to understanding are optional and register-related differences may or may not lead to various shifts (Hansen-Schirra and Steiner 2012, 274, see discussion on the obligatory-optional distinction in Chapter 2.5). At the same time, it is observed that if the differences between the two languages are more semantic rather than formally syntactic in nature, the translations are more likely to adhere to the norms of the SL, and are therefore not necessarily affected by translation shifts (Hansen-Schirra and Steiner 2012, 274). For instance, in German there are limitations on semantic roles that could be mapped onto subjects, whereas in English the range of acceptable semantic roles is wider (Hawkins 1986, 57-61; König and Gast 2009, 108-109). However, in translations from English to German the subjects of the German clauses sometimes realize the semantic roles that are more typical of English: the linguistic feature of English shines through in German translations (Hansen-Schirra and Steiner 2012, 274, see Chapter 5.1.1). The influence of the translation process is sometimes determined by excluding other factors (Klaudy 1998, 82-83). However, it is more reliable to study cognitive processing through psycholinguistic experiments (e.g. Hansen 2003, Alves et al. 2010) in order to collect online information during the actual process of translation. Moreover, the approach also allows controlling more factors or considering the variation across several parameters applying multivariate statistics.

Among other process-related factors influencing translation one should certainly name the strategy employed by the translator, i.e. orientation towards formal or dynamic equivalence.
(Hansen-Schirra and Steiner 2012, 261), translator's attempts to optimize the text, for instance by adding cohesive devices (Becher 2011, 40-41) or by modifying the text based on her/his aesthetic and stylistic norms (Koller 1995, 196). Furthermore, on the more global level, the overall purpose of the translation, determined by translation brief, and other external conditions of translator's work (Koller 1995, 197) play a role in shaping the final translation product.

2.7 Conclusion

Figure 2 summarizes interaction of the concepts discussed in this chapter. Along certain parameters translations can be similar to the corresponding originals, i.e. (to some degree) equivalent, whereas other features taken into account during analysis of the same pairs of texts reveal differences, i.e. translation shifts. Furthermore, translations are typically characterized both by similarities and differences in relation to the comparable originals in the target language, which could be referred to as naturalness and translation drifts respectively. Both types of relations between translations and originals in source and target languages result in a number of features typical of translations, i.e. translation properties. Since the recent empirical evidence reported above questions the existence of universal translation-inherent features, the notion of translation properties is used here in a more neutral way to refer to features that are not translation-inherent or -unique but are still characteristic of this group of texts. As suggested above, these translation properties can be identified only in relation to both the corresponding source texts and the comparable non-translated target texts. Translation process and product are influenced by various factors, which can be considered as possible explanatory variables in the studies on translation properties.
Figure 2: Concepts of translation studies
3. Construction Grammar

Translations can be examined not only from a linguistic but also a variety of other perspectives, such as historical, sociopolitical or socio-cultural (Malmkjær 2005, 5). For instance, the work by Pym (2010) is an example of a sociological perspective on equivalence as it focuses on the acceptance of translations within society. Moreover, the majority of studies, which could be classified as linguistic in its broad sense of studying language do not adopt a theoretical framework of a general linguistic theory. While some operate on the level of shallow statistics (e.g. Laviosa 2002; Grabowski 2013), others are limited to the theoretical notions of traditional grammar, such as parts of speech or grammatical functions, without their integration into a unified linguistic theory (e.g. Ćulo et al. 2012). However, there is a number of studies that investigate translations using, for instance, concepts from Systemic Functional Linguistics (e.g. Steiner 2002; Neumann 2012; Teich 2003) or Relevance Theory (e.g. Gutt 1991; Alves and Gonçalves 2013). The present study adopts the perspective of Construction Grammar (CxG) (Goldberg 1995, 2006) drawing on related work in this area, e.g. by Szymańska (2011) and Halverson (2007). As discussed in Chapter 2.3, looking at translation phenomena from a perspective of a particular linguistic theory may extend the explanatory potential of the research. Therefore, the present study aims at enriching our understanding of various translation phenomena, such as translation units, translation shifts and translation properties, by examining these from the point of view of constructionist approaches.

The main assumptions of CxG are described in detail in Chapter 3.1. It is followed by a discussion of reasons for selecting this particular linguistic theory (see Chapter 3.2). Chapter 3.3 presents the state of the art in the application of this theoretical framework to cross-linguistic analyses as well as to the study of translations. Finally, the main ideas of this chapter are summed up in Chapter 3.4.

3.1 Characteristics of Construction Grammar

The underlying theoretical framework of Construction Grammar (CxG) encompasses a family of cognitive linguistic approaches to grammar (Croft and Cruse 2004; Fischer and Stefanowitsch 2008; Goldberg 2013), which share the assumption that language consists of
constructions, i.e. form-meaning pairings, interpreted as the basic units of analysis (Goldberg 1995, 6; 2006, 5). As can be seen from this definition of construction, CxG is a non-modular approach to language description: it is argued that form and function are two poles of a linguistic unit, which cannot be analyzed separately (e.g. Langacker 1987, 12; Bybee 2010, 9). The link between these two poles is considered to be symbolic, as it is established by convention (see Figure 3).

Figure 3: The symbolic structure of a construction (adapted from Croft and Cruse 2004, 258)

Thus the analyses of constructions should consider phonological, morphological, syntactic, semantic, pragmatic and textlinguistic factors. However, not all of these are equally important for an adequate description of every individual construction. Moreover, not all levels have been taken into account in the previous analyses: for instance, Fischer and Stefanowitsch (2008, 9) note that interactional features of constructions have not been sufficiently investigated. It has to be mentioned, however, that recent constructionist research, especially of German spoken discourse (e.g. Günthner and Imo 2006), but also to some extent of English (Bergs and Diewald 2009) considers also this dimension.

What are the motivations for treating the linguistic knowledge as an inventory of constructions? One of the earlier descriptions of CxG recognizes the necessity to account not only for the core phenomena typically described by grammar, but also for the idiomatic expressions, whose form and/or meaning is not fully predictable. Fillmore, Kay and O’Connor (1988, 504) claim that their “proper understanding will lead us to significant insights into
the workings of language in general”. Grammatically complex idiomatic patterns, which cannot be split into their components without a change in meaning, are analyzed in the same manner as individual lexical items, i.e. as combinations of syntactic form with semantic and pragmatic features (Fillmore, Kay and O’Connor 1988). A possibility to generalize from these analyses of idiomatic structures to more regular ones presupposes that the whole language could consist of constructions. The non-compositionality of constructions, which is present in idioms, has been considered necessary in order to postulate a new construction (Goldberg 1995, 4). The importance assigned to non-compositionality as a defining feature of constructions, as well as the understanding what it means for a pattern to be (non-)compositional differs to some extent in different studies within the framework (Fischer and Stefanowitsch 2008, 8). Croft and Cruse argue that

[c]onstructions other than idiomatic phrases are compositional, that is, the meanings of the parts of the construction are combined to form the meaning of the whole construction. The reason why they must be represented as independent constructions is that semantic interpretation rules associated with the construction are unique to that construction, and not derived from another more general syntactic pattern (2004, 253). The quote indicates that whereas a construction as a whole is said to have a certain function, it can still be analyzed into its parts (Bybee 2010, 36). It is claimed that this is true about schematic constructions like [NP be Adj], as well as the so-called idiomatically combining expressions like spill the beans, where spill is associated with “divulge” and the beans with “information” (Croft and Cruse 2004, 252-253). Here the distinction between compositionality and analyzability is useful: the former refers to the degree to which the meaning of the whole is transparent, and the latter describes the possible identification of its elements and their role in the constructional function (Bybee 2010, 44). Thus the phrase spill the beans in its figurative sense would be an example of a non-compositional but analyzable construction. This combination of (non-)compositionality and analyzability, which considers the whole construction as a chunk with internal structure, is essential for an idea that the findings based on more lexically filled patterns can be extended to the more general ones, as is mentioned above.

Subsequent development of CxG further stresses the existence of meaning on the level of clauses independently of the lexical items that are interchangeably used to fill the slots of the construction. In particular, the fact that the same verbs can often occur in combination with
a range of different argument types has been identified as an argument in favor of constructionist approaches. Whereas within the projectionist approaches the solution is the postulation of polysemous verbs (see Chapter 3.1.4, p. 67, Footnote 9 for a description of the projectionist approaches in relation to argument structures), CxG claims that these verbs are integrated into various constructions (Goldberg 1995, 11, Bencini and Goldberg 2000, 640; Levin and Rappaport Hovav 2005, 189-193).

(9)  a. He sliced the bread. (transitive)
 b. Pat sliced the cattors into the salad (caused motion)
 c. Pat sliced Chris a piece of pie. (ditransitive)
 d. Emeril sliced and diced his way to stardom. (way construction)
 e. Pat sliced the box open. (resultative) (Goldberg 2006, 7)

(10) a. He sneezed. (intransitive)
 b. He sneezed the napkin off the table. (caused motion)

For instance, according to CxG, in the set of examples presented in (9) the verb *slice* occurs in five different argument structure constructions taking different number and types of arguments. The point can be further supported by the famous example of the verb *sneeze*: the verb is typically considered to be intransitive, because it takes only one argument, namely the subject, as is illustrated though (10a). However, Goldberg suggests that (10b) is also a perfectly acceptable English sentence, which can be easily interpreted by a hearer. In this and similar cases it is assumed to be more plausible that the verb is used in the so-called caused-motion construction rather than it having the sense 'X causes Y to move Z by sneezing'. Moreover, assigning certain meaning to a pattern rather than an additional sense to every verb occurring in this pattern contributes to a more parsimonious account of language organization (Goldberg 1995, 9-16; 2003: 220-221; 2006: 6-7) (see further discussion in 3.1.4).

Constructionist approaches assume that every utterance is a combination of several constructions (Goldberg 2003: 221; 2006: 10, 21-22; Fischer and Stefanowitsch 2008: 7): to name just two constructions involved, every clause integrates specific meaning of a verb as well as the semantic component contributed by an argument structure construction (Goldberg 1995: 16). For this integration to be successful, the semantic information of these two construction types has to be coherent: in other words, functional features of the constructions have to be compatible. Due to this relation, constructions are said to attract
certain types of verbs and repulse others. Stefanowitsch and Gries (2003, 236) suggest that these associations could be seen as additional proof that sentence patterns are meaningful (see Chapter 4). In some cases, e.g. in (10b) above, the valence pattern of the verb differs from the number of arguments licensed by a construction it is used in. This mismatch leads to coercion of the semantic features of the verb to correspond to those of the construction. This process is also referred to as an override principle (Michaelis 2004, 25). However, staying with this example it is important to mention that the semantics of the verb sneeze does have some properties in common with the caused-motion construction: the action of sneezing involves movement of air that could lead to motion of light objects.

Moreover, a construction-based view of language allows researchers to account for processing of linguistic structures in terms of domain-general cognitive mechanisms, such as chunking or analogy. For instance, it is argued that repetition of multi-word sequences leads to their chunking: through this process the sequence acquires the features of a unit which can be stored in memory and extracted as a whole during speech production or perception (Bybee 2010, 34) (see also 3.1.3). In usage-based constructionist accounts high frequency of linguistic structures, linked to unit-status of these expressions in the speaker’s mind, can lead to postulation of new constructions. Therefore, at least one of the two criteria, non-compositionality or high frequency, has to be fulfilled for a pattern to be considered a construction:

Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency (Goldberg 2006, 5).

In this account the notion of non-compositionality is understood to involve not only functional but also formal unpredictability. In comparison to earlier versions of CxG by Goldberg (1995, 4), in this description non-compositionality is not the only possible prerequisite for identification of a new construction. However, Goldberg (2005, 17-18) notes that the researcher has to motivate the existence of each postulated construction by studying why a particular form is associated with particular functions in the language under analysis. This could be accomplished by establishing a link to e.g. constraints on acquisition of constructions or to discourse-related factors (see also Chapter 3.1.2).
Since constructions combine form and function, a constructionist account of language provides “a uniform representation of grammatical knowledge” (Croft 2005, 275, see also Langacker 1991, 343). Analyzing a particular pattern as a sign allows a researcher to produce an integrated account of its morphosyntactic, semantic, pragmatic and textlinguistic features (Hilpert 2010, 25). This characteristic of a construction as a unit of analysis can be certainly seen as yet another motivation for subscribing to the tenets of CxG (see also Chapters 3.2 and 3.3).

The following sub-chapter (3.1.1) discusses the tenets of constructionist approaches in more detail by looking at them from the perspective of descriptive and explanatory adequacy. After that the interrelation of constructions (3.1.2), their frequency effects (3.1.3) and a group of argument structure constructions (see 3.1.4) are accounted for.

### 3.1.1 Descriptive and explanatory adequacy of constructionist approaches

While some of the main arguments for assuming the existence of constructions have been summarized above, it is also essential to consider whether a family of constructionist approaches fulfills the main requirements of a scientific theory, namely descriptive and explanatory adequacy. According to Goldberg (2005, 17-18), CxG is descriptive as its aim is to analyze the entirety of language, both its core and periphery (see Chapter 3.1.2), and it is explanatory in the sense that it tries to find reasons for the existence of every language-specific form-function pairing.

Butler (2009) proposes a more detailed set of criteria of adequacy designed for functional approaches to language analysis, which specifies both types of adequacy. Here explanatory adequacy is further split into cognitive, discoursal, sociocultural and acquisitional adequacy. Butler (2009, 6-15) suggests that a functional theory should account for the way language is stored and processed, consider not only individual sentences but also whole texts, study how texts function within their sociocultural background as well as how language structures are learned. These criteria are considered to define the requirements for the type of data that should be described by a theory. He also notes that a functional linguistic theory should be based on the descriptions of authentic data covering different languages and different stages of language development. In an attempt to compare a number of functional and cognitive approaches, among them a number of constructionist accounts, a more comprehensive list of
features has been suggested (Gonzálvez-García and Butler 2006, 49-51). Certain groups of these features can be directly linked to the criteria of adequacy discussed here. Thus the first group dealing with the communication and the motivation of the linguistic system is comparable to the sociocultural adequacy. Even though certain social aspects of language use might not be the main focus of constructionist approaches, Gonzálvez-García and Butler (2006, 55-58) observe that almost all constructionist theories, which are being compared in the study, do account for the role of communication. Here it is important to take into account that this type of adequacy includes such criteria as considering the role of usage in the development of the language system and acknowledging commonalities between linguistic and other cognitive systems.

Usage-based linguistics emphasizes the dependence of language on its usage linking the sociocultural conditions of language use to the way the language is structured (Bybee 2010, 204). Blumenthal-Dramé (2012, 28-33) observes that by relying on corpus data the usage-based linguistics generalizes over a more or less heterogeneous speech community to describe the prototypical cognitive representation of language. While this methodology can be viewed positively as it considers social aspects of our language use, it has been also criticized for creating an ideal model that cannot be found in reality by investigating individual language use. Moreover, Langacker (1987, 128; 1991, 12) argues that particular constructions are selected by the speaker to trigger certain interpretations on the part of the hearer. This kind of manipulation which could to a large extent take place on the subconscious level is possible due to a number of alternative constructions that could be used to transport meaning. These alternatives are believed to be conventionally associated with different perspectives on the same situation or different pragmatic nuances. Therefore, especially if the speaker and the hearer come from the same sociocultural background, the communication exploiting these possibilities is expected to be more successful (Wray 2008, 20-21, 69; Wiechmann 2010, 42-45). Wray (2008, 18) further suggests that a speaker is often under some pressure to use the same patterns as the rest of the social group s/he belongs to. Moreover, by considering how the same constructions are used in a range of registers (e.g. Gilquin 2010, chapter 9) or at least specifying within which register the construction is studied, CxG certainly takes into account that language varies depending on the social situation in which it is embedded. Also a CxG perspective on interactional aspects of spoken
discourse (e.g. Günthner and Imo 2006) contributes to our understanding of sociolinguistic aspects related to the use of constructions. Gonzálvez-García and Butler (2006, 57) suggest that the Scene Encoding Hypothesis put forward by Goldberg (1995, 39; 1998, 205) further accounts for the sociocultural dimension of the CxG. According to this hypothesis, different argument structure constructions are used to transport basic types of human experience, among them such social situation as giving something to somebody.

Another sub-type of the explanatory adequacy, namely the discoursal adequacy, could be interpreted as containing features connected to the role of text and context within a theory (Gonzálvez-García and Butler 2006, 50). Whereas CxG deals mainly with individual sentences or smaller chunks, there have been some suggestions on dealing with linguistic material on the level of texts. For instance, Östman (2005) considers discourse patterns as constructions. Furthermore, descriptions of information structure within constructionist approaches deal with linguistic material beyond single sentences (e.g. Lambrecht 1994, 2004; Leino 2013). Goldberg (2005, 27) considers pragmatic factors to account for constructional motivation. Moreover, interactions between constructions and context are also examined in the edited volume by Bergs and Diewald (2009).

As discussed in more detail in Chapter 3.1.3, constructionist approaches postulate a link between usage and cognitive processing. Language structures, their acquisition and change are explained through general cognitive processes, such as categorization, analogy and chunking (Langacker 1987, Bybee 2010). Therefore, the criteria of cognitive adequacy are fulfilled by this family of theories as well.

Accounts of language in terms of constructions often take into account how both children and adults acquire the form-meaning pairings. Within this framework language is assumed to be learned through its use with the help of various cognitive skills, which are not restricted to language acquisition, encompassing proficiency in intention reading and pattern-finding (Tomasello 2003, 3-5; Goldberg 2006, 12). In the process of language acquisition a special role is assigned to frequency of input. It is often claimed that token and type frequencies influence learning of linguistic structures and generalizations to more abstract categories (Tomasello 2003, 107; Ellis and O’Donnell 2012: 270; cf. Brandt et al. 2011), as well as lead

---

5 With respect to constructions token frequency refers to the number of occurrences of a specific constructional form, whereas the measure of type frequency describes with how many items the constructional slots can be filled (e.g. Bybee 2010, 9; Ellis and O’Donnell 2012, 270; Blumenthal-Dramé 2012, 10-12).
to the formation of prototypes, which are reflected in children’s early utterances (Goldberg, Casenhiser and Sethuraman 2004). Children are believed to be capable of extracting statistical information from the input (Diessel 2007, 112), which can help them to recognize, for instance, word boundaries and word forms (Goldberg 2006, 70-71). It is also important to mention that according to CxG early language acquisition involves both individual words and complex grammatical structures. The process is modeled in the following way: “children learn simultaneously from adult utterances meaningful linguistic structures of many shapes and sizes and degrees of abstractness, and they then produce their own utterances on particular occasions of use” (Tomasello 2003, 100, see also Bybee 2010, 35). Thus both children and adults can learn certain complex constructions as one unit and at a later stage, if necessary, further separate these into their elements. If this re-analysis into smaller chunks is not required by the communicative situations, multi-unit sequences are assumed to remain stored at the same level of complexity (Wray 2008, 17-18). Also research on second language acquisition has considered how constructions are learned through input and what role frequency plays in this process (Ellis 2012).

Finally, descriptive adequacy includes the range of theoretical coverage in terms of analyzed linguistic phenomena and sources of data, amount of cross-linguistic studies as well as synchronic and diachronic comparisons (Butler 2009, 12-15). As mentioned above, CxG does not make a distinction between core and periphery in that it studies both general and idiosyncratic linguistic structures. The theory often uses corpus (e.g. Wulff 2006) or experimental evidence (e.g. Wonnacott et al. 2012), as well as a combination of both (e.g. Gries, Hampe and Schöneweierf 2005, 2010). It also takes into account examples cited by others, dictionaries, literary words and introspection (González-García and Butler 2006, 61-62). Moreover, there have been some studies on cross-linguistic variation of constructions, even though these types of research questions have received less attention within constructionist approaches. Since the specific constellations of formal and functional features are considered to be language-specific (e.g. Goldberg 2003, 222), the cross-linguistic analyses of constructions have to address the issue of comparability of constructions across different languages. Here different starting points for comparisons have been suggested, namely formal or functional aspects as well as a combination of both (e.g. Boas 2010b, see Chapter 3.3.1 for a detailed overview).
The interconnection between usage, synchronic and diachronic variation within the constructionist approaches is considered by the studies on grammaticalization and lexicalization (Traugott 2008; Hilpert 2008; Trousdale 2008; Bybee 2003, 2006, 2010). For instance, Bybee (2010, 28-29) suggests that through repeated usage a more specific instance of a general construction becomes more autonomous and develops into a new independent construction. In a study of future constructions in a number of Germanic languages a gradual semantic change of the verbs, e.g. *shall* and *will* for English, and an accompanying process of grammaticalization is investigated by looking at the co-occurring verbs. The diachronic distinctive collexeme analysis (Hilpert 2008, 41-45), which is a variant of collostructional analysis, is employed using the corpus data from different periods of language development (Hilpert 2008). While the details of collostructional analysis (for synchronic studies) are introduced in Chapter 4.3, here it is sufficient to say that the method is based on frequencies of both lexical items and more complex constructions into which these are integrated. Thus, in usage-based constructionist accounts language change is also linked to frequency.

The present sub-chapter has applied the criteria of adequacy to the family of constructionist approaches to language. It has been shown that CxG addresses, at least to some extent, all issues that are associated with the categories of descriptive and explanatory adequacy. This review indicates that the aim of the theory, namely to provide a unified description of language covering different phenomena on different levels of linguistic analysis, is being achieved in the wide range of studies covering various areas. Taking this into account CxG could be said to represent a solid theoretical foundation for the present study. At the same time some of the areas, such as discoursal and cross-linguistic adequacy, though considered, have received less attention within the theory than other aspects, e.g. those related to language acquisition. Since the application of CxG in translation studies involves cross-linguistic comparisons, this dimension will be addressed and further developed in the present study. For this it is necessary to consider previous research connected to this type of adequacy, which is the topic of Chapter 3.3. But first some theoretical assumptions of CxG, which are especially relevant for the research questions of this study, are scrutinized in the following sub-chapters.
3.1.2 Language as a network of constructions

Form-meaning pairings are postulated on all levels of linguistic description (Goldberg 2003, 220): individual morphemes, words, phrases, clauses, but also discourse patterns (Östman 2005) can be analyzed as constructions within this theoretical framework.

<table>
<thead>
<tr>
<th>Construction</th>
<th>Form/Example</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morpheme</td>
<td><strong>E:</strong> anti-</td>
<td><strong>E:</strong> e.g. the opposite of something</td>
</tr>
<tr>
<td></td>
<td><strong>G:</strong> -keit</td>
<td><strong>G:</strong> a derivational suffix, changes an adjective into a noun</td>
</tr>
<tr>
<td>Word</td>
<td><strong>E:</strong> Avocado</td>
<td><strong>E:</strong> a particular fruit</td>
</tr>
<tr>
<td></td>
<td><strong>G:</strong> Tisch</td>
<td><strong>G:</strong> a particular piece of furniture</td>
</tr>
<tr>
<td>Complex word</td>
<td><strong>E:</strong> tablespoon</td>
<td><strong>E:</strong> a large spoon</td>
</tr>
<tr>
<td></td>
<td><strong>G:</strong> Pizzen</td>
<td><strong>G:</strong> type of food, plural</td>
</tr>
<tr>
<td>Idiom (filled)</td>
<td><strong>E:</strong> Going great guns</td>
<td><strong>E:</strong> doing something extremely well</td>
</tr>
<tr>
<td></td>
<td><strong>G:</strong> die Kurve kratzen</td>
<td><strong>G:</strong> leave quickly and inconspicuously</td>
</tr>
<tr>
<td>Idiom (partially filled)</td>
<td><strong>E:</strong> Jog (someone’s) memory</td>
<td><strong>E:</strong> to make someone remember something</td>
</tr>
<tr>
<td></td>
<td><strong>G:</strong> In (jemandes) Fußstapfen treten</td>
<td><strong>G:</strong> to follow somebody (metaphorically)</td>
</tr>
<tr>
<td>Ditransitive (double-object)</td>
<td><strong>E:</strong> Subj V Obj₁ Obj₂ (e.g. He gave her a Coke)</td>
<td><strong>E:</strong> Transfer (actual)</td>
</tr>
<tr>
<td></td>
<td><strong>G:</strong> Subj V Obj₁ Obj₂ (e.g. Emmi kochte Lili eine Suppe)</td>
<td><strong>G:</strong> Transfer (metaphorical extension)</td>
</tr>
</tbody>
</table>

Table 1: Examples of constructions (adapted from Goldberg 2003,220, Boas 2014, 54).

As is shown in Table 1, constructions can vary not only in terms of size, but also in the level of abstractness. Thus a construction can have all or some slots that are lexically fixed (filled

---

6 The table contains partly different examples and is supplemented by definitions based on Longman online and DWDS.
or partially lexically filled constructions), or it can be fully schematic (Bybee 2010, 9). An example of the latter is the ditransitive construction, which combines the types of the grammatical functions and the semantic roles (for more detailed analysis see Goldberg 1995, 141-151; Stefanowitsch and Gries 2003, 227-230, see also Table 1). In contrast to this, the way construction being of the same size and having been classified as an argument structure construction as well, includes the lexically fixed element way. Its form has been described as [Subj V POSS way OBL], involving not only the fixed noun, but also some schematic slots (for more detailed analysis see Goldberg 1995, 199-218; Jackendoff 1990, 211-223; Israel 1996; Stefanowitsch and Gries 2005, 18-22). Filled idioms are characterized by no variation: only the occurrence of specific lexical items in specific grammatical form corresponds to the meaning of a particular fully lexically filled idiom (consider the example going great guns in Table 1). This description is particularly true for English, whereas the German idioms are less fixed. Due to flexible word order, the elements of the idiom do not always form a continuous unit in German. Moreover, the German verb is modified in agreement with the person and number of the subject. Both of these characteristics could be illustrated using the following sentence taken from Boas (2014, 54): Um 23 Uhr kratzten wir die Kurve ‘At 11 p.m. we are leaving’. The three-way distinction in terms of degrees of abstractness (i.e. between schematic, partially lexically fixed and lexically fixed constructions) might be considered simplified. For instance, constructions with schematic slots differ with respect to the allowed variation: whereas some slots can be filled with a wide range of lexical material, others are very restricted. The latter constructions are rather specific even though their slots are not completely fixed (Bybee 2010, 80). Therefore, it is better to view abstractness of constructions in terms of a continuum (Blumenthal-Dramé 2012, 204).

A strand of research, which is compatible with the study of constructions, concentrates on item-specific multi-unit sequences, which encompass idioms but also other more regular types of collocations (Wray and Perkins 2000, Wray 2002, 2008, Corrigan et al. 2009a, 2009b). These specific units are often referred to as formulaic sequences or lexical bundles (see Wray 2002, 9 on a complete list of alternative terms7). A range of studies dealing with such formulae are corpus-driven: the co-occurring lexical items are extracted from the

---

7 Note, however, that some minor differences in the use of these terms are assumed to exist (Wray 2002, 8).
corpus based on certain criteria, such as frequency or size, and in the second step further
categorized to formulate some generalizations (e.g. Biber, Conrad and Cortes 2004; Simpson-
Vlach and Ellis 2010; Liu 2012). These formulaic sequences could be referred to as lower-
level constructions.

Whereas this point is not shared by all strands of CxG (cf. Croft and Cruse 2004, 285), it is
often claimed that the memory of a language user is redundant. The same linguistic
structures are assumed to be stored both in isolation and integrated into larger chunks (Wray
2008, 13-15), specific utterances and abstract generalizations over them could co-exist
(Langacker 1987, 42; Tomasello 2003, 106; Dąbrowska 2004, 203). Assuming that memory
storing capacities could cope with the multiple representations, this type of organization of
linguistic knowledge results in faster processing, as whole chunks can be extracted during
speech production and perception (Langacker 1987, 57-60; Tomasello 2003, 106-107;
Blumenthal-Dramé 2012, 11).

All stored constructions are believed to be connected with each other forming a network
referred to as a constructicon (e.g. Goldberg 1995, 5; Hoffmann and Trousdale 2013, 3). This
network is typically described as having a hierarchical structure with abstract
generalizations at the top and specific constructions at the bottom. It thus represents the
syntax-lexicon continuum (Hoffmann and Trousdale 2013, 1) between substantive and
schematic structures (Croft and Cruse 2004, 255-256; Langacker 2009, 2): lexically specific
and grammatically abstract structures are interconnected within the network of
constructions. This view provides additional justification to the study of idiosyncrasies:
irregular structures are analyzed not only to account for the whole range of linguistic
material as is required by descriptive adequacy of a theory (see Chapter 3.1.1), but also
because their properties may be generalized to more regular and productive patterns
(Goldberg 2006, 14). For instance, the idiom \[\textit{kick the habit}\] is said to be connected to the
more abstract construction \[\textit{kick Obj}\], which is in turn an instantiation of the pattern \[\textit{Verb
Obj}\] representing the general category \[\textit{VP}\] (Croft and Cruse 2004, 263).

When postulating a construction on one of the levels of complexity or abstraction, it is
important to consider the two criteria contained in the definition of constructions, namely
non-compositionality and frequency. For instance, the idiom \[\textit{kick the habit}\] could be
classified as a lexically-fixed construction rather than simply an instance of the abstract
construction because it is non-compositional. While idiomatic units, due to their lexical fixedness and non-compositionality, are rather straightforward candidates for constructions, the identification of other patterns as constructions requires more argumentation. Goldberg (2005, 38) stresses that the existence of a construction should be motivated to avoid postulation of new constructions as “ad hoc means of accounting for exceptional cases”. Such motivation could consider, for instance, specific pragmatic functions of a pattern. In the present study I classify a complex pattern as a construction in its own right if a structure contains slots characterized by a restricted set of lexical items and has independent function(s) distinct from combinations of different lexical material presented in the same configuration of grammatical functions. Moreover, association of a pattern with specific registers is also taken into account as a functional feature supporting an existence of an independent form-function pairing. However, this type of motivation required for postulation of constructions does not undermine the argument that CxG accounts for the entirety of language. If a certain combination of linguistic elements does not function as a unit, its components can still be analyzed on a lower level of complexity, for instance as individual words, or the pattern can be studied on a higher level of abstraction, for instance as a general transitive construction.

The hierarchical network of constructions could be analyzed both in a top-down and a bottom-up way. From the first perspective the lower-level constructions are said to inherit features from more general form-meaning pairings (Goldberg 1995, 67, 2006, 13-14). Goldberg (1995, 73) suggests that formal and functional aspects of a node in the hierarchy could be motivated either by one or several higher-level structures. An assumption that item-specific constructions arise from more general patterns due to their repeated production (Bybee 2010, 28, Blumenthal-Dramé 2012, 201) can also be classified as taking a top-down perspective on the constructicon. Moreover, in terms of language acquisition it is claimed that certain multi-word units are learned as chunks and may or may not be analyzed into their individual components at a later stage (Wray 2008, 18). The syntax-lexicon continuum or constructicon has been defined above as linked to the relation between abstract and specific items. However, the notion of the syntax-lexicon continuum is sometimes also used to refer to structures of different degrees of complexity, i.e. a continuum between atomic and complex constructions (Croft and Cruse 2004, 255-256). It seems plausible to assume that more
complex structures are located higher in the hierarchy than more atomic ones. Taking this into account, acquisition of certain multi-unit sequences and their possible re-analysis into smaller parts can be considered among the processes, which involve moving top-down along the network's dimension of complexity.

According to the bottom-up view, the schematic slots of the more general constructions develop through the cognitive process of categorization: a language user is assumed to generalize over a set of item-specific instances to arrive at higher-level structures (Bybee 2010, 78). This position is supported by accounts of language acquisition: for instance, Tomasello (2003, 120-126) suggests that children first learn item-specific constructions before the process of abstraction takes place. Moreover, the bottom-up perspective exists also on the dimension of complexity: Langacker (2009, 6) studies how several more atomic units are connected to form a more complex structure.

The preceding discussion has shown that the two readings of the hierarchy do not have to be exclusive. Just as it is claimed in CxG, the same situation or entity can be looked at and described from different angles (cf. Langacker 1987, 128). The conceptualization of the constructional hierarchy depends on the research question. The present study considers a top-down perspective: the more abstract patterns are classified to identify lower-level, i.e. lexically more specific constructions.

Whereas the existence of the continuum between specific and abstract constructions is agreed upon in CxG, it is not clear how abstract the postulated constructions can be to be considered psychologically plausible. Thus higher-level abstract structures can be interpreted as only idealizations that are not among the stored constructions of an individual language user (Blumenthal-Dramé 2012, 29). Tomasello (2003, 98) argues that “[t]he level of abstraction at which the speaker is working in particular cases may or may not correspond to the most abstract level the linguist can find; it is in all cases an empirical question”. We will return to this question in Chapter 3.1.4, after considering how CxG links usage data to the level of cognition.

### 3.1.3 Corpus to cognition principle

Constructionist approaches to grammar, which are considered in the present study, try to account for linguistic phenomena through general cognitive processes, such as
categorization, analogy, chunking, schematization and association (Langacker 2008, 16, Bybee 2010, 7): to achieve psychological plausibility this family of theories pays special attention to the cognitive processing of language (Langacker 2008, 16). One of these cognitive processes is chunking, i.e. turning structures of various internal complexities into units. This process leads to automatization or entrenchment of this pattern in the human mind. In other words, entrenchment is at work when “through repetition or rehearsal, a complex structure is thoroughly mastered to the point that using it is virtually automatic and requires little conscious monitoring” (Langacker 2008, 16). This process could be illustrated through a process of learning how to drive a car: while the beginner has to think what should be her/his next action, an experienced driver performs the actions automatically. In the latter case, the driving situation is not split into individual steps, but rather becomes entrenched as one complex whole, in which the transition between the components is smooth and does not require extra processing effort.

When talking about entrenched constructions the driving example is transferred to language. Thus an entrenched construction is produced and comprehended automatically due to its frequency and is stored as a unit (Langacker 1987, 57-60; Ellis 2002, 2012; Diessel 2007) with schematic and/or lexically fixed slots. The use of such entrenched constructions is claimed to facilitate both language production and perception (Bybee 2010, 34): as mentioned in 3.1.2 it is assumed that patterns of various size and abstractness can be holistically stored in language user’s mind to allow their fast extraction during linguistic processing. On the production side this process of entrenchment is often accompanied by phonological reduction making it faster to pronounce the unit. As a hearer, a language speaker can anticipate what is going to come next by using her/his knowledge of the entrenched constructions (Bybee 2010, 34).

As stated above, entrenchment is a cognitive process. However, since it is difficult to study cognitive activities directly, it is necessary to establish a link between behavioral data and cognition. In usage-based linguistics this link is based on the corpus-to-cognition principle, i.e. the assumption that different frequency distributions of linguistic patterns reflect different levels of cognitive entrenchment (Schmid 2000, 39; 2010, 101-102, cf. e.g. Langacker 1987, 59; Hilpert 2010, 22, 37). Thus one of the possible conditions for a pattern to be identified as a construction is its “sufficient frequency” (Goldberg 2006, 5). The frequent
use of a structure ensures that the individual elements form an entrenched unit, or a construction, even if the slots of the pattern are schematic. However, the question arises whether there are any specific frequency thresholds required for a structure to be considered entrenched (Blumenthal-Dramé 2012, 38), or what other methods could be used to distinguish between entrenched and non-entrenched patterns. One of these methods is the collostructional analysis that aims at establishing entrenchment of lexical items in the individual slots of a construction (for details see Chapter 4.3). Moreover, Wiechmann (2010) suggests analyzing complex patterns, such as relative clauses, into a set of features and applying statistical methods of association rule mining and configural frequency analysis to identify entrenched combinations of variables. Höche (2009, 135) observes that “so far no measure for a pattern’s degree of entrenchment is available”: existing methods could be used to formulate hypotheses related to entrenchment in the speakers’ minds.

Recent studies in the area of cognitive linguistics advocate for converging evidence and warn against exclusive reliance on frequencies (e.g. Gries, Hampe and Schöenefeld 2005, 2010; Goldberg, Casenhiser and Sethuraman 2004; Blumenthal-Dramé 2012, see also 5.1). Comparisons of the corpus-based results concerning frequencies with evidence from experiments do not always support a straightforward interpretation of the frequency data in terms of the level of entrenchment (Gilquin 2010, 164-165). However, the diverging evidence could be explained taking into account issues of ecological validity of corpus-based and experimental results as well as different proportions of idioms in the two sources of data. For instance, elicited sentences, in contrast to naturally produced corpus-based examples, are typically not integrated into a larger linguistic context and are, therefore, characterized through different types of referents and less idiomatic expressions (Gilquin 2010, 165).

It has been suggested that frequency is not the only factor that can have an effect on cognitive entrenchment. In order to get a more comprehensive picture of this cognitive process investigation of recency of activation (Langacker 1987, 59; Bybee 2010, 28; Ellis 2012, 14-15), age of acquisition, the length of a string, rarity of the stimulus as well as high mental imagery (Blumenthal-Dramé 2012, 39-40) have been suggested. A recent study on entrenchment identifies a further shortcoming of the studies concentrating on acquisition or processing of multi-word constructions. Blumenthal-Dramé (2012, 47-48) indicates that whereas the aim of these investigations is to show that high-frequency items are stored
holistically in the minds of language users, the results allow only conclusions on the easier processing of these patterns, not excluding the possibility that high frequency leads to faster composition of these sequences from individual elements. The author tests entrenchment-related assumptions by extracting highly frequent words from a corpus and conducting a series of psycholinguistic and neuroimaging experiments that use these stimuli. Based on a review of existing studies on entrenchment two types of features are identified, that are taken to define the analyzed phenomenon. Depending on the degree of entrenchment, the processing of entrenched structures is said to involve their more or less automatic production and comprehension requiring a different amount of conscious monitoring. On the level of mental representation, an entrenched expression is believed to be stored as a unit (Blumenthal-Dramé 2012, 67): “existing entrenchment definitions can be taken to suggest that there must be a point at which gradual differences in processing ease turn into qualitative differences in representation” (Blumenthal-Dramé 2012, 68).

Her experiments using multi-morpheme words are thus designed to test both the gradual phenomenon of the ease of processing and the assumed categorical effect of chunking. The masked priming paradigm involving interaction between derivatives and their bases\(^8\) revealed that the log transformed relative frequency, representing the frequency of the derivative divided by the frequency of its base, has a significant effect on entrenchment (Blumenthal-Dramé 2012, 147, 180-181). This finding is interpreted as a possible indicator that token frequencies on different levels of linguistic analysis, in this case on the levels of complex words and more basic lexical morphemes, all contribute to the process of entrenchment. Moreover, it can be seen as evidence that both simple and complex units are stored in the minds of language users (Blumenthal-Dramé 2012, 150-152). The regression slopes for the variable of log relative frequency have not been characterized by any breakpoints, which would indicate frequency thresholds beyond which complex structures are stored holistically. Therefore, Blumenthal-Dramé (2012, 153) argues that entrenchment is clearly a gradient phenomenon. The study has also shown that all aspects of entrenchment are better conceived as a continuum, also the holistic storage of units, which is initially

\(^8\) The reported study distinguishes between part-to-whole and whole-to-part conditions, depending on whether the base or the derivative functions as a prime. The prime is shown only for 60ms, i.e. is “masked”: the participants are assumed to process the shown word only "preconsciously" so that the results are not influenced by conscious strategies (Blumenthal-Dramé 2012).
characterized as a categorical one: “each use of the inventory [of linguistic structures] slightly modifies it in terms of association weights and strengths of representation” (Blumenthal-Dramé 2012, 199). It is further assumed that for every complex structure both processing ways could compete, i.e. the (de)composition from smaller elements and the holistic retrieval from memory (Blumenthal-Dramé 2012, 199). The conclusion that storage of a pattern as a unit is a gradient phenomenon confirms an assumption postulated within CxG (cf. Langacker 1987, 59-60).

High frequency reflecting a high level of entrenchment is often considered as an explanatory factor for selecting individual constructions by the language users. However, Wiechmann (2010, 9-10) notes that additional explanation is required to account for the motivation of using certain constructions more frequently than others in the first place. A speaker is believed to choose those form-meaning pairings that present the expressed meaning from a certain perspective suitable for the speaker’s communicative goals. If a certain construction turns out to be successful for communicating specific meaning, it is likely to be used more often for these purposes (Wiechmann 2010, 42-43). Therefore, the explanatory chain could be summarized as follows. In the first step construction-specific features, which result from a unique combination of form and function, increase its chances to be selected by the speaker of a given language. The more often it is selected, the more entrenched the construction becomes in this language. The more entrenched constructions are easier to process and are thus even more likely to be selected (Wiechmann 2010, 61; Wray 2002, 25). It is important to consider the social aspects that play a role in a spread of constructions within a language community: the more often a language user encounters a particular construction, the more likely s/he will start producing this construction influencing in turn the language of others. Certain constructions can be limited to or be especially typical of individual registers (Kerz and Haas 2009, 103), i.e. certain linguistic situations could be more compatible with the specific characteristics of constructions. Thus the initial motivation that facilitates the use of a construction may be stronger in one register than in another. It is claimed here that the same is true with regard to different languages: the construction-specific features may be favored in one language but not in another due to linguistic and cultural differences. As a result, comparable constructions in two languages could be characterized by different levels of entrenchment.
3.1.4 Argument structure constructions

In the present study a group of complex clause-level structures, referred to as argument structure constructions, is examined in more detail. Before the empirical investigation of these constructions in originals and translations is presented in Chapters 5-6, this sub-chapter discusses the previous research on argument structure patterns, with a special focus on analyses within the constructionist approaches to language.9

From a constructionist perspective argument structure constructions are defined as “a special subclass of constructions that provides the basic means of clausal expression in a language” (Goldberg 1995, 3). Some of the examples discussed in the literature are ditransitive (Goldberg 1995 for English; Ekberg 2013 for German), resultative (Goldberg 1995 for English; Boas 2003 for English and German), caused motion (Goldberg 1995 for English; Boas 2003 for English and German), denominative (Hampe 2011 for English), and intransitive motion (Rohde 2001 for English; Stefanowitsch 2008 for German) constructions. Table 2 provides an overview of these and a number of other schematic argument structure constructions. Since some of them have received less attention than others, some of the information could not be found in the literature: the function of the intransitive construction has been selected from possible event types described in Goldberg (1995, 39), and the example of this construction has been extracted from the CroCo corpus (see Chapter 4.1 for its description). However, it should be further investigated whether this function is the prototypical one for the intransitive construction. As is shown, Goldberg (1998, 206) distinguishes at least two different constructions with the general form [Subj V Dobj]. It can be assumed that there are other sub-types of both intransitive and transitive constructions, which are not listed here. Table 2 contains examples in German only for those constructions, which have been analyzed for this language in the previous research.

---

9 Another treatment of argument structure patterns is presented by the so-called projectionist approaches, which identify selected verbs rather than constructions as the elements that determine the number and types of arguments in a clause (Goldberg 1995, 11; Bencini and Goldberg 2000, 640, Levin and Rappaport Hovav 2005, 189-193). Among this family of approaches are such theories as head-driven phrase structure grammar (HPSG) and valency theory (Jacobs 2008, 10; Welke 2009a, 515). The constructionist and projectionist approaches seem to be compatible: they deal with the same phenomena either from the perspective of abstract argument structure constructions or individual verbs and verb classes (Jacobs 2009, 495-496; Welke 2009a, 514-515; 2009b, 83-84). Thus, a number of German linguists have suggested that linguistic description could benefit from a combination of valency (or another item-based) theory and CxG (e.g. Nikula 2007; Welke 2009a, 2009b; Herbst 2011; Stefanowitsch 2011). Moreover, linguistic accounts within the HPSG (Müller 2008) as well as an approach by Boas (2002, 2003, 2010, 2011, 2013, 2014), that stresses the importance of individual verbs and their frame-semantic knowledge, are considered within constructionist approaches.
<table>
<thead>
<tr>
<th>Argument structure construction</th>
<th>Form</th>
<th>Function</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intransitive</td>
<td>Subj V</td>
<td>X is in a state</td>
<td>E: They adapted. (CroCo EO_ESSAY_028 s17)</td>
</tr>
<tr>
<td>Intransitive Motion</td>
<td>Subj V Oblique</td>
<td>X moves to Y / along Y</td>
<td>E: The fly buzzed into the room. (Goldberg 1995, 3) G: Die Flaschenpost trieb ans Ufer. (Stefanowitsch 2008, 249)</td>
</tr>
<tr>
<td>Transitive</td>
<td>Subj V Dobj</td>
<td>X acts on Y</td>
<td>E: The farmer killed the duckling. (Taylor 1998: 177)</td>
</tr>
<tr>
<td>Possessive</td>
<td>Subj V Dobj</td>
<td>X acquires/possesses Y</td>
<td>E: Sam secured a good job. (Goldberg 1998, 206)</td>
</tr>
<tr>
<td>Ditransitive</td>
<td>Subj V Iobj Dobj</td>
<td>X causes Y to receive Z</td>
<td>E: Pat faxed Bill the letter. (Goldberg 1995, 3) G: Sie schickt dem Kind eine Karte. (Ekberg 2013)</td>
</tr>
<tr>
<td>Resultative</td>
<td>Subj V Dobj ComplAdjP</td>
<td>X causes Y to become Z</td>
<td>E: She kissed him unconscious. (Goldberg 1995, 3) G: Er trat den Ball kaputt. (Boas 2003, 302)</td>
</tr>
<tr>
<td>Caused Motion</td>
<td>Subj V Dobj Advdir</td>
<td>X causes Y to move Z</td>
<td>E: Pat sneezed the napkin off the table. (Goldberg 1995, 3) G: Er trat den Ball in das Tor. (Boas 2003, 302)</td>
</tr>
<tr>
<td>Denominative</td>
<td>Subj V Dobj ComplNP</td>
<td>X says Y is Z</td>
<td>E: They elected me senior treasurer. (Hampe 2011, 222)</td>
</tr>
<tr>
<td>Predicate adjective</td>
<td>Subj V ComplAdjP</td>
<td>X has a characteristic Y</td>
<td>E: The house is large. (Taylor 1998, 177)</td>
</tr>
</tbody>
</table>

Table 2: Schematic argument structure constructions

As mentioned earlier, according to the Scene Encoding Hypothesis at least some basic types of the argument structure constructions “encode as their central senses event types that are basic to human experience” (Goldberg 1995, 39). The underlying principle of this hypothesis can be related to the notion of embodiment in cognitive linguistics, according to which our
bodily and social concrete experience shapes both our conceptualization and language (Lakoff 1987, 12, 206-207). Since these experiences are relevant to human beings independently of their language, these functions can be expected to be realized in one way or another cross-linguistically (cf. Goldberg 1995, 39). Based on this discussion, the definition of argument structure constructions can be modified reflecting their main function: a subclass of constructions, which refers to events and its participants (cf. Wellens 2011, 184) as well as states. Since argument structure constructions have been referred to earlier as “the basic means of clausal expression” (Goldberg 1995, 3), it is not surprising that also clauses describe events and states as well as such elements as participants, circumstances or attitudes (Biber et al. 1999, 120). At the same time, Goldberg (2006, 43) observes that since clauses always involve a combination of constructions, also other constructions realized by verbs and individual arguments should be considered in interpreting the meaning of a clause. The theory distinguishes between argument and participant roles. The former are more general semantic roles, which represent the slots licensed by argument structure constructions, such as Agent, Theme, Patient, whereas the latter are verb-specific roles, which encode frame-semantic knowledge evoked by a particular verb integrated into constructions (Goldberg 1995, 43). CxG employs the theory of Frame Semantics, according to which words activate networks of concepts, i.e. frames10: Fillmore (2006, 373) claims that the whole frame is required to interpret the meaning of one of its elements. Frames can also be referred to as the context of experience (Fillmore 1976, 23), i.e. as the extended context on the conceptual level into which the semantics of a lexical item is integrated. While conceptual categories represent the paradigmatic dimension of knowledge, frames organize individual concepts on the syntagmatic level. Coming back to constructions, the merging of participant and argument roles can be illustrated using the example of the ditransitive construction. As was mentioned earlier, the form of the construction can be expressed as a combination of Verb, Subject, Indirect Object and Direct Object. The central meaning of the construction, namely "cause to receive", licenses three argument roles, namely Agent,

---

10 Fillmore (1976) distinguishes between interactional and cognitive or conceptual frames. The first type refers to social situations with particular linguistic material associated with it, e.g. a greeting frame, which specifies not only the forms of greetings but also other social conventions surrounding this interaction. The cognitive frames are the ones we are concerned with when considering how lexical items access the conceptual networks. An example used to illustrate this type of frame is a commercial frame, activated e.g. through the verbs sell or buy (Fillmore 1976, 25). It is the conceptual frames that are referred to in CxG.
Recipient and Patient. If the verb *hand* is used in this construction, the specific participant roles linked to the argument roles are Hander, Handee and Handed (Goldberg 1995, 51). The process of merging can alter the characteristics of the verb’s participant roles: for instance, as mentioned above, if the construction licenses more argument roles than the verb, the additional argument is imposed onto the verb by the construction (Goldberg 1995, 52-56; 2006, 42; Michaelis 2004, 25, see Chapter 3.1 for an example of the verb *sneeze* and the caused-motion construction).

Just like other constructions, the group of constructions under analysis is considered to be organized in a structured network. There have been no attempts to represent all postulated argument structure constructions in a unified taxonomy; only individual segments concentrating on certain sets of structures have been linked together (Goldberg 1995, 109; Höche 2009, 141-142; Wellens 2011, 188). According to a part of this hierarchy presented by Goldberg, both transitive and intransitive constructions are instances of a more abstract subject-predicate construction, which, for English, specifies the SVO word order (Goldberg 1995, 109-110). In particular, when discussing German argument structure constructions, the level on which word order is specified in the hierarchy becomes highly relevant: Müller (2008, 190) observes that in German, depending on the type of the clause, the finite verb could occur in different positions and that the order of obligatory and optional arguments could vary as well.

Goldberg (2006) notes that word order is not indicated at the level of the lower-level constructions, such as the ditransitive one: according to this view, the construction exists whenever the specific grammatical functions are mapped onto the specific semantic roles, as discussed above. The argument in favor of this position is that the same construction characterized by this set of grammatical functions and the meaning of the transfer of possession is also present in questions. For instance, in example (11) the question construction is assumed to be combined, among other things, with the ditransitive construction, where *Liza*, the subject, is mapped onto the Agent, *Zach*, the indirect object, functions as the Patient, and the question word is the direct object with the role of the Theme (Goldberg 2006, 10, 20). Just like in this concrete utterance the word order is specified by the question construction, in the declarative version, illustrated through (12), this role is assigned to the subject-predicate construction (Fischer and Stefanowitsch 2008, 7), and in
it is the topicalization construction that determines the positions of the arguments (Goldberg 2006, 21). Thus, though the word order is relatively fixed in English, utterances may include constructions that change the order of elements (Kuningas and Leino 2006, 301). These constructions, which could be referred to as ordering constructions, pair specific word order with specific function, for instance, connected to information structure (Kuningas and Leino 2006, 303).

(11) What did Liza buy Zach? (Goldberg 2006, 10)
(12) Liza bought Zach a book.
(13) A dozen roses, Nina sent her mother! (Goldberg 2006, 21)

Another approach to word order variation, introduced by Cappelle (2006), considers structures characterized by different word order as allostructions, i.e. “variant structural realizations of a construction that is left partially underspecified” (Cappelle 2006). Thus, taking the phenomenon of the particle placement alternation as an example, it is suggested that the two possible realizations shown in (14b) and (14c) are two allostructions of the construction in (14a), in which the position of the particle is not specified. The connection between the so-called “super-category” with non-fixed word order and its realizations should not be mistaken for one of inheritance (cf. Goldberg 1995): Cappelle (2006) claims that such underspecified constructions can exist on various levels of abstractness, and, therefore, the inheritance links hold between the form-meaning pairings on different levels rather than between the unspecific constructions and their allostructions. Coming back to the example, the inheritance link is argued to exist between the abstract verb-particle construction and the lexically specific [drum up] construction. Moreover, Cappelle stresses the existing connections between the individual allostructions on the different levels of abstraction, i.e. between [V Particle NP_{dobj}] in (14b) and the more specific allostruction [drum up NP_{dobj}] in (15b) as well as between [V NP_{dobj} Particle] in (14c) and [drum NP_{dobj} up] in (15c). For instance, he argues that though the construction [drum up] strongly prefers the word order presented in (15b), the word order in (15c) should be used when the NP is mapped onto an unstressed pronoun. This applies not only for the partially lexically fixed allostruction [drum NP_{dobj} up] but also for other instances of the more general allostruction [V NP_{dobj} Particle].

(14) a. V {Particle} NP_{dobj} {Particle}
    b. V Particle NP_{dobj}
    c. V NP_{dobj} Particle (adapted from Cappelle 2006)
(15)  a. drum {up} NP_{dobj} {up}
b. drum up NP_{dobj}
c. drum NP_{dobj} up (adapted from Cappelle 2006)

It is important to stress that this approach also argues for treating different word order patterns as instances of the same construction, rather than introducing new constructions based solely on the difference between the positions of the elements. However, the approaches by Goldberg and Cappelle differ with respect to alternation patterns: thus while Goldberg (2006) argues for the same ditransitive construction being present in utterances involving questions, subject-predicate and topicalization constructions, she does distinguish between the ditransitive and the prepositional paraphrase constructions, just like e.g. Gries (2003) differentiates between the two particle placement constructions. One of the motivations behind this view is the claim that none of the constructions in the corresponding pairs can be considered as having “a privileged status” (Goldberg 2006, 44). To stress that both patterns are equally entitled to be analyzed and to find certain generalizations on the occurrences of each of the constructions, she suggests that they should be studied independently. Research in this area has shown that such alternations differ in their meaning: example (16) is assumed to map the semantic role of Recipient onto 
Mel, whereas in (17) the same element is connected to the role of Goal. As a result, the use of the ditransitive construction presupposes that Mel will receive the book, and the prepositional paraphrase could also imply that Mel was supposed to buy the book himself but was not able to, the final recipient may be somebody other than Mel (Goldberg 2006, 26-29).

(16)  Mina bought Mel_{recipient} a book. (ditransitive) (adapted from Goldberg 2006, 27)
(17)  Mina bought a book for Mel_{goal}. (prepositional paraphrase) (adapted from Goldberg 2006, 27)

It is also claimed that the ditransitive and the prepositional paraphrase differ with respect to their information structure: in contrast to the Goal in the prepositional paraphrase, the Recipient in the ditransitive construction is more likely to be given than new in relation to the preceding discourse (Collins 1995, Goldberg 2006, 138-143). Previous studies observe that the indirect object itself is less likely to move in the clause: for instance, questions in which the indirect object is substituted by a question word or sentences with the indirect object in the topicalized position are considered less acceptable than the declarative versions
where the indirect object occurs before the direct object: therefore, it is concluded that “the primary determinant of indirect object status in English is positional” (Collins 1995, 36). It could be argued that this claim is in conflict with the view described above according to which example (11) contains the same ditransitive construction as the sentence in (12): while the indirect object may be more fixed in its position in the clause, the same is not assumed for the direct object. If the direct object occurs at the beginning of the clause, the relative position of direct and indirect objects is changed. Since it is the word order that determines the information structure, the indirect object in questions like (11) is likely to be not restrictive to the given information.

While arguing for treating these and similar variants as distinct constructions, Goldberg (2006: 43-44) admits that the existence of alternations does play a role in language processing: in other words, the alternative versions can be assumed to be in some way connected in the speakers’ minds. This claim is in line with the observation by Cappelle (2006), according to which some lower-level schemas, as illustrated through example (18), allow variation of word order. Since both patterns have the same idiomatic meaning, they are likely to be linked on the level of cognitive representation.

(18) buck up one’s ideas / buck one’s ideas up (Cappelle 2006)

Moreover, Cappelle suggests that this interrelationship also plays a role in language acquisition: idioms with one possible word order, e.g. drum up support, could be acquired through the process of indirect negative entrenchment during which the acquisition of the correct word order pattern is facilitated through the lack of exposure to another verb-particle ordering. According to Cappelle, this fact indicates that the two orderings are characterized by semantic similarity that has an effect on language learning (Cappelle 2006). These arguments could be used to account not only for syntactic alternations, but also for word order differences of the argument structure constructions in general: taking into account that the ditransitive construction is associated with the same general meaning irrespective of the word order pattern it is integrated in, one could expect that different orderings of arguments discussed above are related on the levels of cognitive representation, processing and acquisition (Serbina 2013, 173).

The account of the ditransitive form-meaning pairing by Leino (2010) indicates that in English it is the word order that determines the assignment of grammatical functions and,
therefore, should be considered as part of the argument structure constructions. Thus, in (19) *the dog* occurs in the initial position and receives the grammatical function of subject in combination with the semantic role of Agent, whereas in (20) the same NP occurs between the finite verb and the direct object and, as a consequence, is no longer the subject, but rather the indirect object with the role of Recipient (Leino 2010, 110).

(19) The dog\textsubscript{SUBJ\_AGENT} gave the cat\textsubscript{OBJ\_RECIPIENT} a mouse. (adapted from Leino 2010, 109)
(20) The cat gave the dog\textsubscript{OBJ\_RECIPIENT} a mouse. (adapted from Leino 2010, 109)

In contrast to that, in Finnish and other languages, which use morphological means to disambiguate the distribution of grammatical functions in the sentence, the situation is different. It is, therefore, suggested that since in Finnish word order does not influence the assignment of grammatical functions, it is possible to treat all word order patterns with the same grammatical functions and semantic roles as instances of the same ditransitive construction (Leino 2010, 106). The same argument could be applied to German, which also uses the grammatical category of case to express grammatical relations: as shown in (21) and (22) the subject, realized by a NP in the nominative case, remains the same even though the word order changes. However, even in German as a language with a rich case-marking system, in some cases the order of elements still has a direct effect on the assignment of grammatical functions, e.g. if the arguments are realized by proper names, because these nouns are not inflected according to case. In addition, also in these languages the information structure in the ditransitives and other argument structure constructions differs depending on the word order (Leino 2010, 109).

The discussion above indicates that English and German differ with respect to the importance that is attributed to word order in the argument structure configurations. This fact should be certainly reflected in the descriptions of the corresponding constructions in the two languages. Rosti\l a (2007, 62) suggests that the form of the argument structure constructions represents various language-dependent markings of arguments: while in English the argument assignment is linked to word order, in German the formal side of the constructions specifies the case markings. In both cases these markings are connected with the semantic
roles on the functional side. As shown in Figure 4, the English ditransitive construction is assumed to specify not only that the subject is mapped onto the agent but also its position in the clause, namely before the main verb, whereas in German the syntactic pole could include information on grammatical relations and the grammatical cases of NPs with which the arguments could be combined in actual utterances\(^\text{11}\) (see Figure 5). Both Figure 4 and Figure 5 represent prototypical declarative versions of English and German ditransitive constructions.

![Figure 4: English ditransitive construction](image)

![Figure 5: German ditransitive construction](image)

Therefore, based on the discussion of the previous literature, I argue that different word order patterns realizing the same mapping of semantic roles onto grammatical functions can be regarded as one argument structure construction. This is true for both English and German. Though in German the word order is generally more flexible, also English does allow different orderings, for instance in topicalizations. Disregarding whether the orderings are

\(^\text{11}\) Rosti (2007, 62-65) notes that the notation introduced in Goldberg (1995) and modified here can be used as long as it is made clear that grammatical functions belong to the constructional function rather than its formal side (see also Herbst 2011, 351-352). They are interpreted as related to the roles indicating perspectives on a situation (perspektivische Rollen), which are, in turn, linked to the distinctions made in the Cognitive Grammar such as figure-ground and trajectory-landmark (Rostila 2007, 55). At the same time, the author recognizes the necessity of referring to the formal realization of constituents using the functional categories as an abbreviated notation. An alternative would be to present an exhaustive list of all phrasal elements that could fill this slot, e.g. NPs and clauses for the slot corresponding to the subject combined with the information on the nominative case in German or word order in English (which is not treated as part of the argument structure construction by Goldberg) (Rostila 2007, 64-65).
conceptualized as allostructions (Cappelle 2006) or combinations of argument structures with additional ordering constructions (Kuningas and Leino 2006, 301), the word order has an effect on information structure. Therefore, it is important to consider all orders in which the arguments can be arranged in both English and German. Even though different orders are considered to correspond to the same argument structure construction, the analysis should take into account the variation in terms of the information structure.

Previous research has criticized postulation of such highly abstract constructions as subject-predicate construction or a “super-construction” with underspecified word order, in particular on the highest level of abstractness (cf. Hampe 2012, 9-10). Since their existence does not appear to be psychologically plausible (Blumenthal-Dramé 2012, 29), it is considered incompatible with one of the aims of CxG (cf. Hoffmann and Trousdale 2013, 3). To verify these hypotheses and, therefore, confirm the psychological plausibility of these form-meaning pairings it is necessary to conduct psycholinguistic experiments (Tomasello 2003, 98). However, this particular question is beyond the scope of the present study: the translation experiment conducted within the present study is designed to answer different research questions (see Chapter 6).

3.2 Reasons for selecting the Construction Grammar framework
In Chapter 3.1.1 criteria of adequacy have been applied to the family of constructionist approaches to review the general motivation for viewing language in terms of form-meaning pairings. Before it is discussed how the concepts of CxG are used in the context of translation studies, the present chapter concentrates on the arguments for using this particular theoretical framework for empirical analyses of translations. This initial motivation for applying CxG, which is also provided by the previous research in this area, is re-evaluated in Chapter 7 after the discussion of the empirical analysis reported in this study.

Szymańska (2011, 121) claims that CxG “may provide an inspiring conceptual framework and a range of precise tools to address in an integrated fashion the well-established intuitions and observations about translation as well as account for a range of phenomena known to influence the process and product of translation in a vital way” (cf. also Tabakowska 1993, 74). The fact that constructions provide a non-modular view of language combining various formal and functional levels of linguistic analysis is often mentioned as favorable both with respect to cross-linguistic studies (Boas 2010a, 15; Hilpert 2010, 25), as well as to
translation-related investigations (Szymańska 2011, 115, 124). Research on a particular phenomenon across two languages is not restricted by the borders of individual linguistic levels, and can account for all phonological, morphological, syntactic, semantic, pragmatic and discourse-functional factors that are deemed relevant for the construction(s) and language pair under analysis. However, this argument has been questioned by noting that a study can also encompass different levels without relating these to the notion of construction (Wiechmann 2012, 90-91). What is agreed upon is that a form-meaning pairing can be employed as a unit of cross-linguistic comparison, at least for genetically related languages (Wiechmann 2012, 91; cf. Boas 2010b, see Chapter 3.3.1 for further discussion). In the present study the assumed interconnection of constructions representing different levels of abstractness, which are organized in a hierarchical manner, is exploited by comparing more or less specific constructions of the same general type in English and German.

Moreover, the interaction of multiple layers of constructional properties can be used to account for complex phenomena of equivalence and translation shifts (Szymańska 2011, 117, 125-129), which can hold between translation units of various size. Ebeling (1998, 169) assumes that in order to establish equivalence translators would have to find similar patterns in the target language exhibiting most of the properties of the original construction. However, since constructions are language-specific (Croft 2001, 6; Goldberg 2003, 222; Szymańska 2011, 108), translators may come across “constructional resistance”, i.e. “the difficulty (sometimes impossibility) of finding a TL construct representing all the functional properties of a ST construct (semantic, pragmatic, discourse, stylistic) while keeping the formal similarity as well as producing a ‘natural’ TL expression” (Szymańska 2011, 127). The phenomenon of constructional resistance can be one of the factors leading to the introduction of translation shifts. A previous study on translations of resultative constructions from English into Spanish indicates that an investigation of these clause patterns outside of a constructionist framework would result in a list of transpositions, i.e. part of speech changes. As a result, it would possibly miss an important reason for these shifts, namely that this complex construction, which links a specific form to a resultative meaning in English, is not present in Spanish (Rojo and Valenzuela 2013, 286).

Another argument in favor of CxG as a framework for translation studies is its consideration of frequencies of various constructions. Szymańska (2011, 116, 130-132) suggests that this
usage-based aspect of the theory can be linked to the concept of naturalness in translation (cf. also Cappelle and Loock 2013, 255). Discussion related to naturalness based on intuition is very subjective, whereas frequency measures of the corresponding constructions in different languages, be it absolute, relative frequencies or more sophisticated statistical methods, could allow the researchers to operationalize the phenomenon of naturalness in more concrete and verifiable terms.

Furthermore, cognitive aspects of the constructionist theory of language can give important insights into the nature of translation phenomena. Awareness of the cognitive dimension is in line with translation process research, whose aim is to account for cognitive processing during translation. As mentioned in Chapter 2, translation process is considered among the factors that could influence the final translation product, the other ones being contrastive differences and register characteristics (Steiner 2001; Hansen-Schirra and Steiner 2012). It could be argued that cognitive processing is the central factor, since also contrastive and register divergences have to be processed by a translator to play a role in the translation. As observed in previous research, “the most striking consequence of adopting a cognitive perspective is that it places human cognition and human agency at the center of the causal picture” (Halverson 2013, 66). Therefore, it has been suggested that concepts of cognitive linguistics could provide a theoretical background required to explain the occurrence of translation shifts (Tabakowska 1993, 2013; Halverson 2007, 2013) and the existence of translation properties (Halverson 2003, 2010, 2013) by accounting for cognitive aspects involved in the process of translation. The corpus-to-cognition principle used in CxG, and e.g. the eye-to-mind hypothesis (Just and Carpenter 1980), which serves as a basis for the eye-tracking studies employed in translation process research, could certainly complement each other providing a more comprehensive picture of translation.

### 3.3 CxG as a framework for multilingual analysis

Within CxG it is claimed that all languages consist of constructions (Goldberg 2006, 9), and therefore, the constructionist approaches could be applied to the study of different languages. While most of the accounts within this theoretical framework concentrate on English (Boas 2010a, 3), descriptions of constructions existing in other languages, though not so numerous, are developed. In the previous sub-chapters research on German constructions has been considered as well. The next step is to compare constructions across several
languages and to test the potential of the theory in its applications to translation studies. In these investigations English is again often selected as the basis for comparison, and its rich description of constructional inventory is used as the starting point of a comparative analysis (Boas 2010a, 14). The present study considers the language pair English-German. As indicated in Chapter 4, constructions for analysis are identified heuristically both in English and German, so that the investigations are motivated by the previous research in both languages.

The chapter is divided into cross-linguistic and translation investigations conducted within the CxG framework. Since both types of research deal with multilingual analyses of constructions and could draw upon translations as a source of data, there is no sharp dividing line (Granger 2003). In this review, research using parallel corpora\(^{12}\) of originals and aligned translations as well as comparable corpora of originals and translations within the same language are considered in Chapter 3.3.2, which deals with the combination of CxG and translation studies. All other studies, even when working with translations of individual words, are discussed in Chapter 3.3.1, which describes previous work on combination of CxG and contrastive studies. It should be noted, however, that also contrastive comparisons employ parallel corpora (cf. Granger 2003, 21-22). One such study (Ebeling 1998) is included in 3.3.2, because its study design is closer to the investigations reported in this chapter.

### 3.3.1 CxG and contrastive studies

In the introduction to a recent collection of articles on cross-linguistic studies Boas (2010a) notes that there have been only a few attempts to apply constructional approaches to cross-linguistic comparisons. This fact is partly explained through an initial focus of CxG on English to account for both its core and periphery. Boas (2010a, 4-5) suggests that analyses of other languages have adopted the same line of research instead of considering possible cross-linguistic aspects. Furthermore, it is claimed that constructions are language-specific (Croft 2001, 6; Goldberg 2003, 222; Szymańska 2011, 108): this assumption has also an important effect on the number of cross-linguistic comparisons of constructions (Boas 2010a, 5). As has been shown during the detailed discussion of word order in Chapter 3.1.4, similar

---

\(^{12}\) Some case studies conduct manual analyses of individual parallel texts that have not been integrated into electronic corpora.
constructions across languages may display major differences with respect to their formal and/or functional parameters. This fact has to be considered to ensure comparability of items in the investigated languages.

Therefore, taking the language-specific properties of form-function pairings into account, the question arises whether the formal or the functional properties should have a priority for the purposes of comparison. Boas (2010a, 14-15) argues for taking the functional side of a construction as a basis for comparison to see how the same meaning is realized in different languages. Boas (2002, 2013) adopts a bottom-up way of comparing constructions by using the frame-semantic knowledge (Fillmore 2006) linked to individual constructions. After a detailed analysis of verbs or other lexical items in English with respect to the semantic frames they evoke and the syntactic patterns through which these frames are realized, he establishes the German translational equivalents of these verbs using monolingual and bilingual dictionaries. Sentences with these German verbs are then extracted from German corpora. Based on this data the syntactic patterns of these verbs and their associated semantic frames are analyzed. To complete the picture, the corpus evidence is supplemented by native-speaker’s intuitions. Boas suggests that this procedure can be used to create resources for human and machine translation (see also Čulo 2013 for a frame-semantic approach to the study of translation).

An example of a top-down functional analysis of constructions across languages is the study by Hilpert (2010), which starts by examining how particular meaning, namely the category of comparison, is realized in the languages English and Swedish. Since the two languages are genetically related, both contain two comparative constructions. However, their systematic comparison has revealed some differences in their morphological and syntactic properties. Hilpert (2010, 38) notes that the contrastive analysis “can help direct attention to relevant parameters that would not be detected in an analysis of a single language”. In other words, also the description of a single language may be further expanded by performing contrastive studies.

Whether one takes functional or formal common ground as a starting point depends on the research question. This decision does not necessarily have to imply that the corresponding pole is considered more central for analysis, though the opposite has been suggested (cf. Boas 2010a, 15). A study can reveal semantic differences between constructions that are formally
similar in several languages. Timyam and Bergen (2010) analyze caused-motion and ditransitive constructions in English and Thai: the study indicates that these constructions, which are identified in both languages based on the types of grammatical functions, are characterized, among other things, through differences on the functional pole. As in the scenarios described above with semantic similarities as the basis for comparison, the findings of this study could be applied in second language acquisition and translational practice.

Leino (2010, 131) suggests that corresponding constructions in any two languages have to share at least some of the features on both formal and functional poles. Acknowledging this methodological consideration while adopting construction as a unit of comparison may lead to an undesirable restriction of the analyzed data. In particular, in the studies of genetically non-related languages it might be difficult to find many comparable constructions (Wiechmann 2012, 91). Despite this limitation, the cross-linguistic analysis that considers constructions characterized by similar formal and functional features appears plausible as it follows the tenet of non-modularity: taking into account the fact that constructions do differ in some respects, a comparison of form-function pairings with similarities on both formal and functional poles comes closest to ensuring the comparability of units across languages.

3.3.2 CxG and translation studies

Previous applications of constructionist approaches to translations have investigated, among other things, possible reasons for the introduction of shifts in the process of translation. Since the phenomenon of translation shifts is in the focus of the present study, this sub-chapter will first provide an overview of the previous research in this area.

Halverson (2007, 113-115) suggests that different types of construal operations, described as ways to conceptualize extra-linguistic scenes and realize these in linguistic expressions (Croft and Cruse 2004, 40), should be taken into account in the investigation of translation shifts. This approach rests on the assumption that the same situation can be structured differently depending on what aspects are highlighted, from which perspective the scene is presented, how many details are mentioned (Langacker 1987, 116-137), etc. The study of translation shifts presented in Halverson (2007) is based on the classification of construal operations by Croft and Cruse (2004, 46), who relate each of the operations to a general
cognitive process. To name probably the most straightforward example, it is assumed that modulation (Vinay and Darbelnet 1995, 36) could depend on a number of construal operations, including a viewpoint of the conceptualizer (Halverson 2007, 114). Against this theoretical background, three possible types of reasons for translation shifts are proposed, namely reasons connected to cognition, convention and context (Halverson 2007, 118-119; 2013, 49). An example of the first type would be the gravitational pull hypothesis, described below.

Tabakowska (1993, 2013) also links translation shifts to the construal operations and stresses the fact that cross-linguistically constructions may differ in terms of scene conceptualizations. Since a scene is recognized as a translation unit, its realization in the translation using the linguistic material of the TL may lead to various translation shifts (Tabakowska 1993, 73-77). The notion of constructional resistance (Szymańska 2011, 127) or constructional mismatch (Rojo and Valenzuela 2013, 289) exploits the same idea but from the perspective of constructions rather than conceptual structure: it is assumed that constructions in different languages are characterized through different bundles of formal and functional properties. These partial matches between linguistic structures in several languages are used to explain the phenomenon of translation shifts (Szymańska 2011, 215).

As mentioned earlier, a case of constructional resistance has been identified in translations of English resultative constructions into Spanish as there appears to be no corresponding construction in the TL (Rojo and Valenzuela 2013). In an eye-tracking study the subjects were asked to translate at sight short texts with instances of the resultative construction (e.g. he hammered the handle straight) and the so-called predicative construction with a somewhat similar form mapped onto a different meaning (e.g. he hammered the handle until it was straight). The study showed significant differences in eye movements and pupil size for the translation of these two constructions. This result is considered to be an indicator of different effects of these structures on cognitive processing suggesting that complex resultative constructions in English are psychologically real and require more effort to be translated into Spanish (Rojo and Valenzuela 2013, 302). Szymańska recognizes the potential of a CxG application to translation studies in a systematic account of all translation shifts occurring in texts between constructions of different size and level of abstractness. She admits that the
complete inventory of constructions in both languages should be investigated first (Szymańska 2011, 216, see a suggestion on how it could be created in Boas 2010a, 11).

Existential there-constructions have been studied in a cross-linguistic perspective for several language pairs, for instance English-Norwegian (Ebeling 1998), and English-French (Cappelle and Loock 2013). The former investigation uses a bi-directional parallel corpus of originals and translations to learn more about the properties of the corresponding constructions in English and Norwegian originals. The results have shown that in most cases the constructions that have been identified as corresponding are used in originals and their translations. However, in approximately a third of examples in both translation directions the analyzed construction is translated by another structure, indicating that not all instances of the constructions “are felt to be equivalent by the translators” (Ebeling 1998, 174). It is shown that even if two languages have corresponding constructions, the translator cannot always simply rely on them having a one-to-one match as suggested by Rojo and Valenzuela (2013, 289). Assuming that the observed shifts are in fact related to certain properties of the original construction rather than other factors, this could be considered as a potential case of constructional resistance, when some of the properties of the original construction do not match either the TT pattern most closely associated with the original or other alternatives in the TL. The translator, therefore, has to decide what formal or functional aspects of the original constructions should be realized in the translation (Szymańska 2011, 127). Ebeling (1998) assumes that the resulting translation shifts can indicate what features of the source language construction are more central than others, since these are likely to be kept in the aligned translations, whereas others are more easily omitted. Thus, the focus of Ebeling’s study is not on the nature of translations as a special sub-class of texts distinguished from the originals by a set of features, but rather on contrastive characteristics of the two languages under analysis.

In contrast to this, the aim of the study by Cappelle and Loock (2013) is the analysis of the translation property of shining through. After establishing a contrastive difference, which is reflected in the frequencies with which the corresponding constructions are used in the originals, the authors compare the frequencies of the there-construction both in translated and non-translated English and French, and complement this comparison by an analysis of originals and corresponding translations. It should be noted that the study does not explicitly
refer to the framework of CxG, but it still accounts for a complex linguistic pattern and its role in translations. This study is an example of how translation properties can be operationalized in terms of constructions found in originals and translations.

Based on the CxG framework Szymańska (2011) models the translation process as involving the following steps. It is assumed that the translator starts by processing the ST: during this interpretation the ST constructions with all the corresponding formal and functional properties are identified. The model suggests that on the basis of this information a cognitive map of the ST, which contains rich information on these form-meaning pairings and associated cultural knowledge, is created. This map together with the awareness of the function of translation, of the target culture and of other possible factors influencing the translation is said to feed into the cognitive map of the TT. Finally, according to this model, the production of individual linguistic structures in the TL is influenced both by the TT map and the individual sentences in the ST. In other words, following the model by Holmes (1988), two levels are distinguished, a linear, i.e. local, level of ST and TT, which involves a sequential process of constructions’ comprehension and production, and a global plane of cognitive maps. The relation between the linear plane of the TT and the TT map is dynamic: the TT map is believed to store information on the features of the original constructions that are given priority over the others by the translator, and therefore, it dictates what has to be re-expressed in the linear TT production. However, since the latter plane is constantly compared to the parallel level of the ST, the produced TT sentences lead to an update of the TT map (Szymańska 2011, 135-140). The author acknowledges the fact that the original is not always read and analyzed completely prior to translation, and indicates that the TT map is then incomplete and schematic (Szymańska 2011, 139-140). The described model of the translation process pays a lot of attention on the interrelation of the linear level of the ST, the TT map and the linear production of the TT. The ST map does not seem to play a central role in these conceptualizations: for instance, one would expect this map to be dynamic as well, since the linear comprehension of the original, paying more attention to minor details, may result in the reinterpretation of the ST on the global level, having in turn consequences for the translation. However, this aspect is not taken into account, so that the role of this initial map is not sufficiently defined.
The model has been used for qualitative analyses of literary texts and subtitles (Szymańska 2011): even though some of the case studies concentrate on multiple examples of the same construction and its corresponding translations, no quantitative information is provided. Moreover, the analyses of all examples involve their characterization in the broader context of the corresponding works. This perspective is necessary to account for the global plane postulated by the model, which cannot be considered to this extent in quantitative studies. It has not been specified how the subdivision of the translation process into two planes in both languages can be tested empirically. Moreover, even though the model is postulated to reflect the cognitive processing during translation, it is not based on any empirical translation process research. Thus, the model has some drawbacks, but the general idea of applying a constructionist approach to translations is considered fruitful (see Chapter 3.2). The present study performs quantitative analyses of constructions based on the concepts of constructional resistance and the resulting partial matching of the ST and TT constructions introduced by Szymańska (2011).

A further model of the translation process combines the framework of Cognitive Grammar with the cognitive theory of bilingualism (Halverson 2003, 2010). In accordance with both theoretical foundations it is claimed that the formal features of a linguistic construction lead to activation of semantic networks of meaning, which, as is postulated in the model, are shared for all languages. The activated network contains nodes for formal representation of this knowledge in different languages, which a person is familiar with. Certain items within the semantic network are believed to exercise “gravitational pull” (Halverson 2003, 218). These structures are prototypical members of categories, e.g. the most typical chair (cf. e.g. Rosch and Mervis 1975), and the highest level schemas representing the superordinate nodes in the conceptual hierarchy, e.g. a concept of a tree as related to concepts of oak, maple and elm (Langacker 1987, 68). Halverson (2003, 2010) assumes that due to gravitational pull translations are characterized through an increased number of linguistic structures associated with these conceptually more central items. Findings with respect to various translation properties are linked to the gravitational pull hypothesis. For instance, the fact that “the proportion of high frequency words versus low frequency words is relatively higher in translated texts”, interpreted as one of the indicators for the translation property of simplification (Laviosa 2002, 62), is seen as a proof of the gravitational pull towards lexical
items that are used more frequently in the target language, and could have thus acquired the status of prototype (Halverson 2003, 219). Since an explicit connection between the model and previous empirical findings is created, and further steps for testing the postulated hypothesis are proposed (Halverson 2003, 233-234), the model has certainly more empirical validation than the one proposed by Szymańska (2011). Recently Halverson (2010, 358) has stressed that further development of the model is required to separate several “causal factors related to different kinds of cognitive linguistic knowledge”, which have been originally fused.

The present study goes beyond previous investigations in this area, as a quantitative analysis of a wider range of constructions will provide a basis for generalizations on potential problems the translators may encounter dealing with constructions on clause level. Moreover, a combination of corpus-based research and experimental methods can help account for constructions both in translation product and process. For instance, it can be used to investigate constructions’ frequencies not only in originals and translations, but also in the intermediate versions.

3.4 Conclusion

According to the family of constructionist approaches, language is a structured inventory of constructions, i.e. form-meaning pairings. Constructions of different sizes and levels of complexity are considered to exist on all levels of linguistic description. For instance, the argument structure constructions investigated in the present study are complex clause-level patterns. It is suggested that repeated usage of constructions leads to their cognitive entrenchment: when functions of a construction are conceived as especially suitable for particular contexts, language users tend to use the construction more frequently and, as a result, it gradually turns into a unit. Such holistically stored linguistic patterns are assumed to facilitate cognitive processing during language production and perception. Following the corpus to cognition principle formulated in usage-based constructionist accounts, frequencies in corpora are investigated to gain insights into cognition. However, a combination of evidence from corpus-based studies, psycholinguistic and neurological experiments increases the chances of arriving at psychologically plausible hypotheses about cognitive processing.
Constructions are considered to be language-specific: even if they are to some extent comparable in the two languages involved, there are still some differences in their formal or functional aspects. Therefore, translators come across constructional resistance and have to set priorities concerning those features of the original pattern, which have to be kept in the translation. It has been shown that when a particular construction has no corresponding construction in the TL, its translation requires more cognitive effort (Rojo and Valenzuela 2013). Other constructions exist in both the source and the target languages but are characterized by different frequencies of use (Cappelle and Loock 2013). Assuming that such form-meaning pairings are entrenched to different degrees in the languages under investigation, it is worth studying what translation shifts and translation properties are present in the translation process and product data, as well as the amount of cognitive processing during their translation. The present study concentrates on the translation of argument structure constructions and, taking into account the results of the previous studies reviewed in this and the previous chapter, investigates how various translation phenomena can be described and explained using the notion of construction.
4. Methodology

Due to the focus of the present study on corpus analyses, the methodology discussed in this chapter is related to the corpus-based investigations reported in Chapter 5. Chapter 4.1 provides a description of different corpora types and their applications in translation studies concentrating particularly on the CroCo corpus used for the present research. It is followed by Chapter 4.2, which reports on the methodological decisions that have led to the selection of the constructions for a further analysis. Chapter 4.3 deals with the types of analyses that are relevant for all constructions, postponing the discussion of more specific methodological issues related to individual patterns to Chapter 5.

Some of the related research questions that cannot be sufficiently answered with the help of corpus investigations are examined in a translation experiment reported on in Chapter 6. The pilot experiment is based on a small data sample and presents a future perspective rather than an attempt at any generalizations connected to the cognitive processing of constructions during the translation process. For this reason, the methodology for the experiment is discussed in Chapter 6.

4.1 Corpora in translation studies

A corpus can be defined as a large sample of authentic machine-readable texts collected based on a set of criteria, so as to be representative of a particular language or language variety (McEnery, Xiao and Tono 2006, 4-5). Therefore, one of the advantages of corpus linguistics is the potential to study naturally occurring texts in large quantities. It allows researchers to compare different types of language use, be it written vs. spoken language, British vs. Australian English, learner vs. native speaker English or originals vs. translations. To study the linguistic variation analysts do not need to rely solely on their own or the informant’s intuition, but can abstract from individual idiolects to investigate the whole range of linguistic features associated with the population of analyzed texts (McEnery, Xiao and Tono 2006, 6). Similar to other types of empirical methods, a corpus can provide reliable quantitative results that can be generalized from the analyzed sample to the language variety represented by the corpus used in the analysis (McEnery and Wilson 2001, 76-77; McEnery, Xiao and Tono 2006, 19). These, in turn are typically supplemented by more detailed
qualitative investigations. Whereas the quantitative analyses are typically applied to test the hypotheses, the qualitative aspects help us interpret the findings and formulate new assumptions leading to further development or refinement of a theory (Rasinger 2008, 17). Following Baker (1993), corpus methodology has been widely applied in the field of translation studies. The main aspects of this research area are illustrated through the following description of the CroCo corpus, used in the present work.

The CroCo corpus (Hansen-Schirra, Neumann and Steiner 2012) includes English originals (EO) and their German translations (GTrans) along with German originals (GO) and the corresponding English translations (ETrans). It can be used as a multilingual parallel corpus to compare originals and their translations (e.g. English originals and the corresponding translations into German), a monolingual comparable corpus to examine originals and translations produced in the same language (e.g. English originals and English translations) or a multilingual comparable corpus to compare e.g. English and German originals (Neumann and Hansen-Schirra 2012, 26-27; see also Granger 2003, 19-21 on different classifications of corpora in contrastive linguistics and translation studies).

The required type of comparison depends on the specific research questions of individual studies. For instance, a multilingual parallel corpus can be applied to the investigation of equivalence and translation shifts (see Chapters 2.2 and 2.3). The aligned pairs of originals and translations allow the identification of what is changed during translations. A monolingual comparable corpus was mentioned in the preceding chapters in the context of translation properties (see Chapter 2.5). In order to investigate the manner in which translations are different from the comparable originals, the use of various linguistic features linked to the abstract level of translation properties is compared across these varieties. Finally, to account for contrastive differences in the language pair under analysis, a comparative study of originals in both languages can be performed. Since the design of the CroCo corpus allows for all three types of comparisons, the corpus is perfectly suited for the various analyses performed within the present study (see Chapter 4.3).

The main part of the corpus consists of eight sub-corpora each representing a separate register. These include the registers of political essays (ESSAY), fictional texts (FICTION), instruction manuals (INSTR), popular-scientific texts (POPSCI), letters to shareholders (SHARE), prepared speeches (SPEECH), tourism leaflets (TOU) and websites (WEB). Each
register consists of at least ten texts with the total of 31 250 words. The overall size of the core corpus is approximately one million words. In addition, the CroCo corpus includes reference corpora in English and German (ER and GR) consisting of 2000 word samples from 17 registers (Neumann and Hansen-Schirra 2012, 27-32). While one might argue that the corpus of a large size is more likely to include rare phenomena, it is essential to keep in mind that the size of a corpus has a direct impact on the statistical tests performed on its basis. This means that a larger corpus makes it more likely for the alternative hypotheses to be accepted (Baroni and Evert 2009, 783; Neumann and Hansen-Schirra 2012, 31).

The main data sample analyzed within this study has been limited to the registers of popular-scientific discourse and political essays due to the extensive manual correction of the alignment and annotation (see Chapter 4.2 for more details). The first of these registers contains texts, which “aim at informing a general audience in a clear and comprehensible way on findings and developments of a scientific field” (Neumann and Hansen-Schirra 2012, 29; see also Hyland 2010, 118-119; Niederhauser 1999, 48). It is important to mention that the selection of texts in this register differs for the two translation directions. While the English originals and the German translations contain only popular-scientific journal articles, the opposite translation direction includes extracts from popular-scientific books from a larger time period. Neumann and Hansen-Schirra (2012, 29) observe that the German popular-scientific journal articles are less likely to be translated into English. Indeed, the difficulty to sample comparable texts in both translation directions has been recognized as a persistent problem of parallel bidirectional corpora (Johansson 1998, 6; Aijmer 2008, 278). Three texts for the direction English-German have been taken from the corpus compiled by Doherty. This explains a partial overlap of the discussed examples in this study and the research reported by Doherty (1996). It should also be added that the three translations from Doherty’s corpus have been optimized in terms of information structure. This fact is considered in the discussion of the analyzed constructions.

The second analyzed register involves texts on various political topics, for instance on economic issues, which are “often written for a foreign newspaper or other publication and are subsequently published in both English and German by a governmental body” (Neumann and Hansen-Schirra 2012, 27). As indicated in the meta-information files of these texts, they are produced either by (or at least on behalf of) politicians or journalists responsible e.g. for
the business section of a newspaper (see discussion in Chapter 5.1.2). Therefore, the texts from both registers considered here are to some extent influenced by the journalistic prose (e.g. Gläser 1990: 181, 186, 209; Hyland 2010: 120; Fischer 2013: 26). This makes the comparison between the two registers particularly interesting.

In order to perform more sophisticated analyses, it is important to be able to systematically extract information from a corpus related to various levels of linguistic analysis. It is possible to query for these linguistic features if the corpus has been annotated, that is, if appropriate values have been automatically or manually assigned to the textual material in the corpus (McEnery, Xiao and Tono 2006, 29). The CroCo corpus has been automatically annotated for parts of speech, morphology and phrase structure as well as manually for grammatical functions. Moreover, the texts in the corpus have been aligned on the word, chunk, clause, and sentence levels (Hansen-Schirra and Neumann 2012, 36-38). This allows us to extract aligned stretches of text on one or several of these levels. Detailed information on the levels of annotation and alignment relevant for this study are provided below (Chapter 4.2.1).

4.2 Identification of constructions for analysis

The present study starts by exploring the corpus in a heuristic way (Čulo et al. 2012, 91): it identifies the phenomena in the data, which are potentially interesting, rather than starting with a pre-selected set of linguistic structures discussed in the previous studies. Since our research concentrates on construction shifts, the first step involves an operationalization of this phenomenon. Those constructions that are frequently affected by the construction shifts then form the basis for selecting constructions for a further analysis.

4.2.1 Extraction of construction shifts

Čulo et al. (2012) suggest that the alignment links between different levels of analysis, e.g. between words and grammatical functions in the originals and translations could be taken as a starting point of a heuristic analysis of the corpus. Cases in which the meaning of individual words is mapped onto different grammatical functions in the original and the corresponding translation are a type of what can be defined as “crossing lines” (Čulo et al. 2012, 92, 110). For instance, in (23) the prepositional phrase at Wellcome Research Laboratories postmodifies the noun scientists, the whole NP functions as the subject of the sentence. In the translation the corresponding prepositional phrase in den Wellcome Research Laboratories is
mapped, however, onto the local adverbial rather than the subject, since the noun *scientists* is translated as the impersonal pronoun *man*.

(23) **EO:** [At the same time]_{adv_temp} [scientists at Wellcome Research Laboratories]_{subj} [found]_{fin} [that the interferon produced by human white blood cells contained a mixture of the same interferons identified by the genetic engineers]_{dobj}.

(EO_POPSCI_010 s33)

**GTrans:** [Zur selben Zeit]_{adv_temp} [entdeckte]_{fin} [man]_{subj} [in den Wellcome Research Laboratories]_{adv_loc}, [daß [sic] das Interferon der weißen Blutkörperchen des Menschen aus eben jenen von den Gentechnikern identifizierten Alpha-Interferonen besteht]_{dobj}.

Translation pairs characterized through such crossing lines have been extracted from the CroCo corpus with the help of a Perl script. The output of the script is a csv file, which consists of the text ID, the source and target sentence IDs, the complete source and target sentences and the grammatical functions of the source and target sentences. A second script created another table containing the same information for all aligned source and target texts in the corpus, which is used to compare instances with and without construction shifts at a later stage. Both scripts use the multi-level alignment of the CroCo Corpus, particularly on the levels of words and grammatical functions, and the annotation of grammatical functions, which has been performed manually using MMAX2 (Hansen-Schirra and Neumann 2012, 36). The functional analysis of the texts in the corpus has been performed on different levels of the sentence. For instance, the English original in (24), analyzed on the highest level, consists of a subject, a finite verb and a direct object, whereas the annotation on the lower level would analyze the *that*-clause further as including another set of a subject, a finite verb, a predicative, a direct object and a modal adverbial. Similarly, the German translation can be parsed only on the level of the main clause or taking into account a deeper level of the subordinate clause. The search for crossing lines considers only the alignment between elements on the level of the main clause. Otherwise, sentence pairs such as the one shown in (24), would not be retrieved as instances of crossing lines. In this example, the information mapped onto the direct object of the main clause in the original is translated through the subject of the main clause but the original and the translation have the same subject within the subordinate clause, namely *WTO members* in English and *die WTO-Mitgliedstaaten* in German. The translation shift in the main clause is, however, an indicator of a change in perspective, which deserves further analysis. Therefore, the study concentrates on the main
clause and considers subordinate clauses as chunks with the corresponding grammatical functions on the level of the main clause.

(24) **EO:** [The ministerial declaration]_{subj} [states]_{fin} [that [WTO members]_{subj} [will]_{fin} [negotiate]_{pred} [reductions of export subsidies]_{dobj} ["with a view to phasing them out."]_{adv_mod}_{dobj} (EO_ESSAY_007 s26)

**GTrans:** [In der Erklärung des Ministertreffens]_{adv_loc} [wurde]_{fin} [festgestellt]_{pred}, [daß]_{sic} [die WTO-Mitgliedstaaten]_{subj} [den Abbau von Exportsubventionen]_{dobj} ["im Hinblick auf ihre schrittweise Abschaffung"]_{adv_mod} [aushandeln]_{pred} [werden.]_{fin}_{subj}

The annotation of grammatical functions contains a certain number of inconsistencies. The manual functional analysis of chunks has as a consequence that certain decisions are subjective even though the annotators have followed the common guidelines (cf. Hansen-Schirra, Neumann and Kunz 2010). The error rate for this layer of annotation is about 5% (Hansen-Schirra and Neumann 2012, 51). This could have an effect on precision and recall rates, i.e. how many items have been extracted even though they are not the targeted items or not extracted even though they meet the criteria (Manning and Schütze 1999, 268–269). Furthermore, since the script retrieves aligned pairs of chunks, the instances of wrong alignment skew the data. Unfortunately the level of word alignment is especially prone to mistakes, as the alignment tool GIZA++ used in the CroCo corpus has not the required level of accuracy. “The automatic word aligner reaches 78.1% precision and 62.8% recall” (Hansen-Schirra and Neumann 2012, 51). On the chunk level the alignment has been achieved by mapping grammatical functions onto each other (Hansen-Schirra and Neumann 2012, 38). Since the script extracts those sentences that have crossing lines between words and grammatical functions, the precision and recall of this analysis are affected by possible mistakes in both word and chunk alignment. The precision increased as a result of subsequent, partly automatized identification of potentially problematic sentences. For instance, a very long list of grammatical functions suggested that the extracted alignment has been carried out on a lower level of analysis, whereas sentences without subjects or finite verbs could indicate that the annotation contains mistakes (even though there are certain structures in which these functions are missing). This step was followed by a manual analysis of such cases and, if deemed necessary, adjustment of annotation and/or alignment. While these methodological steps improved the precision of the analyzed sample, it should be noted
that, as with all linguistic analyses, some decisions were rather subjective. In future work the categories could be further validated by cross-checking the annotation by several analysts. Identification of crossing lines presupposes that one source sentence is aligned with one target sentence. However, the data also contains cases of sentence splitting and sentence merge, i.e. instances where one source sentence corresponds to two or more sentences in the translation, or when several original sentences are merged into one in the translation (Ramm 2004, 129). While the corpus contains one-to-many or many-to-one alignment links on the sentence level, the script based the retrieval of crossing lines on one of the sentences from the aligned set in the original or the translation. The (missing) IDs of the target sentences contained in the output table were used to determine cases of potential sentence splitting. In the next step the missing sentences were manually retrieved from the CroCo corpus using the corpus query processor (cqp) (Evert and Hardie 2011) (see Chapter 5.1.1 for further discussion).

The crossing lines involve different realizations of the semantic information in the original and the corresponding translation, which could lead to a change in semantic perspective or a shift to a more impersonal structure (see e.g. Chapter 5.1.1). Therefore, the presence of a crossing line between words and grammatical functions, used as one of the concrete indicators for a construction shift, is linked to the change on the functional side of the original construction. This indicator is then combined with changes within the sets of grammatical functions that are used to refer to the formal side of the construction. To apply this criterion, the results of the script are classified further: the patterns of the grammatical functions present in the source sentences are first quantified with the help of the concordance software AntConc. To reduce the variation of potential patterns, the source sentences have been grouped together according to the first four grammatical functions, four being the statistical mode of the construction length, measured in the number of grammatical functions per sentence. It should, however, be mentioned that a number of elements have been merged or not considered as the basis for the classification.

First of all, the categories of negation, particle and conjunction are not included in the list of grammatical functions used for the classification of sentences according to constructions. Because of this methodological decision, sentence in (25) would be an example of the same construction even if the initial conjunction was not present. These functional elements are
typically realized by function words and are not considered to change the argument structure constructions they are added to.

(25) **EO**: [But] \text{conj} [recently] \text{adv_temp} [researchers] \text{subj} [have] \text{fin} [discovered] \text{pred} [that fluoride can affect metal-free enzymes.] \text{dobj} (EO\_POPSCI\_009 s123)

Secondly, the functions of finite verb and predicat are merged together. Since as Goldberg (2003: 221) suggests, a sentence is a combination of several constructions, the division of predicates into finite and non-finite forms can be interpreted as a combination of argument-structure and e.g. future or modal constructions. Since this study concentrates only on argument structure constructions, this division could be disregarded for the classification of constructions. For instance, rather than classifying the construction in (25) as [Adverbial Subject Finite Predicative], I consider it as an example of [Adverbial Subject Finite Direct Object]. As a result, the classification takes into account an important grammatical element, namely the direct object, which would otherwise be discarded. While there are still a number of sentences, which consist of more than four grammatical functions even after the described changes, these methodological decisions still improve the resulting classification. Furthermore, all orders of the first four elements are treated as instances of the same grammatical structures (see Chapter 3 for a detailed discussion). In another step of the study, however, it is important to investigate whether word order variation interrelates with the analyzed phenomena. This methodological procedure is not too exclusive from the start but does not ignore a possible contribution of word order.

To conclude, in the context of the present study a translation pair is classified as an instance of a construction shift when both of the following conditions are met: 1) there is a difference in the first four grammatical functions between the original and the translation, and 2) there is at least one crossing line between words and grammatical functions, which affects functions other than conjunction, particle or negation. Furthermore, the cases when a crossing line is present between a finite and a non-finite verb form are not considered. Taking crossing lines as one of the two indicators of construction shifts I excluded cases in which the mapping of semantic information remains constant but some elements are left out or are added in the translation. For instance, in (26) the first sentence of the original and the corresponding target sentence do not belong to the translation pairs characterized by construction shifts, even though the translated sentence has a different set of grammatical
functions from the original. The temporal adverbial present in the translation does not have any correspondence in the aligned sentence and, therefore, cannot be affected by a crossing line. Instead, it is an instance of an empty link on the level of grammatical functions (Čulo et al. 2012, 92). As can be seen in (26), the temporal information, which is added in the first aligned German sentence, corresponds to the adverbial in the second English source sentence.

(26) **EO:** [But]_{conj} [Woo Suk Hwang of Seoul National University and his colleagues]_{subj} [proved]_{fin} [that it could be done.]_{dobj} The Korean team announced this past February that they had created a human embryo through SCNT, grew it into a blastocyst and derived a pluripotent ES cell line. (EO_POPSCI_006 s68-69)

**GTrans:** [Im Februar 2004]_{adv_temp} [verkündete]_{fin} [jedoch]_{conj} [Woo Suk Hwang mit seinen Kollegen an der Nationalen Universität von Seoul]_{subj} [einen Erfolg.]_{dobj} Sie hatten einen humanen Kerntransfer-Embryo erzeugt, ihn zum Keimbläschen heranwachsen lassen und aus der inneren Zellmasse erstmals auf diesem Wege eine humane pluripotente ES-Zelllinie gewonnen.

The present study concentrates on more critical changes between constructions used in the original and the corresponding translation by considering the instances of construction shifts, which exhibit not only different types of grammatical functions but also a rearrangement of semantic information within a sentence. Inclusion of crossing lines among the criteria for the identification of construction shifts also excludes cases in which the first four elements of the aligned clauses are different only due to differences in word order. For instance, in (27) the connective adverbial *dagegen* ‘in contrast’ is among the first four grammatical functions, whereas in the translation the corresponding element *however* is the fifth function. Thus, based on the formal criterion the translation could be classified as an instance of a different construction. However, since the sentence pair does not have any crossing lines, it is placed in the group without construction shifts.

(27) **GO:** Andere Unternehmen konnten [dagegen]_{adv_con} ihre Position in jüngerer Zeit wieder stärken. (GO_ESSAY_004 s39)

**ETrans:** Other companies have recently managed to strengthen their position, [however.]_{adv_con}

Therefore, both criteria for the identification of construction shifts are crucial to ensure that the formal and functional sides of the source language construction are affected in the process of translation. Based on the frequency of construction shifts and related contrastive
differences, three constructions have been selected for a more detailed analysis. Before they are introduced in Chapter 4.2.3, a number of further methodological decisions connected to the annotation and alignment of the data, which have an effect on the frequencies of construction shifts, are discussed in Chapter 4.2.2.

4.2.2 Special cases of annotation and alignment

This chapter deals with problem cases related to the annotation and alignment of the data. While in some aspects the study follows the decisions made during the compilation of the CroCo corpus, in others it deviates from the guidelines. Moreover, I link the decisions to the classification of constructions into the groups with and without construction shifts, and argue why certain sentences have to be excluded from the final data samples.

The discussion starts with a set of special cases involving expressions of different degrees of compositionality. To begin with, sentences in which the alignment on the word level is difficult due to non-compositional idiomatic expressions are considered as instances without crossing lines. For instance, in (28) the translation contains the idiomatic expression *goes without saying*, which corresponds to the chunk *ist allgemein bekannt* in the original. The alignment between, for example, the direct object *saying* and the complement *allgemein bekannt* does not appear meaningful, just as with the initial parsing of the unit *goes without saying* into smaller components. Since the expression is listed in a dictionary (Longman online), one could argue that it has reached a high degree of fixedness.

(28) **GO:** Es [ist]_fn [allgemein bekannt]_compl, solche Entscheidungen werden nicht aus dem Bauch herausgetroffen. (GO_ESSAY_016 s18)

**ETrans:** It *goes without*_fn *[saying]*_dobj that such decisions are not made off the top of anybody's head, but neither are they altogether free from what are known as "soft factors."

It should be mentioned that the annotation of this English sentence on the level of its constituents contains an example of a phrasal verb, namely *go without*. According to the annotation guidelines of the CroCo corpus (Hansen-Schirra, Neumann and Kunz 2010), the phrasal verbs listed as such in a dictionary are annotated together as one unit, and the following nouns are considered to function as the direct objects. Therefore, the annotation takes into account that certain combinations of verbs and prepositions have reached a high degree of lexicalization. However, the category of phrasal verbs is present only in the English
data, whereas all sentences with prepositional verbs in German are annotated as consisting of verbs followed by prepositional objects (see also Chapter 5.3). In example (29) the same semantic information is mapped onto the direct object in the original and the prepositional object in the translation. Therefore, this could be considered an example of a crossing line between words and grammatical functions. However, this crossing line is created by an introduction of an additional category in English reflecting the fixed nature of the expression. It could also be argued that the prepositional verbs in German, such as bestehen aus, represent rather fixed combinations of verbs and prepositions that are listed in the dictionary (e.g. the expression aus etwas bestehen is recorded as the second verb sense of bestehen in the DWDS). Taking this into account, the present study does not consider translation pairs such as (29), i.e. these sentences are not analyzed further and do not contribute to the frequencies of the corresponding constructions either with or without construction shifts. This decision affects only pairs of sentences, where direct objects preceded by phrasal verbs in English are aligned with prepositional objects in German, as their analysis in terms of construction shifts is considered problematic.

These sentences are treated differently than examples such as (26) and (28) because the latter do not contain any crossing lines and could be, therefore, more readily classified as examples without construction shifts. For instance, the difficulty of creating a direct alignment link between allgemein bekannt and saying in (28) is not a consequence of a different classification of English and German sentences, but rather of a complex fixed expression used in the translation. The non-compositional nature of the English expression goes without saying prevents us from identifying a crossing line on the level of individual grammatical functions. In contrast, sentences such as (29) do contain a crossing line between direct and prepositional objects, which are, however, potentially not justified in all examples.

(29) **EO:** The hydrogen atom [consists of] fin [a nucleus containing a single proton with an electron in orbit around it.] _dobj_ (EO_POPSCI_009 s10)

**GTrans:** Das Wasserstoffatom [besteht] fin [aus einem Kern mit einem einzelnen Proton und einer Hülle mit einem Elektron.] _probj_

(30) **EO:** Countries will continue to make such decisions for themselves. (EO_ESSAY_017 s27)

---

13 It should be mentioned that this decision does not significantly change the overall frequencies of the constructions. For instance, in the translation direction English-German only nine examples of the construction [Subject Verb Direct Object] involving phrasal verbs have been excluded from further analysis, compared to 275 cases of this argument structure construction present in the final data sample.
Moreover, the discussion of (non-)compositionality of certain expressions and their parsing into smaller chunks should consider instances of light verb constructions, which consist of a verb and a noun phrase, but are annotated as one chunk. For instance, the expression *make such decisions* in (30) is an example of a light verb construction. In these cases the whole verb phrase can be substituted by another verb (Quirk et al. 1985, 751; Biber et al. 1999, 1027; Huddleston and Pullum 2002, 290), which is lexically related to the noun. For instance, *make decisions* could be replaced by the verb *decide*, even though there are some meaning differences between the two options (Quirk et al. 1985, 751; Biber et al. 1999, 428). Grammars list the English verbs *take, make, have, do or give* as typical light verbs (Biber et al. 1999, 428; Huddleston and Pullum 2002, 290-296). In combination with certain nouns, these verbs do not carry meaning. As we can see in (30), it is the noun *decision* that refers to an action of deciding rather than the verb *make*. The use of the light verb construction could be motivated by the potential this pattern provides to additionally modify the noun (Huddleston and Pullum 2002, 291). For instance, in the discussed example the semi-determiner *such* works cohesively with the previous text passage.

Idiomatic combinations of verbs and nouns vary in their compositionality. While in some cases a phrase such as *have a look* refers to an action denoted by the noun and has, therefore, a different meaning than the verb *look* used in different combinations, the meaning of both elements could also be retained, for instance in the expression *take a snack* (Biber et al. 1999, 1026-1027). The former are likely to be considered examples of light verb constructions. By contrast, the latter could be described as compositional collocational patterns. However, many patterns occupy intermediate positions in terms of their compositionality (Biber et al. 1999, 1027). The present study relies on the list of the light verb constructions provided in Huddleston and Pullum (2002, 293-296) for English and Helbig and Buscha (2001, 70-83) for German.

In the present study the non-finite clause following the lexical verb is initially treated as one chunk annotated as a direct object or a complement, depending on whether the verb belongs to the category of linking verbs (Hansen-Schirra, Neumann and Kunz 2010). For instance, in (30) the infinitive clause *to make such decisions for themselves* is classified as the direct object.
However, a group of such sentences has been shown to be rather heterogeneous, as is discussed in Chapter 5.1.

Furthermore, most of the coordinated sentences, such as the one in (31), in both source and target texts have not been included into the final data samples, because the script could not automatically differentiate between several main clauses and assign the crossing lines to one of these. Since the study concentrates on the main clause as the level of analysis, this step would be necessary to establish which of the coordinated sentences belongs to the group with construction shifts. For the registers of popular scientific texts and political essays, 538 coordinated sentences have been identified in the translation direction English-German (18% of the total of 2938 sentences belonging to EO_ESSAY and EO_POPSCI) and 494 sentences have been classified as coordinated in the opposite direction (15% of the total of 3324 sentences belonging to GO_ESSAY and GO_POPSCI). Therefore, the interpretation of results of this study should take into account that these sentences have been excluded from the data samples.

(31) **GO:** [Die Analyse soll unter anderem dem Erkennen von Realität dienen,] _main clause_ [das übertragungsbedingte Verkennen ist aber erst einmal nötig, damit ein veränderndes Erkennen folgen kann.] _main clause_ (GO_POPSCI_007 s59)

(32) **GO:** [Vielleicht kann Deutschland wie die Skandinavier die Arbeitslosigkeit bezähmen] _main clause_ und [die Staatsfinanzen ordnen.] _main clause_ (GO_ESSAY_002 s42)

However, some types of coordinated sentences are still present in our data. The first group of these sentences is characterized by ellipsis, where subject and finite verb are realized only in one of the coordinated clauses. For instance, in (32) the second main clause contains only a direct object and a predicator, because it shares the subject and the finite verb with the first clause. As a result, automatic identification of coordinated clauses, based on the repetition of subjects and finite verbs within a sentence, missed these examples. However, these types of coordinated sentences have been excluded from the main data set containing the three selected constructions. The rest of such coordinated sentences, which are examples of other patterns, are kept in the data without being split into several independent clauses. The same applies to the group of sentences, which have been merged in translation: one of the original sentences extracted from the corpus often does not contain an aligned sentence, i.e. is annotated as a case of an empty link on the sentence level. However, a manual inspection of the sample exemplifying the selected constructions revealed that most of the source
sentences characterized by empty links could be in fact aligned to parts of the target sentences, which are often examples of coordinated clauses. Cqp queries were performed to verify these instances of empty links. As is shown in (33), in these cases two English original sentences should be aligned with coordinated sentence. Whereas the cases of empty links were typically included within the group without construction shifts (as there could not be any crossing lines), the cases from the narrower data sample, which correspond to coordinated sentences, were excluded from the data.

The next group of coordinated sentences left in the data includes cases where two elements, such as two finite verbs or two predicates, separated only by a conjunction, are annotated as one element, as in (34) (Hansen-Schirra, Neumann and Kunz 2010). Moreover, a number of coordinated clauses have been parsed into separate sentences within the CroCo corpus. Since these clauses have their own sentence IDs both in the original and the translation, the script could identify the crossing lines connected to the individual coordinated clauses. For instance, in example (35) the original sentence received two IDs. The first one has been assigned to the first main clause, and the second corresponds to the rest of the sentence. In this particular pair of sentences, the parsing on the clause level has been probably carried out to facilitate the alignment to the target text. The translation is an example of sentence splitting, where the first main clause is translated as an individual sentence.

The coordinated sentences belonging to the last two groups, exemplified through (34) and (35), are present in the final data samples. Therefore, the analyses reported in Chapter 5 account for a certain percentage of coordinated sentences.

(33) **EO:** The oxonium ion combines with three further water molecules to form H$\text{9O}_4^+$. This complex ion is held together by hydrogen bonds. (EO_POPSCI_009 s28-29)
    **GTrans:** Das Hydronium-Ion verbindet sich mit drei weiteren Wassermolekülen zu H$\text{9O}_4^+$, und dieses Komplexion wird durch Wasserstoffbrückenbindungen zusammengehalten.

(34) **EO:** One [strengthens]$^\text{lin}$ and [reinforces]$^\text{lin}$ the other. (EO_ESSAY_015 s18)
    **GTrans:** Das eine stärkt und verstärkt das andere.

(35) **EO:** The component molecules that make up different organisms (both at the individual and the species levels) are fundamentally alike: around 99 percent of the proteins in humans have recognizable equivalents in mice, and vice versa; many of those proteins are also conserved in other animals, and those involved in basic cellular processes are conserved in all eukaryotes. (EO_POPSCI_004 s109-110)
Die Moleküle, aus denen sich der Organismus aufbaut, sind selbst bei verschiedenen Spezies sehr ähnlich. Viele dieser Eiweißstoffe haben zudem bei anderen Tieren ihre Äquivalente, und solche für grundlegende Zellfunktionen sind sogar bei allen Eukaryoten konserviert.

The guidelines for annotation of grammatical functions developed for the CroCo corpus propose analyzing instances of direct speech as coordinated sentences. According to this suggestion, example (36) contains two main clauses, namely the clause within the quotation marks and the verb of speaking combined with the NP referring to the speaker, in this case, \textit{verspricht Bundeswirtschaftsminister Müller} "promises Federal Minister of Economy Müller". Hansen-Schirra, Neumann and Kunz (2010) suggest that an alternative analysis of treating the reported message as a direct object is problematic for the analysis of longer stretches of direct speech, possibly consisting of several orthographic sentences. However, the present study argues that splitting the reported message from the verb of speaking and the source of the proposition results in an incomplete clause: the verb introducing the reported message requires a complementation through a NP or a clause functioning as the direct object (Quirk et al. 1985, 1022). Therefore, cases like (36) are analyzed as instances of one main clause realizing the [Subject Verb Direct Object] construction (see Chapter 5.1.3).

(36) \textbf{GO:} ["Wir wollen das unternehmerische Potenzial in Deutschland weiter mobilisieren",]_dobj [verspricht]_fin [Bundeswirtschaftsminister Müller.]_subj

(37) \textbf{GO:} ["Kreativität ist nicht an Arbeitszeit gebunden", sagt Schech.,]_main clause ["und weil ich ortsungebunden bin, kann ich am Ort bleiben." ]_main clause

By contrast, direct speech is split into several main clauses, if all of the clauses can function on their own. In this case, the combination of the verb of speaking and the corresponding subject is linked to the first clause included within the direct speech. For instance, example (37) is annotated as a coordinated sentence involving two main clauses. This annotation decision allows us to deal with even more complex cases of direct speech, separating the reported message into several clauses or sentences. At the same time, this analysis does not systematically exclude examples of direct speech from the [Subject Verb Direct Object] pattern, while including parallel instances indirect speech, as would happen otherwise. Another case where I modified the annotation used in the CroCo corpus is related to the instances of \textit{Zustandpassiv} in German. The sentences realizing this type of passive were
originally annotated as involving the grammatical function of complement. For instance, the German translation in (38) would consist of the subject wir ‘we’, the copula verb sein ‘be’ (modified by the auxiliary werden ‘will’) and the complement mapped onto the verb phrase nicht gezwungen, unsere Währung zu verschrotten ‘not forced to scrap our currency’. Since English does not distinguish between Zustands- and Vorgangspassiv, both types of passive voice clauses in German correspond to the be-passive in English (König and Gast 2009, 131), where the verb, here forced, functions as the predicator, rather than the complement. Therefore, similar to the problem case of alignment between direct objects preceded by phrasal verbs and prepositional objects, the crossing line between the verb and the complement would be a result of different categories existing within the analyzed languages. To make the analysis comparable, the instances of Zustandspassiv are excluded from the [Subject Verb Complement] data set.

(38) **EO:** We will not be forced to scrap our currency. (EO_ESSAY_011 s34)
**GTrans:** Wir werden nicht gezwungen sein, unsere Währung zu verschrotten.

(39) **EO:** The work at these labs was devoted to cells, such as nerve cells, that are fully specialised for electrical signalling. (EO_POPSCI_011 s14)

In order to make sure that the English examples could in fact be classified as sentences in the passive voice, it was verified whether the word following the verb be is listed in a dictionary only as a verb or also as an adjective. If the latter is the case and the agent of the action is not introduced through the following by-phrase, the adjective is considered to realize the function of complement. For instance, in (39) the word devoted has been classified as an adjective, which, together with its postmodification, functions as the complement.

Finally, it should be mentioned that 48 sentences belonging to the text EO_ESSAY_029 and 62 sentences in the GO_ESSAY_022 have not been annotated for grammatical functions and are, therefore, not included into the present analysis. Moreover, because of parsing mistakes, two sentences within the sub-corpus EO_POPSCI, two belonging to GO_POPSCI and 25 in GO_ESSAY represent incomplete sentences and are not considered within the data samples.

**4.2.3 Selected constructions**

After the application of the methodological steps described in Chapters 4.2.1 and 4.2.2, the patterns, which had been classified on the basis of four initial grammatical functions, were quantified in English and German originals. The two constructions that are both frequent in
general, and often undergo construction shifts, are [Subject Verb Direct Object] and [Subject Verb Complement]. These have been selected for further analysis together with the third basic argument structure construction, namely [Subject Verb Prepositional Object]. Though the latter construction occurs less frequently in the analyzed originals, the contrastive difference between the functions of complements and prepositional objects (see Chapter 5 for further details) has motivated its inclusion in the present study.

Therefore, the study is limited to patterns consisting of three grammatical functions. In other words, sentences with additional elements, such as adverbials, are not included as instances of the abstract argument structure constructions under analysis. However, the question whether structures such as [Subject Verb Direct Object] and [Subject Verb Direct Object Local Adverbial] are two constructions or variants of the same construction should be studied further in future work. It is possible that at least certain combinations of objects or complements and adverbials could represent separate constructions. For instance, Goldberg (1995, 3) includes the caused motion construction with the form [Subject Verb Object Oblique] or the intransitive motion construction, realized by [Subject Verb Oblique], among the set of argument structures, where oblique could also be labelled a local adverbial referring to a direction of movement. Moreover, further analysis of these more complex structures could reveal lower-level partially lexically fixed constructions. However, the detailed investigation of these patterns goes beyond the scope of the present study.

The next chapter describes and motivates the types of analyses performed on the set of the selected constructions.

4.3 Types of analyses

The selected constructions were saved and further processed in a set of Excel tables. Instances of every grammatical pattern were grouped in eight tables, depending on whether it is an English or German construction, whether it occurs in originals or translations and whether it is affected by a construction shift or not. Therefore, the eight tables associated with every construction are EO shifts, EO no shifts, ETrans shifts, ETrans no shifts, GO shifts, GO no shifts, GTrans shifts and GTrans no shifts. Depending on the further classification of the construction, additional variables have been annotated within the corresponding tables. Since these classifications are motivated by specific characteristics of the individual
argument structure constructions and their contrastive differences, we discuss the details related to the classification in the relevant sub-chapters of Chapter 5.

Based on the classification, the next step is to identify certain frequent patterns in the data. Patterns are referred to as constructions only if the discussion indicates that the structures could be interpreted as constructions in their own right stored alongside the more general schemas (Goldberg 2006, 55), characterized by a combination of a specific form and function, with certain restrictions on some of the slots and possibly associated with particular registers. In other words, the existence of the lower-level constructions analyzed in the present study is assumed based on independent, in some cases register-specific functions identified for lexically restricted patterns. As discussed in Chapter 3.1.3, entrenchment has been shown to be a gradient phenomenon, making it implausible to postulate a frequency threshold for identification of structures as constructions. Therefore, while I rely on corpus-based frequencies to establish relative entrenchment of several structures (i.e. to determine what construction is more entrenched), the present study does not use frequency to verify whether the postulated constructions are in fact stored separately in the speakers’ minds. It can be argued that psycholinguistic experiments are required to provide empirical evidence for the existence of these constructions. The pilot translation experiment reported in Chapter 6 indicates one way of investigating whether the translators operate with lower-level constructions. However, due to the limited data sample, the potential of this experiment type for the investigation of entrenchment has to be tested further.

In the present study, once an assumption on a potential lower-level construction with specific formal and functional features is formulated, the characteristics of the form-meaning pairing are analyzed further by performing a collostructional analysis. This is a statistical method, which is used to account for association between a construction and lexical items that can occur in it (e.g. Stefanowitsch and Gries 2003, 2005; Gries and Stefanowitsch 2004a, 2004b). According to the authors, “the method provides an objective approach to identifying the meaning of a grammatical construction and of determining the degree to which particular slots in a grammatical structure prefer, or are restricted to, a particular set or semantic class of lexical items” (Stefanowitsch and Gries 2003, 211).

---

14 Whereas the notions of construction and form-meaning pairing are reserved to these structures, the other terms, such as pattern or structure are used more loosely to refer both to sequences of grammatical functions that have and have not been identified as constructions.
account not only frequency of occurrence of a particular lexeme in the analyzed construction, but also frequency of this lexeme in all other constructions, frequency of the construction with all other lexemes, as well as frequency of all other constructions with all other lexemes. These four values form a contingency table. The base-ten logarithm of the p-value, calculated through the Fisher’s exact test, indicates the strength with which this lexeme is attracted to the construction or repelled by it. The log-transformed p-value reaches the significance level of 0.05, if it is over 1.30103 (Stefanowitsch and Gries 2005, 7). The analysis has been performed using the statistical software R (R Core Team 2014) and the script Coll.analysis 3.2a (Gries 2007).

In addition, one of the extensions of collostructional analysis, namely covarying collexeme analysis (Stefanowitsch and Gries 2005), is also used in the present study. It allows for the establishment of what lexical items are likely to co-occur filling two different slots of a given construction. The analyzed contingency tables include 1) frequency with which the two items (L and M) co-occur in a given construction; 2) frequency with which the lexical item L filling the first slot is combined with items other than M in a given construction; 3) frequency with which the lexical item M filling the second slot is combined with items other than L in a given construction; and 4) frequency of a given construction with items other than L and M. Similar to the basic version of the collostructional analysis described above, the log-transformed p-values for every combination are calculated using the Fisher’s exact test performed with the help of the R script Coll.analysis 3.2a (Gries 2007).

A possibility to differentiate those co-occurrences of the construction and the lexeme that are statistically significant from the ones that occur due to chance is named among the main advantages of the method (Gries et al. 2005, 648). One of the major points of criticism stresses the fact that the analysis would be more valid if the semantic aspects rather than only frequencies were considered, especially since the aim of the approach is to account for constructional meaning (Bybee 2010, 98-100). For example, the experiment using acceptability judgments suggested that lexemes occurring in the analyzed construction with low frequency but semantically related to the ones identified as the central members of the construction are also considered relevant to the function of the construction (Bybee 2010, 99-100; Bybee and Eddington 2006, 351). This finding can be explained by reference to general cognitive mechanisms, such as categories for the individual construction slots are
based on high frequency items, and are extended by analogy to similar lexemes, even when these are novel in the environment of the specific construction (cf. Bybee 2010; Bybee and Eddington 2006). Even though the relative ranking of the association strength may, to some extent, depend on the frequency with which the lexemes occur in a given construction, the argument does not undermine the method’s potential to identify the central meaning of the construction, since the prototypical semantic categories are formed around the frequent lexemes. Moreover, the advocates of collostructional methodology argue that the divergences between the rankings based on the raw frequency data and the collostructional strength, particularly with regard to lexemes, which appear at the top of the corresponding lists, provide valuable information on the nuances of the constructional meaning (Gries, Hampe and Schönefeld 2005, 652-653).

The applications of collostructional analysis in addition to information on the use of the identified construction within the two registers and its comparison to the corresponding construction in the other language allow us to explore some of the features associated with each lower-level pattern. However, it is also necessary to consider larger data samples by consulting reference corpora. For this purpose the British National Corpus (BNC) (The British National Corpus 2007; Davies 2004-) and the corpus das Digitale Wörterbuch der deutschen Sprache (DWDS, more specifically, the core corpus of the 20th century) (BBAW 2010) have been selected. The corpora are comparable, as they both include non-translated texts from a similar range of registers written in the 20th century and consist of approximately 100 million word tokens (Burnard 2007, Geyken 2007). I decided against the use of the reference corpora included in the CroCo corpus because, even though these corpora contain the types of annotation necessary for extraction of the constructions, their size is insufficient for more detailed investigations of partially lexically filled constructions, which are in general less frequent than the abstract schemas. Since some of the patterns under analysis are associated with particular registers, the number of examples contained in the CroCo reference corpora is expected to be rather low, as each of the included 17 registers is represented through two thousand-word samples.

The different methods of sampling the data in the CroCo corpus, on the one hand, and the BNC and DWDS, on the other, are problematic with respect to the direct comparability of results. The analysis of the CroCo corpus takes into account nearly all examples of the
analyzed construction in the two registers (with a few exceptions mentioned above, such as coordinated sentences), because the data extraction is based on the available annotation of grammatical functions. Unfortunately, it was not possible to identify any comparable representative corpora of English and German annotated for grammatical functions that are larger than one million words. Therefore, the same data extraction procedure is not possible: the search within BNC and DWDS is limited to specific realizations of the analyzed construction (see Chapter 5 for details on individual search parameters). In most cases the results are limited to random samples of 200 examples. The data extracted from the reference corpora is used to further analyze the corresponding constructions and to confirm the cross-linguistic frequency distributions of the constructions observed in the smaller sample from the CroCo corpus. In some cases, when a pattern under investigation is associated with the register of political discourse, other corpora had to be considered, since the BNC and DWDS do not include this register (see e.g. Chapter 5.1.2).

Taking these limitations into account, the contrastive differences in the use of the corresponding constructions serve as the basis for the analysis of translation shifts from and into the analyzed constructions in both translation directions. Where applicable, the resulting differences in word order (see Chapter 5.1.3) or the degree of impersonality (see Chapter 5.1.1) are discussed. In the case of contrastive differences, it is also studied whether the frequency of the structures differ in the originals and translations belonging to the same language. This discussion is linked to translation properties.

Taking into account different sizes of the analyzed data samples, the results of both multilingual and monolingual comparisons are reported using percentages. Moreover, to establish whether the detected differences are statistically significant or are due to chance, statistical tests for independence of samples were performed using the software R (R Core Team 2014). The tests include the chi-square test for nominal response variables (see Chapter 5) and the t-test for interval response variables (see Chapter 6). Both tests have a number of assumptions that have to be verified prior to the application of the method.

The chi-square test presupposes a set of independent data points, in our case instances of a given argument structure construction. The application of the test allows testing whether there is a significant effect of an independent (predictor) variable on a dependent (response) variable. The test is suitable when both variables are nominal or categorical, i.e. contain two
or more discrete levels that represent labels for different categories and cannot be arranged on a scale, such as the levels of female and male for the variable “gender”. The values of these variables form 2x2 or, in case of categorical variables, larger contingency tables. The observed frequencies are then compared to the expected frequencies calculated for each of the cells. Before the calculation of the test statistic $\chi^2$, Gries (2009, 166) suggests verifying whether all expected frequencies are larger than one and whether 80% of the expected frequencies are larger than or equal to five. At the same time, he refers to previous research according to which “the chi-square test is fairly robust even if this [the second] assumption is violated” (Gries 2009, 152). While this assumption is not met in some of the analyses, it was made sure that the contingency tables do not contain zero values. To avoid these, certain minor categories are grouped together for the purposes of statistical analyses. If one of the analyzed variables has been annotated for more than two levels, the results of the chi-square test are visualized using the mosaic plot that helps establish what level(s) is (are) responsible for the significant result (Friendly 1994).

The t-test is used to establish the relation between an independent nominal variable and dependent interval variable, whose levels can be arranged on a scale with fixed intervals between them. For instance, time is a typical interval variable. Rather than working with contingency tables, this test compares the means of two groups. The assumptions of the test involve normal distribution of the data and the homogeneity of variances within samples (Gries 2009, 206). If these assumptions are not met, the log-transformed values are considered in the calculation of the test statistic $t$.

The study also involves two logistic multivariate regression analyses that help us study the effect of various factors on the translation shifts (see Chapters 5.1.6 and 5.2.3 for more methodological details on regression models). All statistical analyses are performed using the software R (R Core Team 2014). The obtained results are then interpreted in the context of the study.
5. Corpus-based analysis of construction shifts

The present chapter deals with construction shifts from a product-based perspective examining the phenomenon both through quantitative and qualitative corpus-based investigations. As mentioned earlier, the holistic analysis of construction shifts has identified three schematic constructions for further analysis. Compared to other structures, two of these constructions, namely [Subject V Direct Object] and [Subject Verb Complement] are often changed in translations and are at the same time frequent in the data. Moreover, to account for the third basic argument structure and analyze a contrastive difference between the grammatical functions of complements and prepositional objects reported in the previous literature, the construction [Subject Verb Prepositional Object] has been included in the study despite its low numbers in the corpus. Each of the three constructions consists of three elements and does not involve any adverbials or other grammatical functions considered in the initial classification of sentences (see Chapter 4).

The following sub-chapters discuss the findings for the three constructions for the translation directions English-German and German-English. A more detailed classification and selection of patterns for further analysis is motivated through contrastive differences between the two languages. It is shown that the frequencies of constructions, particularly on the lower levels of the constructional hierarchy, and construction shifts they are involved in depend on the language and the translation direction, but other factors such as register also play a role. The sub-chapters are structured in the following way. First, the relevant contrastive differences between English and German are identified leading in some cases to further annotation and grouping of the data. Based on the distribution of individual patterns, lower-level constructions are identified and characterized. It is investigated whether the category of construction shifts depends on various sub-patterns of the construction and what reasons could be identified for translating the original construction with a non-corresponding one. In the next step the effect of translation shifts is examined in more detail taking into account translation properties. The same procedure is repeated for the opposite translation direction.

The research questions for the corpus-based analyses are: 1) how instances of the schematic constructions found in English and German originals could be classified, 2) what the
characteristic features of the identified lower-level constructions are, and 3) what types of generalizations could be formulated on the translation of these sub-patterns. In particular, it is analyzed whether there is a correlation between the type of constructions found in the original, the presence of a construction shift and the effects of the construction shifts, such as information structure differences.

5.1 Subject Verb Direct Object

The first analyzed construction is [Subject Verb Direct Object]15, i.e. typically referred to as a transitive construction. However, it should be kept in mind that the present study concentrates only on the instances of the transitive construction that do not involve any additional elements, such as adverbials. Therefore, the category examined here is more limited than the one discussed in the literature.

On the construction's formal side, English and German differ with respect to the realization of grammatical functions: as discussed in Chapter 3.1.4, the subject in an English declarative clause is marked by its position before the finite verb, whereas the direct object occurs after the finite verb. In contrast, the same grammatical constituents in German are marked by the nominative and accusative cases, respectively. The general semantics of the [Subj V Dobj] construction has been described as one participant affecting another. In other words, the subject is typically mapped onto the semantic role of Agent, whereas the direct object is linked to the Patient (Goldberg 1995, 118 for English; Rostilla 2007, 268 for German, see the discussion below on a more detailed account of the transitivity prototype). Despite some cross-linguistic differences, the [Subj V Dobj] constructions in English and German can be considered corresponding, as they share the set of grammatical constituents and the functional characteristics.

<table>
<thead>
<tr>
<th></th>
<th>[Subj V Dobj] shifts</th>
<th>[Subj V Dobj] no shifts</th>
<th>All constructions shifts</th>
<th>All constructions no shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-G ESSAY</td>
<td>55</td>
<td>98</td>
<td>495</td>
<td>798</td>
</tr>
<tr>
<td>E-G POPSCI</td>
<td>60</td>
<td>62</td>
<td>488</td>
<td>560</td>
</tr>
<tr>
<td>G-E ESSAY</td>
<td>27</td>
<td>82</td>
<td>657</td>
<td>917</td>
</tr>
<tr>
<td>G-E POPSCI</td>
<td>16</td>
<td>56</td>
<td>442</td>
<td>725</td>
</tr>
</tbody>
</table>

Table 3: [Subj V Dobj] and other clause-level constructions with and without shifts

15 From now on, the construction is referred to as [Subj V Dobj].
Table 3 provides an overview of the frequency distributions related to the categories of construction shifts and no construction shifts for the two translation directions and the two registers under analysis. The first two columns show how often the construction [Subj V Dobj] is changed or kept in translations, whereas the third and fourth columns indicate the frequency distributions of all constructions in the two sub-corpora, both with and without shifts. It can be seen that the grammatical pattern [Subj V Dobj] shifts in 36% (55/153) of all occurrences of the construction in the English originals belonging to the register of political essays (ESSAY), accounting for 11% (55/495) of the construction shifts in this sub-corpus in the translation direction English-German (E-G). In the sub-corpus of the English originals extracted from the register of popular-scientific texts (POPSCI) 49% (60/122) of all instances of this pattern are affected by the construction shifts, corresponding to 12% of all shifts (60/488) in this direction. In the opposite translation direction, I have identified 25% (27/109) of instances that are translated by a different construction for the register ESSAY and 22% (16/72) for the register POPSCI.

A comparison between the two translation directions shows that there are less shifts from the given construction in the direction German-English (G-E). While for the popular-scientific register this difference is significant ($\chi^2$=12.7, df=1, p=0.0004, see Table 4\textsuperscript{16}), for the second register the p-value is somewhat above the significance level of 0.05 ($\chi^2$=3.2, df=1, p=0.07, see Table 5). The shifts from [Subj V Dobj], identified in the German originals belonging to the registers ESSAY and POPSCI, correspond to 4% (27/657 and 16/442, respectively) of all shifts for the translation direction German-English. The construction [Subj V Dobj] could be shown to play a larger role in the category of construction shifts for the direction English-German. This could be explained by the fact that the analyzed pattern has different frequency distributions in the originals of the two languages: in relation to the overall number of clause-level constructions belonging to the two registers, the analyzed construction occurs more often in English than in German ($\chi^2$=40.27, df=1, p=2.209e-10, see Table 6). Taking the different frequencies into account, the lower percentage of shifts for the direction German-English in the register POPSCI is not surprising: since the construction is more frequent in

\textsuperscript{16} For convenience, the values from Table 3, which form the basis of this and the following chi-square tests, have been grouped in separate contingency tables.
English, translations into this language do not have to adjust to the requirements of the TL and can keep the original construction.

The higher frequency of \([\text{Subj V Dobj}]\) in English than in German to some extent contradicts the findings of the study by Fischer (2013, 47), which has shown the opposite tendency. However, Fischer’s investigation is more explorative as the corpus consists of three texts (two literary and one political essay, classified as a journalistic text\(^{17}\)) and their respective translations (Fischer 2013, 26-27). As a result, the literary texts have a stronger effect on the general frequency distribution. The author notes that the proportions are different in the political essay, where there are slightly more \([\text{Subj V Dobj}]\) examples belonging to the English data. Therefore, it might be the case that the number of \([\text{Subj V Dobj}]\) sentences depends not only on the language, but also on the register (cf. Fischer 2013, 244, 249). However, it should also be taken into account that Fischer (2013, 23) does not distinguish between originals and translations in his analysis.

<table>
<thead>
<tr>
<th></th>
<th>Shifts</th>
<th>No shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>EO-GTRANS_POPSCI</td>
<td>60</td>
<td>62</td>
</tr>
<tr>
<td>GO-GTRANS_POPSCI</td>
<td>16</td>
<td>56</td>
</tr>
</tbody>
</table>

*Table 4: \([\text{Subj V Dobj}]\) - Contingency table for construction shift & translation direction, POPSCI*

<table>
<thead>
<tr>
<th></th>
<th>Shifts</th>
<th>No shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>EO-GTRANS_ESSAY</td>
<td>55</td>
<td>98</td>
</tr>
<tr>
<td>GO-ETRANS_ESSAY</td>
<td>27</td>
<td>82</td>
</tr>
</tbody>
</table>

*Table 5: \([\text{Subj V Dobj}]\) - Contingency table for construction shift & translation direction, ESSAY*

<table>
<thead>
<tr>
<th></th>
<th>([\text{Subj V Dobj}])</th>
<th>Other constructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EO</td>
<td>275</td>
<td>2066</td>
</tr>
<tr>
<td>GO</td>
<td>181</td>
<td>2560</td>
</tr>
</tbody>
</table>

*Table 6: \([\text{Subj V Dobj}]\) - Contingency table for construction & language*

Since the examples of \([\text{Subj V Dobj}]\) appear to be too heterogeneous to identify any regularities in the occurrence of construction shifts from this pattern, as a next step it is essential to look at the data in more detail. Therefore, based on the contrastive differences between English and German the schematic construction is further categorized to identify

\(^{17}\) The political essay and its translation analyzed by Fischer (2013) are also included into the CroCo register ESSAY.
possible reasons for translation shifts. As mentioned in Chapter 2.6, contrastive differences is one of the main factors influencing translations (Steiner 2001, 9) and is, therefore, a good starting point for formulating initial hypotheses. Since there are more translation shifts from the analyzed construction for the translation direction English-German, the analysis considers this direction first. Moreover, the contrastive difference used for classification of constructions is also expected to have a stronger effect on the translations from English to German.

Contrastive studies for the language pair English-German have noted that English has a wider possible range of semantic functions that can be mapped onto the subject than German. Using the terminology of Quirk et al. (1985) it can be stated that whereas both languages frequently express the semantic roles of Agentive Participant or External Causer through the grammatical function of subject, the mapping of such roles as Temporal, Locative or Instrument onto the subject is more typical of English than of German (Hawkins 1986, 57-61; König and Gast 2009, 108-109; see also Kast 2012, 149 for a contrastive overview of all semantic roles). In the present study the semantic role of Agent(ive) is understood as referring to a volitional, typically human being who brings about an action; External Causer is an inanimate cause of an action (e.g. a natural force); Instrument refers to a tool or method that is used to carry out an action; Temporal and Locative specify time and place of an action (Quirk et al. 1985, 741-747; Saeed 1997, 140-141).

This flexibility with regard to semantic roles in English is often considered to be interconnected with more restrictions on word order in this language. We will return to this discussion in Chapter 5.1.1.

(40) This advert will sell us a lot of dog food.
Mit dieser Werbung werden wir viel Hundefutter verkaufen. (König and Gast 2009, 108)

The difference in semantic roles realized by subject can be illustrated in (40): the subject in the English original is linked to the Instrument, namely the advert through which the action is accomplished. In the translation the same information is mapped onto an adverbial. The personal pronoun wir ‘we’ functions as the subject of this German sentence, i.e. the semantic content, which is realized by an indirect object in the original. Therefore, the example demonstrates that shifts in grammatical functions are often interconnected with a shift in
semantic perspective known as “modulation” (Vinay and Darbelnet 1995, 36). At the same time, it shows that a different mapping of semantic roles onto grammatical functions allows the translator to keep the order of words (and hence information structure) of the original sentence.

The English original sentence is characterized through a mismatch between the inanimate subject this advert and the verb sell, which expresses a volitional action performed by the subject. As described above, it is often claimed that this type of subject is more typical of English. Additionally, focusing on verbs Doherty (2002, 76-78) notes that German does not allow a combination of certain verbs with inanimate subjects. For instance, the verb zerreifen (‘tear apart’), when used in a transitive construction, occurs only with an animate subject.

The present study assumes that it is the volitional nature of the verb zerreifen that makes it incompatible with an inanimate subject in German. Taken these findings into account, the schematic construction [Subj V Dobj] is classified according to the animacy of the subject and the volition expressed by the verb. Based on this classification a systematic analysis of all data points with regard to the features of animacy and volition allows assessing the importance of this contrastive difference on the variable of construction shifts.

It has been suggested that in linguistic studies the binary biological distinction between animate and inanimate categories is insufficient (e.g. Ovrelid 2009, 630). Therefore, the classification of animacy introduced in Garretson (2004) and Zaenen et al. (2004), which differentiates between humans, animate and inanimate objects, is taken as the basis for the data annotation in the present work. The biological category of animate beings is further divided into humans and other animate objects, such as organizations, animals, as well as the two minor categories of machines and vehicles. Examples of the category “organization” are the NPs Microsoft or the company (Garretson 2004), i.e. individuals united by “collective voice” (Zaenen et al. 2004). The groups of intelligent machines, such as robots, and vehicles, e.g. cars, are also categorized as animate entities. The latter are included because they are “treated as living beings in some linguistic contexts” (Zaenen et al. 2004), for instance by referring to ships through the personal pronoun she. It should be noted that this argument might apply only to English that mainly operates with the categories of natural gender but not to German that distinguishes between natural and grammatical gender. However, this cross-linguistic difference in categorization does not affect the classification according to
animacy in our study, since the German data under analysis does not contain any instances of subjects realized by NPs referring to machines and vehicles. The third category of inanimate objects consists of concrete and abstract entities as well as the categories of time and location (Zaenen et al. 2004). It is important to mention that even though humans and animate objects are certainly concrete entities, they do not belong to the group labelled “concrete”, since the latter comprises only inanimate tangible entities, e.g. cup, door, tree (Garretson 2004). According to this annotation, the semantic roles of Locative, Temporal and Instrument that are realized by different grammatical functions in English and German all belong to the category of inanimate.

The detailed guidelines provided by Zaenen et al. (2004) are valuable for objective annotation of our data. However, there are still a number of problems arising during the assignment of the categories. First of all, the reference of linguistic expressions on which the annotation is based is not always clear. For instance, as mentioned by Zaenen et al. (2004), the names of countries could refer either to concrete places or metonymically to the governments of these countries: in the former case the NP is categorized as a location, whereas the latter exemplifies an organization. The linguistic context is used to assign the NPs to one of the animacy categories. Similarly, we can usually consult the context to determine the reference of pronouns. However, the reference of the personal pronoun we is more problematic in this respect, since its reference is typically dependent not on the linguistic context but rather on the register. For instance, in the register of popular scientific texts, similarly to academic writings, the pronoun is often used to potentially ascribe the statement to a larger research community: the specific researcher responsible for the findings moves into the background (Charles 2006, 507). In this case the subject is annotated as HUMAN. In contrast, the same pronoun often refers to the government or some other organizations in the register ESSAY and, therefore, the category of this pronoun in the register ESSAY is by default ORG (organization). In particular, in political discourse it is common for speakers to function as representatives of the political organizations to which they belong. Thus, in terms of the distinction between animator, author and principal (Goffman 1981, 144) the writer of a political essay is not the only principal, i.e. the one whose views are expressed: instead the group of principals often includes all members of the party or the government (Serbina 2010, 8). In academic or popular scientific writing the author of
the article is typically also the principal, even though the use of the pronoun we suggests a larger group of scientists sharing the expressed ideas (see also Chapter 5.1.2).

Another problem with the annotation for animacy concerns a group of subjects in German realized by the impersonal pronoun *man* and the dummy subject *es*, for instance in the partially filled construction *[Es gibt Dobj]*. It is interesting to note that examples of the construction *[Subj V Dobj]* in English originals do not contain impersonal pronouns, such as *one*, as subjects, and the English existential *there*-construction does not belong to the *[Subj V Dobj]* construction (see Chapter 5.1.4 on reasons for these classification differences). Since impersonal subjects do not carry any semantic content or have a very general reference, it is impossible to classify these according to animacy. The special category of *[Subj\_non-ref V Dobj]* is introduced to cover these examples, as shown in (41). Furthermore, subjects may refer to the whole process rather than an individual participant of an action. These subjects are expressed through a clause or a pronoun with an extended reference (Halliday and Hasan 1976, 66). In these cases the subject is also not categorized with regard to animacy and is annotated as *Subj\_process*. The group of the transitive constructions, which consists of sentences with this type of subjects, is illustrated through example (42).

(41) Subj\_non-ref V Dobj

  GO: Man kann sagen, daß in den großen Parteien sämtliche Extrempositionen in all ihren Widersprüchen vertreten sind. (GO_ESSAY_005 s30)

(42) Subj\_process V Dobj

  EO: Being able to choose has enormous important positive effects on us. (EO_POPSCI_005 s137)

In addition to annotation of subjects according to animacy, the main verbs occurring in the construction *[Subj V Dobj]* have been annotated with regard to volition. The distinction is made between Volition and Non-volition. Results of semantic verb annotation may differ depending on criteria playing a role in classification. Lenci (2014) suggests that verbs can be grouped in classes based on either extra-linguistic features or their linguistic distribution: description of verbs provided by *FrameNet* (Fillmore, Johnson and Petruck 2003) is characterized among the former types of categorization, whereas classification by Levin (1993) is attributed to the latter. At the same time it is noted that there is no strict dichotomy between the two classification types: the semantic and syntactic criteria tend to overlap to a certain extent (Levin 1993; Schulte im Walde 2006). For instance, Schulte im Walde (2006,
161) observes that "[e]ven though the classification is primarily based on semantic intuition and not on facts about syntactic behavior, the verbs grouped in one class share certain aspects of their behavior". The suggestion that it is the verb semantics that influences the occurrences of a particular verb in certain linguistic contexts (Levin 1993, 1) is in line with the tenet of CxG, according to which the meaning of a construction and the semantic properties of its collexemes tend to be semantically compatible (Goldberg 1995, 50; Stefanowitsch and Gries 2003, 213). Another related aspect that should be considered is whether the semantic classification is based on multiple verb senses or the most prototypical sense. In the first case the same verb might be assigned to different semantic groups depending on its sense. One of the arguments for this type of class assignment is also connected to the syntactic distribution: different senses occur in different constructions (Gilquin 2010, 197-198). Furthermore, Boas (2011, 50) states that multiple senses could activate distinct extra-linguistic frames: one and the same verb can be associated with different sets of participant roles that can fuse with argument roles of the argument-structure constructions. For example, according to FrameNet, the verb show is integrated into seven different frames, i.e. has seven senses. Looking at sentences from our data set in which the verb show occurs in the construction [Subj V Dobj] we could identify several of the suggested senses. Thus, (43) exemplifies the sense “cause to perceive”: the employment map of Europe would be annotated as the medium that causes the perceiver to observe the phenomenon expressed through the direct object. Example (44) can be interpreted as representing the sense “evidence”, in which the support, in our case several studies, are used as evidence for a proposition, whereas (45) is linked to the frame “reasoning”, where the arguers assert the content.

(43) **EO:** The employment map of Europe shows that inflexible social protection destroys rather than protects jobs. (EO_ESSAY_012 s27)

(44) **EO:** Several studies have shown that two of the factors affecting regret are how much one feels personal responsibility for the result and how easy it is to imagine a better alternative. (EO_POPSCI_005 s82)

(45) **EO:** Daniel T. Gilbert Of Harvard University and Timothy D. Wilson of the University of Virginia and their collaborators have shown that people consistently mispredict how long good experiences will make them feel good and how long bad experiences will make them feel bad. (EO_POPSCI_005 s95)
However, the application of this form of the multiple sense approach, which is based on the analysis of the semantic properties of the sentences into which each verb is integrated, is problematic for the present study. As indicated above, only the role of the arguer is typically filled by an animate entity, whereas the source in the frame “evidence” is realized by a certain type of inanimate referent. Since the present chapter aims at examining how verbs with the semantic feature of volition are combined with subjects characterized by different degrees of animacy, the semantic annotation cannot be based on their co-occurrence with arguments characterized by various semantic properties, one of which being animacy. In our case the semantic aspects of verbs rather than their distribution is in the focus of attention, as the occurrences of the [Subj V Dobj] construction in the CroCo data represent the starting point of the analysis. As postulated by CxG, the number of participant roles of the verb does not always have to correspond to the number of argument roles of the construction (Goldberg 1995, 52-56; 2006, 42; Michaelis 2004, 25). Furthermore, it is claimed that extra-linguistic knowledge activated by the verb allows the language user to integrate the verb into novel contexts (compare Chapter 3, example (10)), i.e. into constructions which may not be typically associated with particular verbs (cf. Goldberg 1995: 29).

At the same time the present study does not follow the prototypical sense approach, which has been attributed to the constructional accounts by Goldberg (Boas 2011, 43). The problem with this approach lies in identifying the most prototypical sense of the verb. Gilquin (2010, 166) suggests that the establishment of prototypical features is an empirical question, which, depending on one’s definition of prototypicality, could be answered through such methods as corpus-based investigations, psycholinguistic experiments or diachronic studies. The first two methods would require comprehensive analyses of every individual verb occurring in the data, and since the prototypicality of verb senses is not the central question of the present study, these methods are not considered. Taking a more diachronic perspective it has been suggested (Sweetser 1990; Meyer 1997) that concrete and abstract senses of polysemous verbs are interconnected through metaphor and metonymy. This classification, summarized in Table 7, takes animacy of subject and the level of abstractness of direct object into account. For instance, when considering agentive senses of the verb show, the prototypical subject is a human agent, while in metonymic senses the semantic role of the subject could be described as a pseudo-agent, since it is non-human and often abstract. In the metonymic cognitive sense
the instruments associated with the reported results (e.g. experiments) tend to be mapped onto the subject. The object is more likely to be concrete in the physical rather than cognitive senses, in which the action is a metaphorical one: the target domain of mental acts is structured as the source domain of, for instance, concrete acts of showing (cf. Meyer 1997, 100-110).

<table>
<thead>
<tr>
<th>Agentive</th>
<th>Physical</th>
<th>Cognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj Human Agent VerbAgentive ObjConcrete</td>
<td>He showed his identity card to the policeman.</td>
<td>Subj Human Agent VerbAgentive ObjAbstract</td>
</tr>
<tr>
<td>Metonymic</td>
<td>Subj Non-human Pseudo-Agent VerbAgentive ObjConcrete/Abstract</td>
<td>The map shows the harbor of Southampton.</td>
</tr>
</tbody>
</table>

Table 7: Senses of the agentive verbs (adapted from Meyer 1997, 100)

Meyer (1997, 100) suggests that cognitive senses tend to develop from physical ones, abstract from concrete and metonymic from agentive. In other words, according to the diachronic perspective, the prototypical human agent in physical agentive senses is typically transferred to other senses of agentive verbs: the physical agentive sense could be considered prototypical. While this approach can be operationalized within the scope of the present study, its application indicates semantic problems arising from a systematic assignment of the most concrete sense of the verb to all examples disregarding the context altogether.

The discussion has shown that the study requires a middle ground between context-dependent and verb-dependent semantic classification: the analysis should reflect the fact that meaning is created through combined contributions of lexical items and constructions (Goldberg 2011). In other words, the item-based lexical features and syntactic generalizations interact (see Chapter 3, Footnote 5): as mentioned earlier (see Chapter 3.1) utterances could be interpreted as combinations of atomic constructions such as individual verbs and syntactically complex argument structure constructions (Goldberg 1995, 16). This combination is another indicator of the syntax-lexicon continuum postulated by CxG.

To account for this interdependence and at the same time avoid the disadvantages of multiple-sense and prototype approaches, the version of the prototype approach described above could be modified using insights of Systemic Functional Linguistics (SFL). Assuming
that the logical subject, namely the doer of the action, does not always correspond to the grammatical subject (Halliday and Matthiessen 2013, 78-82; cf. Eisenberg 2004: 280), the sentences in our data sample are paraphrased to establish a potential logical subject of the sentence, which is in some cases left implicit. Halliday and Matthiessen discuss the distinction between different subject types using the example in (46): while my aunt is the grammatical subject, the duke is interpreted as the logical subject because he initiates the action\(^\text{18}\). The version in (47) is characterized not only by an unmarked information structure pattern, but also by a lack of reference to the logical subject. However, the duke or another doer of the action is implied.

(46) This teapot my aunt was given by the duke. (Halliday and Matthiessen 2013, 79)
(47) My aunt was given this teapot.
(48) a. This program arranges face-to-face professional, scientific, technical and community exchanges between Americans and their counterparts around the world. (EO_ESSAY_023 s31)
   b. Through this program people arrange face-to-face professional, scientific, technical and community exchanges between Americans and their counterparts around the world.

Similarly, in (48a) it could be said that the NP this program, functioning as the grammatical subject of the sentence, does not directly arrange any exchanges and hence is not the logical subject. It is argued that the people responsible for the program are implied as the initiators of the actions attributed to the program. Therefore, the sentence in (48a) could be paraphrased to arrive at its more explicit version in (48b). In this case it is made clearer that some group of people represents the Agent of the action, whereas this program realizes the semantic role of an Instrument. It is essential to state that the paraphrase does not represent a more basic syntactic structure from which others are derived: in other words, the approach should not be mistaken for the transformations in generative linguistics (cf. e.g. Chomsky 1975). The method of paraphrase, developed using certain concepts of SFL, is fully compatible with cognitive linguistics. Depending on various conceptualizations of an extra-linguistic event, the speaker may encode its participants in a number of ways. The subject and the direct object are considered to be focal participants: by mapping certain entities onto these grammatical functions the conceptualizer makes them more salient affecting the image

---

\(^{18}\) Halliday and Matthiessen (2013, 78-82) also discuss a category of psychological subject, realized through the NP my aunt in example (46), which is less relevant for the present analysis.
of an event transmitted through a linguistic clause. Whereas in an unmarked transitive clause subjects refer to agents and direct objects specify patients, the clauses may deviate from this prototypical constellation and profile other participants (Langacker 1991, 283-301). The paraphrases are formulated to uncover a more unmarked assignment of semantic roles to grammatical functions and, therefore, the participants in the so-called “canonical event model” (Langacker 1991, 285-286, see the discussion below on the transitivity prototype).

The current approach to the identification of the (implied) logical subject is closely interconnected with the SFL work on ergativity and grammatical metaphor. In the ergative model of transitivity it is assumed that the bomb in both (49a) and (49b) has the semantic role of Medium (or Affected, i.e. a participant, which participates in an action without causing it (Quirk et al. 1985: 741)), whereas the police, i.e. the participant which is left implicit in (49a), functions as Agent/Initiator (Halliday and Matthiessen 2013, 339). Since the non-ergative variant with the implicit agent as in (49a) corresponds to the intransitive sentence pattern [Subj V] (Halliday and Matthiessen 2013, 340), in contrast to the [Subj V Dobj] construction discussed here, the ergativity model is not directly applicable to the present study. However, it is possible to characterize the analysis of sentences along the lines of ergativity as a special case of the identification of the logical subject. For instance, in the non-ergative sentence (49a) the logical subject is left unspecified.

(49)  a. The bomb exploded.
       b. The police exploded the bomb. (Halliday and Matthiessen 2013, 339)

(50)  a. The creation of complex objects, whether houses or horses, demands two kinds of specifications: one for the components and one for the system that guides their assembly. (EO_POPSCI_004 s106)
       b. Builders demand two kinds of specifications, one for the components and one for the system that guides their assembly, to create complex objects, whether houses or horses.

Moreover, the phenomenon of grammatical metaphor (Halliday and Matthiessen 1999, 2013; Taverniers 2003) also plays a role in the identification of the agent who initiates the action. Whereas the congruent or non-metaphorical variant is assumed to represent the default realization of semantic information through certain grammatical units, the process of metaphorization refers to deviations from these unmarked lexico-grammatical links. The notion of non-congruency (metaphoricity) implies the re-alignment between certain
semantic and lexico-grammatical elements. For instance, a semantic figure, e.g. an event of doing or saying, is typically expressed through a grammatical clause, but in a metaphorical variant this clause could be reduced to a noun phrase. In fact Halliday and Matthiessen (2013, 730) suggest that nominalization is the most prominent mechanism of grammatical metaphor. The method of paraphrase is often applied to identify different degrees of congruency. In example (50a) the NP functioning as the subject corresponds to a clause in the more congruent version of this sentence in (50b): the metaphorical variant is formed by nominalizing the verb create and turning the direct object of the verb into the postmodifying prepositional phrase of the head noun creation. After the clause has been restructured as a NP it can function as the subject of the sentence moving the Agent, namely the builders, into the background. In other words, the sentence, which occurs in the register of popular scientific texts (shown in (50a)), does not explicitly refer to those responsible for the creation of complex objects. But again it could be argued that the creation of complex objects, whether houses or horses is not the logical subject of the sentence.

(51) a. The Rio decade has elevated the world’s understanding that development must be sustainable, that the three "pillars" of sustainable development – environmental protection, economic development, and social development – must go hand-in-hand. (EO_ESSAY_006 s13)
b. During the Rio decade the events have elevated the world’s understanding that development must be sustainable, that the three "pillars" of sustainable development – environmental protection, economic development, and social development – must go hand-in-hand.
c. During the Rio decade people in the world understood that development must be sustainable, that the three "pillars" of sustainable development – environmental protection, economic development, and social development – must go hand-in-hand.

(51a) presents an example from the register of political essays. Here a NP referring to a time period is mapped onto the subject. The example could be reformulated by turning the nominalized structure the world’s understanding into its clausal counterpart. The process of demetaphorization results in a more congruent variant in (51c), where the human referent realized by the noun phrase people in the world functions as the subject and the time reference is realized by an adverbial of time. However, this paraphrase does not contain the verb elevate, which is the main verb of the original sentence. Since the aim of the present analysis is not to identify the most congruent version of each sentence but rather the
volitional status of the verb, another paraphrase using the same verb is required. As shown in (51b) this variant also involves turning the subject into an adverbial of time, which could be considered as a more congruent function for time references. Comparing (51a) and (51b) we could observe that the time reference in the more congruent version in (51b) is realized by a prepositional phrase, which is re-aligned to a nominal group in (51a). As a result, the nominal group could be integrated into a material process type referring to a figure of doing where the element typically functioning as the circumstance is re-interpreted as the actor of the action. It is possible to position the three versions of this sentence on a scale of congruency: the variant in (51b) is more congruent than the corpus example in (51a) but more metaphorical than sentence (51c), which identifies people as the logical subject of the figure of sensing.

Analyzing a similar example Halliday and Matthiessen (2013, 716-717) state that the motivation behind this grammatical metaphor is a textual one: by condensing a clause into a NP it is possible to present the whole clause as one textual unit, which could then function, for instance, as new information. This is true for both (51a) and (51b), since the latter variant contains a prepositional phrase rather than a full clause. It could be assumed that (51a) is a more elegant and more condensed version of (51b): the Rio decade refers metonymically to the events that happened during this time.

So how can this analysis help us classify the verbs according to volition? Volition is defined here as “the capability of conscious choice and decision and intention” (Princeton University 2010). Thus, volitional actions are understood as involving human logical subject. As shown above, the method of paraphrase is used to decide whether the same verb used in the same voice19 could be used with a human referent without changing the meaning of the sentence. Another criterion for a verb to be considered volitional is its reference to an action, causation or a verbal act rather than a mental activity, an existence or a relationship. This classification

---

19 In example (51) it would be possible to paraphrase the sentence to contain the human subject and the verb elevate but the verb would have to be used in the passive instead of the active voice. However, as shown in (7) the subject of the clause in the passive voice is not considered to be the logical but only the grammatical subject.
is based on semantic groups of verbs identified by Biber et al. (1999) for English. To be more precise, mental verbs appear to form a heterogeneous group involving both volitional and non-volitional verbs (Biber et al. 1999, 362). While Halliday and Matthiessen (2013, 249-250) stress that the figure of sensing involves a conscious being, the definition of volition provided above also includes the aspect of intention, which is often missing in the mental verbs of perception, feeling and, to some extent, also cognition. Rather than deciding for every specific mental verb whether it could be classified as volitional or not, I classify the whole category of mental verbs as non-volitional. Since the identified categories are rather general, it could be argued that German verb groups as characterized by Schulte vom Walde (2006) could be integrated into this classification. The assignment of verbs to these groups has been performed after the potential logical subject has been identified. The lexical resources WordNet (Princeton University 2010) for the English language and its German counterpart GermaNet (Kunze and Lemnitzer 2007) have been consulted to ensure that the analyses are based on language-specific categorizations of verbs (see Chapter 7 for an evaluation of this methodological decision).

Therefore, the verbs, which 1) are combined with (implicit) human logical subjects, and 2) refer to volitional actions are categorized as volitional ones. Applying these criteria the verb arrange in (48) is volitional because its paraphrase with the subject people in (48b) describes a volitional action. Similarly the verb demand in version (50b) refers to a verbal action performed by human referents. In contrast, the verb elevate is kept only in (51b), a sentence without a human logical subject. Therefore, the verb in this particular sentence is considered to be non-volitional. Example (52) illustrates a case where the subject refers to a human being but the described action is a mental one resulting in the non-volitional verb believe.

---

20 The present chapter draws upon several theoretical concepts from SFL. However, I decided to adopt a theory-neutral classification of verbs into semantic categories rather than to analyze sentences according to the process types identified in SFL, because within this theoretical framework, in contrast to the present study, the assignment of semantic (participant) roles is performed not on the congruent versions of the sentences. For instance, Halliday and Matthiessen (2013, 250) argue that the empty house in the clause the empty house was longing for the children to return functions as the Senser. However, according to the classification framework used in the present study, the empty house becomes the Senser only through the process of metaphorization and, therefore, does not function as the logical subject of the clause. The more congruent version of this sentence could be people living in the house were longing for the children to return. This paraphrase helps to determine the logical subject, namely people, which functions as the Senser within the congruent assignment of the participant roles (see also a survey by O’Donnell, Zappavigna and Whitelaw (2008) that shows how the same sentences allow for different interpretations leading to different annotations of process types). It should be, however, mentioned that even though I do not explicitly refer to SFL process types, they are comparable to the categories identified by Biber et al. (1999).

21 The GermaNet was queried using GermaNet Explorer 7.3. developed by Marc Finthammer and adapted to the JAVA API 2.0 by Verena Henrich, available online at www.sfs.uni-tuebingen.de/GermanNet/tools.shtml
whereas in (53) both criteria for a volitional verb are not fulfilled. In cases where the first criterion is not met, the verb cannot be classified as referring to a volitional action. Therefore, there are no examples, which conform only to the second requirement.

(52) I believe that the experience and determination that the new member states bring to the EU will give new energy and boost the pace of modernisation, to the benefit of us all. (EO_ESSAY_004 s37)
(53) Lower prices mean a paycheck goes further at the supermarket and the department store. (EO_ESSAY_015 s22)

The present analysis thus recognizes the fact that different uses of the same verb are interconnected through metonymy and (both lexical and grammatical) metaphor and is, therefore, close to the prototype approach discussed earlier. But at the same time, the context does play a role: the classification of a verb starts from its use in a certain sentence and establishes whether a paraphrase could help identify a logical human subject that functions with the same verb without altering the general meaning of the sentence. Therefore, in the analysis I combine certain SFL concepts, namely the notion of logical subject and grammatical metaphor, and research on the verb’s semantics, which according to the CxG approach merges with the functional properties of the argument structure constructions. This integration of SFL and CxG tenets is beneficial for the analysis of semantic features. It takes into account the interaction between verbs and constructions as well as the role of prototype for the conceptualization of meaning. Moreover, the integrated approach indicates that the less prototypical meanings are interpreted against the more congruent, i.e. less marked, mappings between lexico-grammatical and semantic features, which could help identifying the prototypical semantic content of lexical items as well as larger patterns.

It is essential to consider sentences in which the verb in the main clause modifies the action reported in the following non-finite clause. For instance, Biber et al. (1999, 364) identify aspecual verbs as a separate semantic group: this category includes verbs that refer to the stage of development of the action presented in the non-finite clause, such as the verb begin in (54), which specifies that the process of investigation has been initiated recently. This group can be extended to include other types of modification: e.g. in (55) the verb seek does not refer to the action itself but rather to an attempt to perform a certain action, namely to negotiate regional and bilateral agreements to open markets around the world. Halliday and Matthiessen (2013, 584-585) suggest that phrases of the type are beginning to investigate
and *seek to negotiate* imply the actions of investigating and negotiating, whereas similar structures, such as *wants to balance* in (56), are different in this respect. The non-finite clause in (56) is a proposition that is a result of a mental activity of wanting rather than an action modified in some way by the preceding verb. Moreover, the patterns differ with respect to the time references. Thus, both parts of the verb phrase with the verb *seek*, namely *will seek* and *to do* have future reference: one cannot say *yesterday we sought to do it the following day* as seeking cannot be separated from the action it is modifying. In contrast, it is possible to formulate a sentence *yesterday we wanted to do it the following day*: while the mental activity took place in the past, its result may refer to the future. Halliday and Matthiessen state that the dividing line between the two types is not clear and one could treat some cases of the second category as belonging to the first one. These considerations should be kept in mind while classifying verbs occurring in such sentences according to volition.

(54) **EO**: But scientists are just beginning to investigate whether natural regeneration is somehow blocked in tissues that do not repair themselves easily and, if so, whether unblocking their regenerative capacity will be possible. (EO_POPSCI_006 s101)

(55) **EO**: We will seek to negotiate regional and bilateral agreements to open markets around the world. (EO_ESSAY_015 s71)

(56) **EO**: President Clinton wants to balance law enforcement, national security, privacy and commercial interests. (EO_ESSAY_021 s47)

However, the question arises whether in examples (54) – (56) the non-finite clause functions as an independent clause constituent with its own grammatical function, namely the direct object. Halliday and Matthiessen (2013, 567, footnote 1, 585) show that the whole verb phrase, e.g. *will seek to negotiate*, belongs together. This could be established by applying the question test: Halliday and Matthiessen argue that the appropriate question for the element following the verb would be *what will we seek to negotiate?* rather than *what will we seek?* The same could be stated about the verb phrase *want to balance*, even though it differs in other respects. At the same time all of the non-finite phrases following the main verb could be argued to function as one element. In some grammatical descriptions this syntactic unit is referred to as a direct object (Biber et al. 1999, 199-200) or a catenative complement (Huddleston and Pullum 2002, 65-66), whereas others distinguish between catenative and transitive instances of this pattern (Quirk et al. 1985, 146-147, 1061-1067, 1185-1195). According to the latter account, examples (54) – (56) have the structure [Subj V Dobj]
because the non-finite clause can be substituted by the pronoun *it*, a nominal version of the clause or function as the focus in a pseudo-cleft, as is shown in (57). In contrast, the verb *tend* behaves differently: example (58) illustrates that the sentence becomes ungrammatical, if the non-finite form is exchanged by a noun phrase. This sentence would, therefore, be analyzed as a catenative construction in this grammatical description.

(57)  
a. President Clinton wants to balance it.  
b. President Clinton wants a balance of law enforcement, national security, privacy and commercial interests.  
c. What President Clinton wants is to balance law enforcement, national security, privacy and commercial interests.

(58)  
a. Developmentalists have tended to visualise the physical forces that underpin that polarity as a gradient of some critical entity across the cell. (EO_POPSCI_011 s24)  
b.*Developmentalists have tended a visualization of the physical forces that underpin that polarity as a gradient of some critical entity across the cell.

Taking into account that there is no consensus on how to classify the structure consisting of a finite verb followed by a non-finite clause and whether they can be regarded as instances of the [Subj V Dobj] construction, the sentences with this pattern are grouped together and discussed separately from the rest of the data.

The discussion has so far concentrated on English patterns. However, also German grammars distinguish between two types of non-finite verb forms: those that are independent of the main verb of the sentence, thus forming a separate non-finite clause, and the ones that are considered to be parts of a complex VP (Eisenberg 2004, 349-365; Duden 2009, 846-852). In example (59a) the finite verb *genießt* ‘enjoys’ is followed by a non-finite infinitive clause: in this case it is the finite verb that is classified according to volition. One of the characteristics of this type of infinitives is the pronoun *es* ‘it’ used in the main clause, which links it to the following infinitive clause (Duden 2009, 850) integrating the whole non-finite clause into the constituent structure of the main clause: in this and similar cases the infinitive clause functions as the direct object (Eisenberg 2004, 353). It has been observed that verbs introducing non-finite clauses could remain in the position before the infinitive clause (see (59b)) (Eisenberg 2004, 363, Duden 2009, 851-852). In contrast, in (60) the verbs *lässt wachsen* ‘lets grow’ could be interpreted as belonging to one VP with *wachsen* ‘grow’ as its main verb (Duden 2009, 848). To ensure the comparability of analyses in both translation
directions, both types of sentences discussed here are also examined separately from the rest
of the data, as the category [Subj V Clause\textsubscript{non-fin}]. This group excludes those non-finite clauses,
which are integrated into complex NPs, and/or are subjects of the main clause.

(59)  
\begin{enumerate}
  \item a. Der Familienvater genießt es, jeden Tag mit seiner Frau und den beiden Kindern Lara (6) und Kolja (1) zusammenzusein. (GO\_ESSAY\_023 s47)
  \item b. …dass der Familienvater es genießt, jeden Tag mit seiner Frau und den beiden Kindern Lara (6) und Kolja (1) zusammenzusein.
\end{enumerate}

(60)  
Die Unfähigkeit des Partners lässt meine eigenen Fähigkeiten, mich zu spüren und zu verstehen, wachsen. (GO\_POPSCI\_004 s143)

Applying the described classifications of nouns and verbs in terms of animacy and volition,
the following senses of the transitive construction are differentiated in the data (see
examples (61) – (66)).

(61)  
\begin{enumerate}
  \item a. EO: [\textbf{We}Subj\_human [confirmed]V\_volition [that they had identified 85 to 90 percent of the wells correctly.]Dobj (EO\_POPSCI\_002 s84)
  \item b. GO: [\textbf{Die Empfängerin, Frau Schwarz, die} coarse\_human [die Truhe]Subj\_human kann [die]Truhe Dobj [öffnen.]V\_volition
      lock owns, Subj\_human can [the chest Dobj [open]V\_volition (GO\_POPSCI\_008 s25)
\end{enumerate}

(62)  
\begin{enumerate}
  \item a. EO: [\textbf{I}Subj\_human [believe]V\_non-volition [that the experience and determination that the new member states bring to the EU will give new energy and boost the pace of modernisation, to the benefit of us all.]Dobj (EO\_ESSAY\_004 s37)
  \item b. GO: [\textbf{FREUD (1917a)}Subj\_human [nahm an]V\_non-volition [daß das Ich wenig Einfluß auf das Handeln des Menschen hat.]Dobj (GO\_POPSCI\_007 s146)
\end{enumerate}

(63)  
\begin{enumerate}
  \item a. EO: [\textbf{We}Subj\_animate must [forge]V\_volition [partnerships with other governments, with businesses, and with civil society groups that ensure successful on-the-ground implementation.]Dobj (EO\_ESSAY\_006 s10)
  \item b. GO: [\textbf{Die Krankenkasse}Subj\_animate [bezahlt]V\_volition [die Kosten ärztlicher und zahnärztlicher Behandlung, Arzneien, Krankenhausbehandlung und Vorsorge.]Dobj (GO\_ESSAY\_001 s107)
\end{enumerate}
These six types of the [Subj V Dobj] construction do not include examples that belong to the groups [Subj process V Dobj] and [Subj non-ref V Dobj], the latter relevant for the German data (see discussion above and examples (41) and (42)) as well as the pattern with the form [Subj V Clause non-fin] (examples (54) – (56)), which are analyzed separately.

Before we discuss the distribution of the six senses of the transitive constructions in English originals, let us have a look at the prototype of the transitive construction suggested in the literature. Hopper and Thompson (1980) suggest that transitivity can be described along a number of parameters, summarized in Table 8. The authors argue that each of these features contributes to the overall meaning realized by a transitive construction. The prototypical semantics of this schematic construction has been described as “a matter of carrying-over or
transferring an action from one participant to another” (Hopper and Thompson 1980, 253) or as “a volitional actor affecting an inanimate patient – a causative event” (Goldberg 1995, 118). Thus, as can be seen in Table 8 volitional actions involving at least two participants have a high transitivity value and, therefore, are located closer to the transitivity prototype on the cline of transitivity. Moreover, some further aspects realized by subject, verb and direct object, such as agency, mode or affectedness, strengthen or weaken the transitive meaning of sentences. The present study considers only the parameters of agency, operationalized here through animacy, and volitionality, but future work could also take into account further factors from this list.

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Participants</td>
<td>2 or more participants</td>
<td>1 participant</td>
</tr>
<tr>
<td>B. Kinesis</td>
<td>action</td>
<td>non-action/state</td>
</tr>
<tr>
<td>C. Aspect</td>
<td>telic</td>
<td>atelic</td>
</tr>
<tr>
<td>D. Punctuality</td>
<td>punctual</td>
<td>non-punctual</td>
</tr>
<tr>
<td>E. Volitionality</td>
<td>volitional</td>
<td>non-volitional</td>
</tr>
<tr>
<td>F. Affirmation</td>
<td>affirmative</td>
<td>negative</td>
</tr>
<tr>
<td>G. Mode</td>
<td>realis</td>
<td>irrealis</td>
</tr>
<tr>
<td>H. Agency</td>
<td>Agent high in potency</td>
<td>Agent low in potency</td>
</tr>
<tr>
<td>I. Affectedness of Object</td>
<td>Object totally affected</td>
<td>Object non-affected</td>
</tr>
<tr>
<td>J. Individuation of Object</td>
<td>Object highly individuated (proper, human/animate, concrete, singular, count, referential/definite)</td>
<td>Object non-individuated (common, inanimate, abstract, plural, mass, non-referential)</td>
</tr>
</tbody>
</table>

Table 8: Parameters of transitivity (adapted from Hopper and Thompson 1980, 252-253)

Langacker (1991, 298-301) defines the transitivity prototype against an extra-linguistic canonical event model, which is characterized in the following way.

The stage model contributes the notion of an event occurring within a setting and a viewer [...] observing it from an external vantage point. Inherited from the billiard-ball model is the minimal conception of an action chain, in which one discrete object transmits energy to another through forceful physical contact. Moreover, the action-chain head is characterized as an agent, and its tail as a patient that undergoes a resultant change of state [...]. In sum, the canonical event model represents the normal observation of a prototypical action (Langacker 1991, 286, emphasis removed).

Langacker argues that the conceptualizer can realize an event through different grammatical structures profiling certain aspects of the event. As mentioned earlier, the functions of subject and direct object are assumed to represent focal participants, which highlight the entities mapped onto them. In the prototypical transitive clause these functions correspond to the
ones described in the canonical event model: the agent initiating an action is realized by the subject, whereas the patient is mapped onto the direct object (Dowty 1991, 576; Langacker 1991, 294, 301; Goldberg 1995, 118). According to this view, the conceptualizations of the same scene differ depending on the grammatical relations the participants are expressed through and whether they are mentioned explicitly in the clause. As discussed above, the subject of a transitive clause, particularly in English, could realize other semantic functions, such as an instrument. In this case it is the instrument rather than the agent using it that is presented as affecting another participant. The agent may be missing completely in this linguistic realization of a scene. Moreover, an intransitive (or a non-ergative) clause can be used instead to refer to the same event resulting again in a different event structure (Langacker 1991, 297; Höche 2009, 145). However, Langacker (1991, 301-302) also stresses that the number of participants is only one of the factors contributing to the degree of transitivity.

As is shown in the description of the canonical event model by Langacker (1991, 286) and the definition of the prototypical transitive sense by Goldberg (1995, 118), transitivity is often linked to causality: thus, in the billiard ball model the agent causes a certain action that affects the patient. In the prototypical concrete situation the causation results from physical contact (Langacker 1991, 286). Also the model of direct manipulation breaks the causation prototype into a number of features, which are similar to both the transitivity prototype suggested by Hopper and Thompson (1980) and the billiard ball model referred to by Langacker (1991). The causative situation is presented as an event during which a single definite human agent performs an action resulting in a change of a state of a single definite patient. Like in the prototypical transitive sentence the agent is also described as a volitional one. In addition, Lakoff (1987, 54-55) specifies that s/he uses hands, body or an instrument and is aware of the effect of the action. While causative situations can also be realized by other linguistic patterns, such as the periphrastic causative construction (Gilquin 2010), the prototype is typically illustrated through the sentences of the type [Subj V Dobj], i.e. a transitive construction.

It should be mentioned that the described transitivity prototypes by Goldberg, Hopper and Thompson and Langacker refer to English. Rostilla (2007, Chapter 9) describes the German transitive construction in similar terms as containing two arguments: the first of these is
realized by the nominative case mapped onto the Agent, whereas the second is expressed by the accusative case and has the semantic role of the Patient. Therefore, the general semantic properties of the English and German transitive constructions seem to be comparable.

While, as mentioned above, not all of the features discussed in the context of prototypical transitivity and/or causality are considered, the first identified sub-sense of the transitive construction, namely the pattern [Subj human V volition Dobj] illustrated through (61) (repeated as (67) for convenience) meets the criteria of a human agent and volitionality and, therefore, could be considered the most prototypical. In particular, example (67b) is a prototypical transitive sentence: the human agent, Frau Schwarz 'Ms. Schwarz', performs an action, which affects the patient, in this case the inanimate object die Truhe 'the chest'. However, taking into account the registers under analysis, this prototype is less likely to be frequent in our English and German data. While the prototype characterizes very concrete actions, at least the register of (popular)-scientific texts has been shown to rely on the use of grammatical metaphor and inanimate subjects (Halliday and Martin 1993; Hansen-Schirra 2012). Even example (67a), which has been also identified as being close to the prototype, is not very high on the scale of transitivity as it does not refer to a physical action of a single agent affecting a single concrete patient. Instead the sentence denotes a verbal act with a finite that-clause functioning as the direct object, which describes what has been confirmed. This example could be considered causative only in a metaphorical sense: through her/his action the agent ‘causes’ a scientific community to accept the results as being true (Serbina 2013, 180).

(67)  

a. EO:  We confirmed that they had identified 85 to 90 percent of the wells correctly. (EO_POPSCI_002 s84)


(67) (GO_POPSCI_008 s25)

The suggested categories of animacy, volition as well as the additional tags Subj_process, Subj_non-ref and Clause_non-fin have been assigned to the total of 275 examples of the schematic construction [Subj V Dobj] extracted from the sub-corpora of English originals and 181 examples belonging to the German originals from the registers of popular-scientific texts and essays. The general research question is whether certain identified sub-senses of the
transitive construction in the originals are more or less often affected by construction shifts (Serbina 2013\textsuperscript{22}). In particular, taking into account the contrastive difference discussed above according to which the combination of an inanimate subject and a volitional verb is more typical of English than of German, it is expected that the sub-sense [Subj\text{inanimate} V\text{volition} Dobj] occurring in the English originals will be more frequently changed than translated with a comparable construction in the German translations. Before testing this hypothesis, our study first verifies whether similar frequency distributions of this sense can be detected in the English and German originals belonging to the CroCo corpus. At the same time, this comparison allows us to identify further differences in the occurrences of all eight senses.

Figure 6 shows the frequencies of each sense relative to the number of the transitive constructions in the respective data samples. The chi-square test reveals that the general effect of the variable of language on the variable of transitive sub-senses is significant ($\chi^2=35.29$, df=6, p-value=3.781e-06)\textsuperscript{23}. In Figure 6 we can see that the sub-sense [Subj\text{inanimate} V\text{volition} Dobj] occurs twice as often in English than in German originals. As visualized through a mosaic plot in Figure 7, the observed values for this pattern are significantly higher than the expected ones in English and significantly lower in German. The light blue and light red colors indicate that the differences are significant at approximately the 0.05 level. This distribution is in line with the contrastive difference outlined above. Moreover, if we compare the number of inanimate subjects in general in the two languages, they are more common in English than in German, amounting to 47\% (117/250) and 40\% (68/172) respectively.

\textsuperscript{22}Serbina (2013) presents an earlier version of this analysis. The differences in numbers are due to the refined classification according to volition and further manual correction of the data annotation and alignment.

\textsuperscript{23}To avoid zero values, the categories [Subj\text{process} V Dobj] and [Subj\text{non-ref} V Dobj] have been grouped together for this calculation.
Taking into account that the sense [Subj\(\text{inanimate}\) Volition Dobj] is indeed associated with the English data, the original assumption that this pattern might play a special role in translations involving the language pair English-German deserves further attention. Therefore, in the next step the pattern is examined with respect to translation shifts in the translation direction English-German. To analyze this question, the data with and without construction shifts has been correlated with the sub-senses in English originals. Figure 8 shows the distribution of all patterns identified in the data.
We can observe that almost all constructions are more frequently kept in translations than changed to other structures. The only exception is the transitive sense [Subj_animate V_non-volition Dobj], which has almost the same number of instances that are affected by construction shifts and those that are not. This comparison shows that the effect of the identified seven\textsuperscript{24} sub-senses of the transitive construction on the variable of construction shift is not significant ($\chi^2=7.34$, df=6, p-value=0.29). Against our expectations, the particular sense [Subj_animate V_volition Dobj] is translated by the comparable pattern more often than it undergoes construction shifts. However, apart from being significantly more attracted to English than German, it is also the most frequent sense in the English originals and has the second highest percentage of shifts among the seven sub-senses in the translations from English. Due to these reasons, the pattern is investigated in more detail in Chapter 5.1.1.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8}
\caption{Construction shifts from the transitive patterns in the translation direction E-G}
\end{figure}

Returning to Figure 7, it is important to acknowledge contrastive differences with respect to two other senses, namely [Subj_animate V_non-volition Dobj] and [Subj_human V_volition Dobj]. While the former is significantly associated with English, the latter occurs more often in the German originals. Figure 8 shows that the transitive sense [Subj_animate V_non-volition Dobj] has the third

\textsuperscript{24}As discussed earlier, the pattern [Subj V Clause\textsubscript{non-fin}] is rather heterogenic and will, therefore, be discussed separately. At this point the analysis considers only the six combinations of animacy and volition, as well as the sense [Subj\textsubscript{process} V Dobj] (the sense [Subj\textsubscript{human-ref} V Dobj] does not occur in the English originals). Therefore, if not indicated otherwise, the statistical tests for the translation direction English-German reported in this chapter are based on 250 rather than the total of 275 observations. Also the tests for the opposite translation direction exclude the indicated pattern from the calculations.
highest proportion of shifts, following [Subj\text{inanimate} V\text{volition} Dobj]. Therefore, even though the pattern is more frequently kept in translations, it has a rather high proportion of shifts among the seven main senses identified in the present study. Similarly, the contrastive difference with respect to the pattern [Subj\text{human} V\text{volition} Dobj] is reflected in the number of shifts for the translation direction English-German. Because the pattern is more typical of the target language, this sense is characterized by the lowest proportion of shifts from English to German. It should also be mentioned that, according to Figure 8, the pattern [Subj\text{human} V\text{volition} Dobj] has a rather low number of occurrences and, therefore, as assumed, cannot be considered prototypical for the registers under analysis. Due to the contrastive differences associated with them, the two patterns are examined further in Chapters 5.1.2 and 5.1.3, respectively.

Moreover, since the sense [Subj\text{non-ref} V Dobj] occurs only in German originals, it also deserves further discussion (see Chapter 5.1.4). Chapter 5.1.5 gives a brief overview of the translations of sentences belonging to the category [Subj V Clause\text{non-fin}]. Finally, Chapter 5.1.6 concludes the discussion of shifts from the construction [Subj V Dobj] on a more general note taking into account other factors that could lead to the change of this abstract argument structure construction in translations.

### 5.1.1 Subj\text{inanimate} V\text{volition} Dobj

As shown in Figure 8 there are 32 instances of the sense [Subj\text{inanimate} V\text{volition} Dobj] that are changed in translations. A closer look reveals that 19 (59%) of these examples could be classified as the structure [NP\text{research} V\text{show clause/NP}], which has been previously discussed both in monolingual studies of English non-translated writings (Biber et al. 1999; Charles 2006; Kerz 2007; Liu 2012) as well as in contrastive and translation studies of English and German (Doherty 1996; Hansen-Schirra 2012). A high relative frequency of this pattern within the construction shift data motivates its more detailed analysis.

The construction is illustrated through example (68) below: in this sentence the noun study and the verb show fill the slots NP\text{research} and V\text{show}.

(68) **EO:** Several studies have shown that two of the factors affecting regret are how much one feels personal responsibility for the result and how easy it is to imagine a better alternative. (EO\_POPSCI\_005 s82)
Previous research on this construction has been taken as a starting point in the operationalization of its individual slots. However, even though the structure has been discussed in the literature before, the scope of the nouns and verbs included is different and is not clearly defined. Some of the studies listed above are corpus-driven investigations of lexical bundles (Biber et al. 1999; Liu 2012) and mention only the most frequent patterns and/or account only for the verbs occurring in this construction, such as suggest, show, argue, indicate, believe and claim (Liu 2012, 31). Other studies concentrate on inanimate subjects and include a discussion of certain examples of the pattern [Subj inanimate V volition Dobj] (Hansen-Schirra 2012), or focus on one of the research verbs and its translation depending on the reference of the subject (Doherty 1996). The present study uses the previous research as the starting point but develops further criteria to delimit the exact scope of the construction [NP research V show clause/NP].

To operationalize the NP research slot in the English pattern, the nouns occurring in the examples listed in various studies as well as some semantically related nouns identified with the help of WordNet have been considered. A few nouns have been included that are not specific to research but can be used in the research context. The selected nouns could refer either to a research process, e.g. study, or to a research product, e.g. result. The complete list of the nouns in this position in the English originals and translations extracted from the CroCo corpus includes the following lemmas:

account, assessment, calculation, comparison, consideration, data, discovery, dogma, evidence, experiment, figure, finding, investigation, measurement, number, outcome, projection, proof, research, result, rule, statistics, study, theory, treatment, work

While the verbs mentioned in the previous studies have been again taken as the starting point to operationalize the slot V show, some of these are excluded: for instance, the verb believe, frequently occurring in academic discourse in the pattern [NP believe that], belongs to the group of non-volitional verbs. Also the verbs find and conclude have been identified as referring only to mental activities. Since in the present study the analyzed structure is considered to be a more specific instance of the schematic transitive pattern [Subj inanimate V volition Dobj], the non-volitional verbs are not included into this category. The study considers all the verbs that according to Francis, Hunston and Manning (1996) belong to the groups of showing, saying and adding. Moreover, the group has been extended based on the semantic
relations provided by *WordNet*. In the analyzed sets of English originals and translations, the following lemmas have been identified among verbs of showing:

- back up, bear (witness), dictate, explain, imply, indicate, paint (a picture), prove, reveal, show, state, suggest

For German the initial group of nouns and verbs established on the basis of the previous literature comprises only a few items, since the comparable German construction has received less attention. Therefore, *GermaNet* was used to expand the relevant categories. As is the case with the English data, the lists consist of the nouns referring to scientific processes and their results. According to *GermaNet*, some German verbs corresponding to the English verbs of showing and saying refer only to mental activities: for instance, *GermaNet* does not list the sense “inform” for the verb *belegen* 'prove', in contrast to the English verb *prove*, as indicated by *WordNet*. Thus, these verbs have received the value “non-volitional” and sentences containing such verbs could not be considered as instances of \[NP_{\text{research}} \ V_{\text{show}} \ \text{clause/NP}\]. While it is certainly possible that the list of existing senses both in *WordNet* and *GermaNet* is incomplete, the present analysis relies on these resources, in order to make annotation decisions more transparent and replicable. The \[NP_{\text{research}} \ V_{\text{show}} \] slots are realized through the following lemmas in German originals and translations belonging to the CroCo corpus:

- Beurteilung, Daten, Ergebnisse, Extrapolation, Forschung, Messung, Studie, Untersuchung, Vorhersage, Zahl
- Ausweisen, liefern, signalisieren, verdeutlichen, zeigen

33 combinations of the research nouns and verbs of showing followed by a *that*-clause or an NP have been identified in the English originals. A larger share of these examples, namely 27 occurrences corresponding to 82%, belongs to the register of popular scientific texts. This frequency distribution could be explained taking into account the function of this structure, namely to report and explain the outcome of the studies (Liu 2012, 30). The construction can be found in those sections of academic writing, which are used for referencing to the previous work or to the findings being reported in the paper (Kerz 2007, 196-197). Thus, the structure is typically included in the lists of patterns specific to academic discourse (Biber et al. 1999; Charles 2006; Kerz 2007; Liu 2012) but can also be considered characteristic of the popular scientific writing (Hansen-Schirra 2012): the texts belonging to both registers frequently
refer to the results of various studies. A specific feature of this pattern is the impersonality effect that is created by using research nouns rather than researchers themselves as subjects of the clauses: the authors aim at presenting the findings in a more objective way by letting “facts speak for themselves” (Charles 2006, 501). Hyland (2010) claims that popular scientific writing is less impersonal than academic discourse and, therefore, the use of non-agentive subjects, including the research nouns, is more typical of the latter (Hansen-Schirra 2012, 98-99). However, the function of the analyzed structure is still more compatible with the POPSCI, rather than the ESSAY register (Hansen-Schirra 2012, 101-102).

It is essential to consider the question whether the pattern \([\text{NP}_{\text{research}} \text{V}_{\text{show}} \text{clause}/\text{NP}]\) could be a construction in its own right with an independent mental representation. As mentioned in Chapter 3, within the CxG framework it is often suggested that the constructicon is characterized by a certain degree of redundancy storing form-meaning pairings on different levels of abstractness: the specific “highly conventional” instances of the more general schemas are also recognized as constructions (Goldberg 2006, 55). Since the structure under analysis includes two slots, namely \(\text{NP}_{\text{research}}\) and \(\text{V}_{\text{show}}\), which are defined through a finite set of semantically related lexical groups, and can be distinguished by a specific function, mostly compatible with the academic and popular-scientific registers, this pattern can be postulated to be a separate form-meaning pairing, i.e. construction. The requirement of scientific discourse to present the results objectively can be assumed to motivate (cf. Goldberg 2005, 17–18) the existence of this construction (Serbina 2013, 181). As mentioned in Chapter 4.3, the identification of specific formal and functional features of a certain pattern allows us formulating a hypothesis related to an existence of this independent lower-level construction, which has to be tested further in psycholinguistic experiments (see Chapter 6). The degree of association between the slots of the construction and the lexical items filling these slots in our data sample is determined using collostructional analysis (Stefanowitsch and Gries 2003, see Chapter 4.3 for methodological details). Table 9 shows the significantly attracted collexemes for the construction slots \(\text{NP}_{\text{research}}\) and \(\text{V}_{\text{show}}\), i.e. items with the negative logarithm to the base of ten p-value above 1.30103. This value corresponds to the significance level of 0.05. It should be mentioned that for this construction there are almost no differences between the observed frequencies and the collostructional strength. The analysis does not identify any items that are repelled by the individual slots of this
construction, whereas the noun *study* and the verbs *suggest* and *show* are most strongly attracted to the slots NP\textsubscript{research} and V\textsubscript{show} respectively. Moreover, the covarying collexeme analysis suggests that the lexemes *study* and *show* are likely to co-occur in the construction [NP\textsubscript{research} V\textsubscript{show} clause/NP] (p-value log10=2.02). The patterns [*dogma states* clause/NP], [*projection paints* clause/NP] and [*theory dictates* clause/NP] could also be shown to be significantly attracted (p-value log10=1.52). However, in the latter three combinations both lexical items have been attested only once in the present data sample and, therefore, these results should be treated with more caution. These combinations yield significant results because the analysis is based not only on the frequencies with which the two items, e.g. *dogma* and *state*, co-occur in the construction [NP\textsubscript{research} V\textsubscript{show} clause/NP], but also instances of the analyzed construction including e.g. the noun *dogma* with verbs other than *state*, the verb *state* with other research nouns and occurrences of the construction in which both slots are filled with other items (Stefanowitsch and Gries 2005, 9, see Chapter 4.3). Therefore, even if the combination occurs only once but the individual items do not occur in this construction with other elements, the resulting covariance is considered significant. Based on the available numbers some possible combinations of nouns and verbs, most of which have not been observed at all in the two registers of the CroCo corpus, are classified as being repelled by the construction under analysis (see Appendix I for the top ten combinations that are attracted and repelled).

<table>
<thead>
<tr>
<th>Collexeme_NP</th>
<th>Freq. in the corpus</th>
<th>Freq. in the constr.</th>
<th>Coll. strength</th>
<th>Collexeme_V</th>
<th>Freq. in the corpus</th>
<th>Freq. in the constr.</th>
<th>Coll. strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>study</td>
<td>31</td>
<td>6</td>
<td>6.12</td>
<td>suggest</td>
<td>25</td>
<td>9</td>
<td>11.80</td>
</tr>
<tr>
<td>evidence</td>
<td>17</td>
<td>4</td>
<td>4.55</td>
<td>show</td>
<td>58</td>
<td>11</td>
<td>11.04</td>
</tr>
<tr>
<td>assessment</td>
<td>12</td>
<td>3</td>
<td>3.58</td>
<td>reveal</td>
<td>14</td>
<td>4</td>
<td>4.91</td>
</tr>
<tr>
<td>experiment</td>
<td>13</td>
<td>2</td>
<td>2.05</td>
<td>imply</td>
<td>3</td>
<td>2</td>
<td>3.44</td>
</tr>
<tr>
<td>dogma</td>
<td>3</td>
<td>1</td>
<td>1.48</td>
<td>indicate</td>
<td>14</td>
<td>3</td>
<td>3.36</td>
</tr>
<tr>
<td>work</td>
<td>29</td>
<td>2</td>
<td>1.39</td>
<td>back up</td>
<td>2</td>
<td>1</td>
<td>1.65</td>
</tr>
<tr>
<td>research</td>
<td>30</td>
<td>2</td>
<td>1.36</td>
<td>paint</td>
<td>2</td>
<td>1</td>
<td>1.65</td>
</tr>
<tr>
<td>finding</td>
<td>4</td>
<td>1</td>
<td>1.35</td>
<td>dictate</td>
<td>2</td>
<td>1</td>
<td>1.65</td>
</tr>
</tbody>
</table>

Table 9: [NP\textsubscript{research} V\textsubscript{show} clause/NP] - Significantly attracted collexemes for NP\textsubscript{research} & V\textsubscript{show}
Before examining the construction shifts from this English construction, it is essential to establish whether English and German differ not only with respect to frequency distributions of the sense [Subj\_animate V\_volition Dojb], but also of the lower-level construction [NP\_research V\_show clause/NP]. Based on the co-occurrences of the verb *show/zeigen* with various types of subjects, Doherty (1996) distinguishes different degrees of lexicalization of these verb senses in English and German. She suggests that while the patterns [Subj\_human *show*] and [Subj\_representation *show*], e.g. *figure shows*, are common in both languages, the structure [Subj\_event/process *show*] is more typical of English. Doherty (1996, 602) claims that experiments *show* would normally not be translated as *Experimente zeigen* ‘experiments show’: instead, this combination can be translated as an adverbial (Doherty 1996, 603). Both the second and the third senses are more abstract realizations of the concrete action, during which a human agent shows something. However, a figure or another form of a result representation is more closely related to the concrete meaning of the verb *show* than an abstract research process.

Doherty (1996, 605-607) concludes that this pattern is lexicalized to a lesser degree in German than in English. It is interesting to note that even though Doherty (1996) distinguishes between different senses of the verb *show* rather than different constructions, the discussion of the verb’s co-occurrences with various subjects suggests that more complex patterns play a role in language use and should be considered in linguistic analysis. Therefore, the reported study could be interpreted as providing indirect arguments in favor of the CxG approach.

As mentioned earlier, English and German differ in the types of semantic roles that can be mapped onto the subject. Hawkins (1986, 69) argues that the more rigid word order in English, which dictates that the subject’s unmarked position is before the finite verb, has been compensated by more flexibility with regard to possible semantic content expressed through the subject. This, in turn, is explained through a necessity to have several information structure options. In other words, whereas German could use constituent order to express the information in a certain way, English could accomplish the same by an extended set of semantic roles of the subject (see also Rohdenburg 1974, 412; Hansen-Schirra 2012, 95). Doherty suggests that the processing factor should also be taken into account. She argues that in German more flexibility with regard to the semantic roles of the subject would result in more processing effort on the part of the listener or reader, because the main verb often
occurs in the final position and, therefore, cannot be used to interpret the initial NP (Doherty 1996, 607-608). This factor is especially relevant if several NPs have identical morphological case markings. In these cases the language users cannot rely on the case assignment as a reliable indicator of the corresponding grammatical function.

Based on the discussion in Doherty (1996) it could be assumed that the analyzed construction is more frequent in English than in German originals. To test this hypothesis, the occurrences of the construction are first compared in the English and German originals of the CroCo corpus. Figure 9 shows the distribution of the construction in both English and German non-translated texts: we can observe that the English construction occurs 33 times, while the corresponding German construction is present only three times. While the English construction is more often shifted than kept in translations to German, all occurrences of the German pattern are not changed in translations to English.

![Figure 9: [NP_research V_show clause/NP] in EO and GO](image)

<table>
<thead>
<tr>
<th></th>
<th>NP_research V_show clause/NP</th>
<th>Subj V Dobj_other</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>33</td>
<td>217</td>
</tr>
<tr>
<td>German</td>
<td>3</td>
<td>169</td>
</tr>
</tbody>
</table>

Table 10: [NP_research V_show clause/NP] - Contingency table for construction & language

As can be seen, our data from the registers of essays and popular scientific texts confirms the hypothesis: the construction [NP_research V_show clause/NP] occurs more frequently in English than in German ($\chi^2=15.7$, df=1, p=7.42e-05, see Table 10). Since, as discussed in Chapter 3, the higher frequency of a construction is typically associated with a higher degree of its entrenchment in the speakers’ minds, these results indicate that the analyzed construction is more entrenched in English than in German. It can be assumed that the writers of the German texts tend to select other patterns to express the same function, i.e. to report on the results of the studies. Some possibilities are discussed below while looking at translations of the
English \([\text{NP}_{\text{research}} \ V_{\text{show}} \ \text{clause}/\text{NP}]\) construction. A more detailed account of these German structures is beyond the scope of the present study.

As there are only few examples of the construction \([\text{NP}_{\text{research}} \ V_{\text{show}} \ \text{clause}/\text{NP}]\) in the two selected registers of the CroCo corpus, it is necessary to verify the frequency distributions using reference corpora, namely the British National Corpus (BNC) and the corpus das Digitale Wörterbuch der deutschen Sprache (DWDS) (see Chapter 4.3). Since the corpora are not annotated for grammatical functions, a data extraction procedure comparable to the CroCo corpus is not possible. This is unfortunate because an extraction of all instances of the schematic \([\text{Subj} \ V \ \text{Dobj}]\) construction would ensure that all examples of the analyzed construction are considered, even though, taking into account the overall size of the reference corpora, a creation of random samples would be required. The extraction of the relevant construction has been based on the individual lexical items occurring in the slots \(\text{NP}_{\text{research}}\) and \(V_{\text{show}}\): (one of) the most frequent lemmas filling the NP slot and the most frequent verb lemma occurring in the instances of the analyzed construction belonging to English and German originals and/or translations extracted from the CroCo corpus have been queried in BNC and DWDS. Only the most frequent verb lemma has been selected due to frequencies of verbs in the CroCo sample of the German data, in which only one verb occurred more than once in the analyzed construction, namely \(\text{zeigen} \ 'show'\). Moreover, only one lemma for the noun slot has been investigated in the reference corpora because of high numbers of hits retrieved for every combination of a research noun and a verb of showing in the English data. This limitation was necessary to perform a subsequent manual analysis of the data.

For English the combinations of the noun \(\text{study}\) and the lemma \(\text{show}\) were considered: besides being the most frequently attested lexical items, the covarying collexeme analysis also identified the combination of these lemmas as being most strongly attracted to the analyzed construction. Taking into account that English is an SVO language, the BNC query targeted all sentences in which the selected noun occurs before the verb. The span between the elements was set to four words: a verb could occur either in the position immediately following the noun or be separated from the noun with up to four lexical items. Therefore, also cases with several auxiliaries or additional elements were extracted. This query resulted in 932 sentences that were further analyzed. A small number of duplicates were excluded. To make this investigation comparable with the main data sample from the CroCo corpus, the
structure had to fulfill certain requirements to be considered as a true hit. The focus of this sub-study is also on the construction \([NP_{\text{research}} \ V_{\text{show}} \ \text{clause}/NP]\) that does not involve any additional grammatical functions, such as an adverbial, occurs on the level of the main clause and is the only main clause of the sentence. However, the frequencies of these three variations of the pattern were also determined. In contrast, the sentences in which the selected items happen to occur in close vicinity but do not form an instance of the analyzed construction were excluded from further analyses. Among the group of excluded sentences are also cases in which combinations of the selected noun and one of the verbs have been condensed into a NP with the noun *study* as the head.

The German corpus was queried for combinations of the lemmas *Ergebnis* ‘result’ and *zeigen* ‘show’. Even though several research nouns occurred twice in the constructions extracted from the German data of the CroCo corpus, the selected noun is the only one that occurs not only in the translations but also in the originals. Due to the extremely small number of instances of this construction in the German originals of the CroCo corpus, no collostructional analysis for German has been conducted. Similar to the English queries, the span has been set to four elements, but this time the verb could occur either before or after the noun, as the word order is less fixed than in English. Moreover, the query also considered compounds with the head noun *Ergebnis*. A total of 134 sentences have been extracted from the DWDS corpus. However, it is important to mention that this number excludes 20 hits that have been identified but not shown by the corpus interface due to copyright restrictions (cf. Geyken 2007, 37-38).

<table>
<thead>
<tr>
<th></th>
<th>EO (BNC)</th>
<th>GO (DWDS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(NP_{\text{study/Ergebnis}} \ V_{\text{show/zeigen}} \ \text{clause}/NP)</td>
<td>371</td>
<td>23</td>
</tr>
<tr>
<td>Other sentence-level constructions</td>
<td>6,025,913</td>
<td>5,171,305</td>
</tr>
</tbody>
</table>
To compare the frequencies of the construction in the two reference corpora, the overall number of sentences in the two reference corpora has been taken into account. As shown in Table 11, the English and German constructions investigated in this sub-chapter account for less than 1% of all sentences in the respective reference corpora. These low numbers can be explained taking into account that the construction under analysis is a lower-level form-meaning pairing associated with specific registers. Moreover, since the queries were restricted to only one combination of lexical items that can occur in this construction, the proportion does not reflect the actual distribution of the pattern in the corpora.

The results from BNC and DWDS are in line with the findings reported above: the construction variant \([\text{NP}_{\text{study}} \text{V}_{\text{show}} \text{clause}/\text{NP}]\) is more frequent in English than in German originals \((\chi^2=256.42, \text{df}=1, \ p<2.2e-16)\). To reduce the possibility that this difference is due to the different selected lemmas in the two languages, the analysis has been repeated with the lemma \text{Studie} ‘study’ and resulted in only 14 true hits as well as 10 sentences involving additional grammatical functions and one sentence involving the analyzed construction as one of the main clauses. Since this query has even less hits and no differences with regard to the distribution among the variants of the construction as compared to the structure with the lemma \text{Ergebnis} (see Figure 10), the remainder of the discussion concentrates on the patterns selected initially, namely \([\text{NP}_{\text{study}} \text{V}_{\text{show}} \text{clause}/\text{NP}]\) and \([\text{NP}_{\text{Ergebnis}} \text{V}_{\text{zeigen}} \text{Clause}/\text{NP}]\).

As can be seen in Figure 10, 54% (371/684) of English and 52% (23/44) of German sentences involving the construction under analysis are classified as true hits in both languages: as mentioned earlier, these instances consist of subject, verb and direct object as the only three grammatical functions on the level of the main clause. The other three types of the construction, which are not considered in more detail in the present study, have lower
frequencies in the reference corpora comparable across the two languages. During the classification it was first established whether the analyzed construction is present at all and if yes, whether it is located on the level of the subordinate clause: for instance, in examples (69) and (70) the pattern is embedded into the direct object. In these cases, even if the construction has additional elements or coordinated clauses, the sentences have been considered to belong to the category of subordination and thus counted as instances of the variant involving subordination. If, however, the construction has been identified on the level of the main clause, the analysis proceeds to verify whether it is the only main clause of the sentence. In (71) and (72) the patterns [NP\text{study} V\text{show} clause/NP] and [NP\text{Ergebnis} V\text{zeigen} Clause/NP] are coordinated with another main clause through a coordinating conjunction. Such cases are included into the category “several main clauses”. Finally, if the sentence consists only of the examined structure, it was established whether there are any additional grammatical functions involved: for example, (73) and (74) are characterized by additional adverbials and do not belong to the final sample. Future studies could complement the present research by considering in more detail these three variants not included in the category of true hits.

(69) He says [[all the studies]Subj [show]V [sunburning is linked to cancer]Dobj]Dobj. (BNC K1M W\_news\_script)


(71) [[Repeated studies]Subj [have shown]V [low use]Dobj]Clause but [little effort has been made to find out why]Clause. (BNC H0S W\_misc)

(72) [Es ist noch nicht lange her, seit begonnen wurde, diese Beschlüsse in die Tat umzusetzen]Clause, aber [[schon die ersten Ergebnisse]Subj [zeigen]V, [daß wir auf dem richtigen Wege sind]Dobj]Clause. (DWDS adg19671109 Zeitung)


In the next step let us have a closer look at the true hits extracted from the reference corpora. Figure 11 shows that in English the construction is most frequently realized through the pattern [studies have shown that-clause] illustrated by (75). In this pattern, similar to the rest of the data sample, the head noun studies is typically premodified through a quantifier (e.g. several, many, some), a determiner (these or other), an adjective providing a time reference (previous or recent), an adjective or noun referring to the type of the study (e.g. population studies, epidemiological studies), or through a combination of these elements. The instances in which the lemma study is postmodified with or without additional premodification occur much more seldomly. These can refer to the object, time or place of the study as well as those responsible. Moreover, there are also some examples without any type of modification. (75) More recent studies have shown that hepatic denervation causes significant changes in the biliary lipid composition. (BNC HU2 W_ac_medicine) (76) Diese Ergebnisse zeigen, daß mit Hilfe des Enzyms die Basenzusammensetzung der primer-DNA kopiert wird. (DWDS bresch Wissenschaft)

As shown in Figure 12, in German the patterns [Ergebnis zeigt Clause] and [Ergebnisse zeigen Clause] have a higher frequency than the other variants. For instance, (76) could be considered a typical representative of the German construction. Similar to English, the NP Ergebnis occurs more often only with premodifying and without postmodifying elements (14 out of 23 instances). However, premodification is typically limited to a determiner. This difference could be at least partly explained through different lemmas selected in English and German: the lemma Studie 'study' filling the construction’s NP slot displays more variation in terms of premodifiers. Also the postmodifiers depend on the concrete noun. When present, the postmodification of Ergebnis is typically realized through a genitive NP characterizing results in terms of their source, e.g. das Ergebnis der Viehzählung ‘the result of the livestock census’ or die Ergebnisse der Umfrage ‘the results of the survey’, whereas the postmodification of Studie usually names those responsible and/or the type of the study. This information is expressed through a prepositional phrase or a relative clause.
Finally, the distribution of the analyzed construction across different registers of the reference corpora was compared. In the English sample, as expected, the construction mostly occurs in the registers of academic and popular-scientific texts, consisting of the BNC categories of academic and non-academic texts, the latter group being roughly equivalent to popular-scientific discourse. 80% (296/371) of all instances under analysis belong to these registers. In German, the small samples for both [NP Ergebnis zeigen Clause/NP] (23 true hits) and [NP Studie zeigen Clause/NP] (14 true hits) indicate an almost even distribution of examples among the (popular-) scientific and newspaper/magazine prose. One explanation would be the claim that the construction [NP research show clause/NP] was introduced to German through journalistic texts (König and Gast 2009, 108). However, as with the other analyses for German, the small overall number of instances allows only tentative statements.

Since comparisons of English and German originals in the CroCo and the reference corpora have confirmed the assumption that the construction [NP research show clause/NP] is more

---

25 Since the reference guide for the British National Corpus (Burnard 2007) does not provide descriptions of registers included into the corpus, the equivalence between the category of the non-academic discourse and the popular-scientific register is assumed based on the meta-information of the individual texts that belong to this part of the BNC.
frequent and, therefore, potentially more entrenched in English than in German, the translation direction English-German is expected to be characterized by a higher number of shifts from the analyzed construction than translations through a corresponding German pattern.

<table>
<thead>
<tr>
<th>Construction shifts</th>
<th>NP_{research} \ V_{show} \ clause/NP</th>
<th>Subj \ V \ Dobj_other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>19\textsuperscript{26}</td>
<td>82</td>
</tr>
<tr>
<td>No construction shifts</td>
<td>14</td>
<td>135</td>
</tr>
</tbody>
</table>

Table 12: [NP_{research} \ V_{show} \ clause/NP] - Contingency table for construction & construction shift

As mentioned above 19 instances of the construction [NP_{research} \ V_{show} \ clause/NP] extracted from the CroCo corpus are affected by a construction shift in contrast to 14 cases identified within the data set without shifts. The p-value for effect of this specific construction on the variable of construction shift equals the conventional 0.05 level of significance ($\chi^2=3.87$, df=1, p-value=0.05, see Table 12\textsuperscript{27}). Thus, even though the p-value fails to reach the significance level, it could be assumed that there is a certain tendency to change this construction in the translation direction English-German. Due to the small data sample of the analyzed construction, the finding should be tested further. In the next step the examples with and without construction shifts are discussed in more detail.

Construction shifts affecting the construction [NP_{research} \ V_{show} \ clause/NP] often go hand in hand with changes of subjects. The present study uses the classification introduced in Kast (2012, 150-151), which identifies three main types of variation. Thus, Variation A is present when the semantic information mapped onto the original subject is translated through a different grammatical function, as in (77), whereas Variation B refers to sentence pairs in which the semantic information mapped onto the original subject is not translated (see (78)). Moreover, Kast (2012) distinguishes variation C, which considers sentences with no subject

\textsuperscript{26} It should be noted that 5 out of 19 examples characterized by construction shifts and 3 out of 14 examples of this construction from the group without construction shifts occur in the three texts taken from Doherty's corpus (see Chapter 4.1). Since these sentences have been modified to optimize the information structure, it is possible that the changes affected the presence or lack of a construction shift. However, because examples from these texts occur in both categories (with and without shifts), we assume that the changes did not specifically target this particular construction.

\textsuperscript{27} Due to some minor changes in the classification of verbs according to volition and further manual corrections related to the annotation of grammatical functions and alignment, the result of the chi-square test presented here differs from the one in Serbina (2013, 183). Even though in the earlier version of this study the p-value was just below 0.05 and, thus, based on convention, could be considered significant, the reported Phi coefficient was very low ($\varphi=0.12$) pointing towards little or no association between the two variables.
in the original. Since all occurrences of the analyzed construction have to contain a subject, there are no examples of this type in the data sample.

(77) **EO:** [A growing library of results]$_{\text{subj}}$ reveals that the central dogma is woefully incomplete for describing the molecular biology of eukaryotes. (EO_POPSCI_004 s17)

**GTrans:** [Wie mehr und mehr Forschungsergebnisse der letzten Jahre zeigen]$_{\text{adv}}$, beschreibt das zentrale Dogma die Molekularbiologie der Eukaryoten nur sehr unzureichend.

(78) **EO:** [This discovery]$_{\text{subj}}$ suggests that delta FosB may be responsible for the added spines. (EO_POPSCI_003 s85)

**GTrans:** Ist dafür die noch vorhandene erhöhte Konzentration von delta-FosB verantwortlich?

In the construction shift data Variation A occurs in 10 of 19 pairs of clauses. Similar to (77) the subject is often shifted to an adverbial. Other examples include such adverbials as *nach Untersuchungen verschiedener Wissenschaftler* [...] ‘according to investigations of different scientists’, *der Beurteilung zufolge* ‘in accordance with the assessment’ or *durch weitere gezielte Versuche* ‘by means of further targeted experiments’. The more flexible word order in German allows the translator to keep the order of the lexical information constant, which is often preferred to preserving the order of grammatical functions (Čulo et al. 2008, 49; Solfjeld 2008). In five further instances of the construction [NP$_{\text{research}}$ V$_{\text{result clause}}$/NP] undergoing a construction shift Variation B has been identified: the head noun or the whole NP functioning as the subject in the original does not have any correspondence in the translation. For instance, in (78) the NP *this discovery* as well as the verb *suggest* have been dropped in the aligned clause. To reproduce the hedging effect of the original structure, the mood is changed from indicative to interrogative, affecting the word order. Moreover, the anaphoric reference of the English pronoun *this* is realized through the pronominal adverb *dafür* ‘for that’, a category more typical of German (e.g. Kunz and Steiner 2013, 225). The cases in which the original subject is split and realized through several grammatical functions in the translation are classified in a number of ways. If the head of the NP functioning as the subject in the original does not have any correspondence in the translation. For instance, in (78) the sentence pair is not regarded to involve subject variation. If all parts of the original subject are realized in the translation through several clause constituents other than subject, then it is an example of Variation A, whereas the examples in which the head of the original subject is not present in translation.
are considered to belong to Variation B. For instance, in (79) the subject of the source sentence, taken from one of Doherty’s texts, is split: even though in the German translation most of the semantic information corresponding to the original subject is realized through the prepositional object and the adverbial of place, the head of the original subject, namely *studies*, is missing in the target sentence. Therefore, this sentence exhibits Variation B. These changes result in different orders of semantic information in the original and the corresponding translation (which could arguably be the result of local text modifications performed by Doherty’s team).

(79) **EO:** Studies at Stanford of one subtype of interferon, 2, show anti-cancer activity and side effects very similar to those of the mixture. (EO_POPSCI_010 s104)

**GTrans:** Für den Interferontyp Alpha-Zwei konnten in Stanford fast dieselben Antikrebs- und Nebenwirkungen nachgewiesen werden wie für das Alpha-Interferon-Gemisch.

(80) **EO:** But no strong evidence has backed up that presumption. (EO_POPSCI_004s98)

**GTrans:** Überzeugende Belege, die diese Annahme stützen, fehlen aber.

The present study adds a further category of subject variation not discussed in Kast (2012), namely the cases of subject merge (Variation D) illustrated through example (80). In this case several grammatical functions of the original, namely the NP head and its premodifying adjective functioning as the subject, the volitional verb and the direct object are merged into the target subject. The verb introduced in the translation corresponding to the negation particle in the original subject is a non-volitional one. Therefore, the construction [NP*research V*show Dobj] is changed to the intransitive pattern [Subjinanimate V*non-volition*. While in our data sample this is the only example of the analyzed construction affected by this variation type, the category is of potential interest for future studies.

Furthermore, the phenomenon of sentence splitting also plays a role in the analysis of subject variation. In contrast to Solfjeld (2008, 118), who also accounts for sentences split into several coordinated clauses connected through the conjunction *and* or separated by a comma, only cases in which one sentence in the original corresponds to several sentences separated by a full stop in the translation are regarded as instances of sentence splitting in the present work. As mentioned in Chapter 4.2.2, sentences consisting of coordinated main clauses are not included into the analysis. However, coordination could be present on a lower level integrated into the subject or the direct object. These cases as well as other types of
complex NPs may trigger sentence splitting in translation (Solfjeld 2008, 117). Among the examples of the construction \([\text{NP}_{\text{research}} \ V_{\text{show}} \ \text{clause}/\text{NP}]\) affected by construction shifts sentence splitting occurs four times.

In the cases of sentence splitting the source sentence is aligned to all corresponding target sentences. However, for practical purposes of identifying construction shifts, the Perl script has only extracted the first of the aligned sentences in the translation. For qualitative analyses of individual constructions the missing target sentences have been extracted as well (see Chapter 4.2.1). Since the subject plays an important role in sentences in general (Quirk et al. 1985, 724) and in the discussed constructions in particular, the decision which of the several aligned sentences to use for identification of construction shifts has been based on the subject of the original sentence. This decision is to some extent problematic, especially if the rest of the sentence has different lexical content than the original construction, as in example (81). Here certain elements of the original subject are translated in a different sentence than the rest of the source sentence. As a result, the first target sentence concentrates on the fact reported in the original that-clause, conforming to the tendency of German texts to be content-oriented (House 1997, 84). The uncertainty of the original formulation is conveyed through the particle wohl ‘probably’, i.e the word class, which is more typical of German than of English (König and Gast 2009, 239, 245). Thus, the translator, similar to example (78) above, uses linguistic features typical of German to shift from the English construction. The second target sentence involves parts of the English subject: logic corresponds to the adjective logische ‘logical’ and evidence is translated by Indizien ‘evidence’, both of which are integrated into a structure premodifying another research noun, namely Hypothese ‘hypothesis’. The latter research noun functions as the head noun of the subject in the translated sentence, which is combined with another verb of showing or saying, namely erklären ‘explain’. In other words, despite the multiple changes involved in the translation of the original construction, the translation exhibits a variant of the analyzed construction \([\text{NP}_{\text{research}} \ V_{\text{show}} \ \text{clause}/\text{NP}]\) with an additional grammatical element. Since the second target sentence contains lexical items corresponding to the original subject, it is aligned with the English source sentence. In contrast, when the source subject is missing in the translation, the first of the aligned sentences is considered. As a consequence of this decision all instances
of sentence splitting could be classified as Variation A or B, or as examples not affected by subject variation.

(81) **EO:** [Logic and the available evidence]$_{subj}$ [suggest]$_{verb}$ that the rise of multicellular organisms over the past billion years was a consequence of the transition to a new control architecture based largely on endogenous digital RNA signals. (EO_POPSCI_004 s128)

**GTrans:** Der Aufstieg der Vielzeller, der vor etwa einer Milliarde Jahren begann, ist wohl dem Übergang zu einer neuartigen Steuerarchitektur zu verdanken, die weitgehend auf eigen digitalen RNA-Signalen beruht (siehe Grafik rechts oben). Zumindest konnte [diese logische, aus verschiedenen Indizien ableitbare Hypothese]$_{subj}$ die so genannte kambrische Explosion [erklären]$_{verb}$.

<table>
<thead>
<tr>
<th>Construction shifts</th>
<th>Variation A</th>
<th>Variation B</th>
<th>Variation D</th>
<th>No subject variation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>No construction shifts</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 13: [NP$_{research}$ V$_{show}$ clause/NP] - Subject variation for the translation direction E-G

Table 13 summarizes the types of subject variation in both groups of sentences, with and without construction shifts for the translation direction English-German. In contrast to the former group, the majority of examples in the latter have the same subjects in originals and translations. Thus, these cases, illustrated through example (82), often present a translation through a comparable German construction, i.e. a combination of a research noun and a verb of showing. However, as shown in Table 13, there are also two cases of subject variation within the data without construction shifts. For instance, (83) is interpreted as an example of Variation A: part of the original subject, namely the head noun *findings*, is integrated into the direct object in the translation, where it corresponds to the noun *Anzeichen* ‘evidence’. Thus, this example is characterized through a crossing line between subject and object, but, since it has the same types of grammatical functions, it is still classified as an example without a construction shift. As a result, two translations of the construction [NP$_{research}$ V$_{show}$ clause/NP], which belong to the group without construction shifts, do not involve the comparable German construction. Moreover, one example within the shift data is characterized by the combination [NP$_{research}$ V$_{show}$] in both the original and the corresponding translation: as can be seen in (84), this construction shift can be explained by the change from
the direct to the prepositional object. The prepositional object construction has been shown to be more typical of German than English (see Chapter 5.3).

(82) **EO:** [New work]NP\_research [shows]V\_show [immune reactions in the body yield interferons.]dobj (EO\_POPSCI\_010 s77)

**GTrans:** [Neueste Untersuchungen]NP\_research [zeigen]V\_show, [daß bei den Immunreaktionen des Körpers auch Gamma-Interferone entstehen.]dobj

(83) **EO:** [Some findings]subj suggest that they can also influence the synthesis of certain glutamate receptors. (EO\_POPSCI\_003 s98)

**GTrans:** Es gibt [Anzeichen, daß sie auch auf die Herstellung einiger dieser Rezeptoren Einfluß nehmen.]dobj

(84) **EO:** These studies suggest [that openness to trade leads to declining absolute poverty rates and does not increase income inequality.]dobj (EO\_ESSAY\_008 s14)

**GTrans:** Diese Studien weisen [darauf] hin, [daß Offenheit für den Handel zum Rückgang der absoluten Armut führt und Einkommensunterschiede nicht verstärkt.]probj

The above discussion has indicated that the English construction [NP\_research V\_show that-clause/NP] is often shifted in translations. In particular, the research nouns functioning as subjects are often changed to other grammatical categories or deleted. I assume that the different number of occurrences of this construction in English and German originals is at least one of the major motivations for the change of this structure during the process of translation.

In order to analyze whether there are any differences between the occurrences of the analyzed construction in originals and translations in both languages, the translated texts should be considered in more detail.

<table>
<thead>
<tr>
<th></th>
<th>NP_research V_show clause/NP</th>
<th>Subj V Dobj_other</th>
</tr>
</thead>
<tbody>
<tr>
<td>German originals</td>
<td>3</td>
<td>169</td>
</tr>
<tr>
<td>German translations</td>
<td>9</td>
<td>181</td>
</tr>
</tbody>
</table>

Table 14: [NP\_research V\_show Clause/NP] - Contingency table for construction & translation status, German

Even though the difference between the frequencies of this construction in German originals and translations is not statistically significant ($\chi^2=1.68$, df=1, p=0.2, see Table 14), there are somewhat more instances of the construction [NP\_research V\_show that-clause/NP] in German translations. It is interesting to note that these examples are in most cases motivated by a comparable construction in the English originals, so that the feature of the SL could be said to “shine through” in the translations. Translations from English along with the status of
English as a lingua franca have been named as potential reasons for the extension of possible semantic roles mapped onto subjects in German (König and Gast 2009, 108; Hansen-Schirra and Steiner 2012, 274). I refer to this phenomenon as construction borrowing (Serbina 2013, 187). However, at the same time the construction is also frequently changed in the translations from English to German. Since the analysis of the CroCo and reference corpora has shown that the construction is more frequent in English than in German, shifts from this construction can be interpreted as the translation property of normalization. Moreover, normalization is often accompanied by either implicitation or explicitation through deletion or addition of certain semantic content present in the original28. Also such features of the original construction as information structure and the degree of impersonality may be affected in the process of translation of the analyzed construction. For instance, in example (85), classified as a case of subject variation B, both the research noun study and the verb show are missing in the translation. The translation is, thus, characterized both by normalization and implicitation. It is also more impersonal than the original: thus, the impersonal English construction is exchanged to increase the degree of impersonality even further. This finding is in line with another contrastive difference between English and German, according to which German tends to focus on content rather than persons (House 1997, 84). Finally, information structure is affected less often than the semantic content and the degree of impersonality. As reported in previous studies (Čulo et al. 2008, 49; Soljfeld 2008), the information structure is mostly kept constant.

(85) **EO:** Studies of mutant mice that produce excessive amounts of delta FosB in the nucleus accumbens show that prolonged induction of this molecule causes animals to become hypersensitive to drugs. (EO_POPSCI_003 s78)  
**GTrans:** Mäuse, die dieses Protein wegen einer genetischen Mutation im Nucleus accumbens im Übermaß bilden, neigen von allein zu Überempfindlichkeit auf Drogen.

<table>
<thead>
<tr>
<th></th>
<th>NP&lt;sub&gt;research&lt;/sub&gt; &lt;i&gt;V&lt;sub&gt;show&lt;/sub&gt;&lt;/i&gt; clause/NP</th>
<th>Subj V Dobj&lt;sub&gt;other&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>English originals</td>
<td>33</td>
<td>217</td>
</tr>
<tr>
<td>English translations</td>
<td>13</td>
<td>218</td>
</tr>
</tbody>
</table>

Table 15: [NP<sub>research</sub> V<sub>show</sub> Clause/NP] - Contingency table for construction & translation status, English

---

28 We adopt a broad definition of explicitation considering cases where some information is added in the translation that is not present in the corresponding original.
In translations into English, the construction \([\text{NP}_{\text{research}} \text{V}_{\text{show}} \text{clause}/\text{NP}]\) occurs 13 times, in contrast to 33 instances in the English originals. This difference is statistically significant \(\chi^2=7.11, \text{df}=1, p=0.008\), see Table 15) and suggests that the English translations are also characterized by the translation property of shining through: similar to the German originals, this group of texts is assumed to include other constructions when referring to the outcome of the studies. From the perspective of the aligned data, we could observe that in the translation direction German-English the construction \([\text{NP}_{\text{research}} \text{V}_{\text{show}} \text{clause}/\text{NP}]\) remains unchanged. This finding is not surprising taking into account that the construction is frequently used in the target language. Thus, in three cases the analyzed construction in the English translations corresponds to the comparable construction in the German originals. Moreover, one further instance of the analyzed pattern corresponds to the German original involving the construction \([\text{NP}_{\text{research}} \text{V}_{\text{show}} \text{clause}/\text{NP}]\) with an additional grammatical function, while in another example the construction shift has been caused by a crossing line between a prepositional object in the original and a direct object in the translation. Thus, also this German sentence is characterized by the combination \(\text{NP}_{\text{research}}\) and \(\text{V}_{\text{show}}\). All of these instances, in which the whole construction, or at least the subject and the verb are left unchanged, cannot be analyzed in terms of the translation properties of shining through or normalization. The former property is present, if the feature, which is more characteristic of the SL, is left unchanged in the TL. Assuming that the construction is more entrenched in English than in German, the cases cannot be interpreted as interference from German. Moreover, also the opposite property, namely normalization, should be assigned to those target sentences, which are normalized, i.e. changed to the more typical TL construction, which is not the case in these sentence pairs as the corresponding construction is present in the aligned originals and translations. In contrast, eight further examples, illustrated through (86), are characterized by shifts into the English construction \([\text{NP}_{\text{research}} \text{V}_{\text{show}} \text{clause}/\text{NP}]\) and could, therefore, be interpreted as normalization.

(86)  
\textbf{GO:} Aber nachweislich wird die Leipziger Messe zur gleichen Zeit erwähnt wie die Messen in der Champagne und in Frankfurt am Main. (GO\_ESSAY\_011 s51)
\textbf{ETrans:} But \([\text{the proof}]_{\text{NP}_{\text{research}}} [\text{reveals}]_{\text{V}_{\text{show}}} [\text{that the fair in Leipzig was being mentioned at the same time as those in Champagne and Frankfurt am Main}]_{\text{clause}}\).
In this pair of sentences the modal adverbial *nachweislich* ‘evidently’ in German original is translated as the [NP_research V_show], whereas the content of the main clause corresponds to the *that*-clause functioning as the direct object. In this and similar examples it could be argued that the translator normalizes the structure towards a more typical one in English. Thus, in five out of these eight examples the research noun is either missing completely in the aligned German original or is expressed through an adverbial, as in (86). In the other three cases, the original contains the combination of an inanimate research noun and the non-volitional verbs *besagen* ‘indicate’ and *belegen* ‘inform’. While *besagen* is classified as a non-volitional verb because it cannot occur with subjects referring to human beings, *belegen* does not fulfil the second requirement, as it can only denote a mental activity (see Chapter 5.1 for the criteria of volition).

The analysis has shown that the transitive pattern [Subj_inanimate V_volition Dobj] and its more specific subset corresponding to the construction [NP_research V_show clause/NP] are used more frequently in English than in German. Thus, based on the corpus-to-cognition principle the English lower-level construction could be assumed to be more entrenched in the speakers’ minds, especially its more specific realization [NP_study V_show clause/NP]. Though the analysis of construction shifts involving this construction has not reached the level of significance, there is a certain tendency to change the pattern under discussion to another German construction typically by translating a research noun by an adverbial, even though in other cases the pattern is translated using a comparable TL construction. This combination leads to an interplay of the two translation properties, namely normalization and shining through in the German translations of this construction. The former property also plays a role in the English translations, when the more typical English construction is introduced in the target text. Moreover, when the research noun functions as the subject in the German originals, it is likely to be kept in the English translations: this could be explained taking into account that inanimate subjects in general are more typical of English than of German and, therefore, in these cases the translator has good reasons for using a comparable noun in English. In the next step, the transitive pattern with animate subjects is accounted for.
5.1.2 Subjanimate Vnon-volition Dobj

As indicated in Figure 7 above, the pattern [Subjanimate Vnon-volition Dobj] occurs in the English data significantly more often than would be expected given the standard normal distribution. Thus, while in English this transitive pattern accounts for 12% (31/250) of the construction [Subj V Dobj], in German it reaches only 3% (5/172). To explain this difference, the structure is examined in more detail.

In the English originals 61% (19/31) of the instances, all belonging to the register of political essays, could be classified as [NPcountry/government Vfeel/think clause/NP]. Apart from the nouns government and country (or its name), the subject of this pattern is often realized through the personal pronoun we. As mentioned earlier, I consider the use of this pronoun in the register of political essays inclusive, i.e. encompassing the speaker and the party, the Government and potentially also the addressees constituting the people of the country (cf. Pyykkö 2002, 233; Fetzer 2008, 389; Fetzer and Bull 2008, 275). In other words, the pronoun can be said to denote a group with a “collective voice” and is, therefore, annotated as an organization, which is, in turn, categorized as an animate entity (Garretson 2004; Zaenen et al. 2004).

The discussed pattern could be illustrated using the second sentence in (87). The example is taken from an essay by Tony Blair, which addresses British EU policy as a response on proposals for the EU budget. The preceding sentence explicitly mentions Europe and its people, reinforcing the function of the inclusive we that follows: it is implied that the opinion is shared not only by the Prime Minister but rather by all the people of Europe. Thus, the aim of this political essay is to persuade as many people as possible of the necessity of a different budget. The use of this pattern can be connected to the two strategic functions of political discourse, namely legitimization and coercion (Chilton and Schäffner 1997, 212-213; see also Marín Arrese 2011, 194). The functions typically co-occur as in democratic societies the politicians try to convince the people to agree with their assertions and actions (Chilton and Schäffner 1997, 213).

(87) EO: Spending four out of every ten euros on the Common Agricultural Policy - and on less than five per cent of the population - simply isn’t a budget which meets the needs and challenges of Europe and its people. [We]country/Government [need]feel/think [a budget which concentrates on the future.]NP (EO_ESSAY_003 s9-10)

(88) EO: We [hope]effective_stance that steel companies worldwide will follow that lead to produce a healthier, freer international steel market. (EO_ESSAY_016 s26)
The combination of the NP\textit{country\text{;/}government}, a verb of feeling or thinking and a NP or a clause could be used to express the author’s stance to the events and transmitted propositions. Stance involves implicit or explicit communication of “personal feelings, attitudes, value judgments, or assessments” (Biber et al. 1999, 966) and, therefore, deals with the reflection of author’s subjectivity and evaluation (Englebretson 2007, 16-17). Marín Arrese (2011) distinguishes between two general types of grammatical stance devices used in political discourse, namely effective and epistemic stance. Effective stance is defined as “speaker’s attitudes towards an event”, the perceived desirability or necessity realized by e.g. \textit{should}, \textit{it is necessary to}, \textit{I wanted to}. The speaker is said to express epistemic stance when s/he assesses the probability or communicates the certainty that an event will take place using such expressions as \textit{I am confident that}, \textit{I think}, \textit{we all know} (Marín Arrese 2011). These groups are comparable to the categories of attitudinal and epistemic stance markers identified by Biber et al. (1999, 972-975).

For instance, (87) and (88) illustrate effective stance: while the former expresses the necessity of an action, the latter refers to the desirability of an event. In contrast, (89) is an example of epistemic stance as it describes the assurance that the proposition expressed in the following \textit{that}-clause is true. It should be mentioned that both types of stance and their sub-categories can be realized by a variety of grammatical features (see e.g. Biber et al. 1999). However, the present study accounts only for the examples of the transitive construction. To be more precise, I focus on the expression of stance that is attributed to an organization rather than an individual author. Marín Arrese (2011, 214-215) suggests that the political figure may use certain linguistic devices to imply shared responsibilities for the evaluation of the expressed propositions (cf. Nuyts 2001, 34-36 on the notion of the intersubjective evidentiality), in particular by the use of the pronoun \textit{we}. Therefore, it can be argued that the structure under analysis, with the form \textit{[NP\textit{country\text{;/}government} V feel/think clause/NP]}, is connected to the general functions of legitimization and coercion achieved through the more specific functions, namely communication of stance in combination with shared responsibility. Since
the pattern, similar to the \([\text{NP}}_{\text{study}} \; V_{\text{show}} \; \text{clause}/\text{NP}\] construction discussed in the previous sub-chapter, is characterized through a register-specific function as well as lexical restrictions on its slots, I assume that it represents an independent lower-level construction. Furthermore, based on the distinction between the two groups of stance markers, two variants of this construction are suggested, namely \([\text{NP}}_{\text{country/government}} \; V_{\text{desirability/necessity}} \; \text{clause}/\text{NP}\] and \([\text{NP}}_{\text{country/government}} \; V_{\text{certainty}} \; \text{clause}/\text{NP}\], expressing effective and epistemic stance respectively. Marín Arrese (2011, 210) argues that the former pattern is used to convince the audience by presenting the attitude towards the realization of an action, whereas the latter has the function to achieve “the acceptance of the information as valid on the part of the hearer”.

As indicated earlier, the nouns mapped onto the subject of the construction under analysis refer to either the Government, or a country. All of the verbs filling the verb slot of the construction belong to the semantic groups related to thoughts and feelings (Francis, Hunston and Manning 1996, 51-52, 100). Thus, the investigated structure has been identified by verifying whether the nouns and verbs occurring in the instances of the \([\text{Subj animate} \; V_{\text{nonvolition}} \; \text{Dobj}]\) pattern can be assigned to these categories. The classification of verbs referring to effective or epistemic stance is based on Marín Arrese (2011). The following lemmas occur in the slots \(\text{NP}}_{\text{country/government}}\) and \(V_{\text{feel/think}}\) in our data sample of the English originals: Administration, country, Europe, Government, we, the United States;
effective stance: (can’t) accept, hope, need, (should) remember, (have to) understand;
epistemic stance: believe, know, recognize

The verbs accept, remember and understand have been included into the category of effective stance due to their combinations with modals and semi-modals denoting deonticity (Marín Arrese 2011, 202).

Results of the collocational analysis of the English construction based on our data from the CroCo are summarized in Table 16. We can observe that only one lexical item is significantly attracted to the initial NP slot, namely the personal pronoun we. Due to the different uses of the pronoun in the two registers under analysis and taking into account that the construction has been attested only in one register, the analysis has considered only the register of political essays for the calculation of the values required for the contingency tables, such as the number of occurrences of the lexical items in the corpus as well as the
overall size of this sub-corpus. Our small data sample also suggests that the lemmas *country, Europe* and *United States* are repelled by the subject slot of this construction meaning that these lexical items are much more likely to occur in different patterns. In contrast, all verbs encountered in the construction have been classified as being attracted to the V slot, five of them, presented in Table 16, significantly. From these, only the verb *believe* is used to express the epistemic type of stance. As mentioned above, the verbs *remember, accept* and *understand* express stance only in combinations with modal verbs. It is interesting to note that all of these three verbs have lower p-values than *believe, need* and *hope*, which have a more straightforward association with stance. Finally, it should be mentioned that none of the combinations of the lexemes reached the level of significance in the covarying collexeme analysis.

<table>
<thead>
<tr>
<th>Collexeme_NP</th>
<th>Freq. in the corpus</th>
<th>Freq. in the constr.</th>
<th>Coll. Strength</th>
<th>Collexeme_V</th>
<th>Freq. in the corpus</th>
<th>Freq. in the constr.</th>
<th>Coll. strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>We</td>
<td>363</td>
<td>13</td>
<td>4.42</td>
<td>believe</td>
<td>21</td>
<td>4</td>
<td>4.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>need</td>
<td>106</td>
<td>7</td>
<td>3.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hope</td>
<td>10</td>
<td>3</td>
<td>3.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>remember</td>
<td>1</td>
<td>1</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>accept</td>
<td>3</td>
<td>1</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>understand</td>
<td>4</td>
<td>1</td>
<td>1.31</td>
</tr>
</tbody>
</table>

Table 16: [NP country/government V feel/think clause/NP] - Significantly attracted collexemes for NP country/government & V feel/think

In the German originals, only one instance of the analyzed construction could be identified, which is shown in (90). This example can be characterized as expressing epistemic stance, more specifically experiential evidentiality. By using a verb of perception in combination with the inclusive *we* the author suggests that both she/he and the audience have “direct personal sensory access to the evidence” (Marín Arrese 2011, 208). Therefore, this pattern could also be interpreted as a device employed to persuade the audience.

(90) Wir erleben, wie Grenzen ihre Bedeutung als physische, rechtliche und kulturelle Barrieren - aber auch als Rahmen - für den Austausch von Waren, Menschen, Wissen und Ideen verlieren. (GO_ESSAY_009 s7)
Since both reference corpora BNC and DWDS do not include the register of political discourse, i.e. the group of texts associated with the function of the analyzed construction, the Europarl corpus of parliamentary debates (Koehn 2005) has been selected as an additional resource to extract further examples of this construction. For this purpose, I have used the web interface provided by the Centre for Translation Studies at the University of Leeds. It allows querying the Europarl corpus by using the cqp syntax (Evert and Hardie 2011). However, the interface does not support all features of this querying language. For instance, it was not possible to specify the exact span between the two lemmas or determine the overall number of hits returned by a query. Furthermore, the display appears to be limited to 1000 examples. Most importantly, even though the search options suggest that it is possible to limit the query to English or German originals, the opposite is the case: the queries for the combination *we need* in the English originals and *wir brauchen* ‘we need’ in the German originals returned partly the same examples. Therefore, we cannot be certain whether the 222 (of 1000) true hits corresponding to the construction [*Wir brauchen* NP] identified in the German data belong to the originals or the translations from English. If it is assumed that a large share of these instances does occur in the German originals, then the low number of its occurrences in the CroCo corpus could be perhaps explained by a particular selection of texts. Thus, it is plausible that the construction is used more or less frequently based on the topic and/or the author of the text.

The construction has been also queried in another CroCo sub-corpus classified as the register of prepared speeches by government officials (Neumann and Hansen-Schirra 2012, 29), which could be considered as also belonging to the political discourse. Because the extensive manual correction of annotation of English and German originals, English and German translations as well as their alignment was not performed for this register of the CroCo corpus, we did not use the annotation of grammatical functions to extract the instances of the construction but rather limited the search to specific lexical items. For this query those lemmas that occur in the English or German construction most frequently have been selected, namely [NP*we V need* clause/NP] and [NP*wir V brauchen* NP/clause]. The modal uses of the verb

---

29 The interface is available at http://corpus.leeds.ac.uk/paraquery.html, accessed on 29.05.2014.
30 As mentioned earlier, the collocational analysis of the English construction has shown that the verb *believe* is more strongly attracted to the verb slot of this structure. However, since the collocational analysis could not be performed for the corresponding German construction (due to only one occurrence of this construction in the German originals), the queries are based on the most frequent items in both languages.
need to have been excluded, because in these cases it is not the verb need but rather the following verb that is used to determine the semantic aspect of volition. Since only one instance of the construction has been identified in the German originals belonging to the registers of popular-scientific texts and political essays, the frequencies of lemmas have been determined based on their occurrences in the German translations. Taking into account the more flexible word order in German, the German queries have also considered the reversed order of lemmas. In all queries the span between lemmas has been set to four words.

As shown in Figure 13, the construction variant realized by the three grammatical functions of subject, verb and direct object as the only main clause (forming the category of true hits) occurs in German but not in English originals. Comparing these values with the frequency distributions in the corresponding translations, it could be assumed that the frequent use of the construction in the German originals is reflected in its increased use in the aligned English translations. Similarly, the construction is (almost) missing both in the English originals and the corresponding German translations. Moreover, it is also interesting to note that the frequencies in almost all groups of texts are higher in the variant involving additional grammatical functions. Also this category is characterized by comparable values for the originals and the aligned translations (see below for a more detailed discussion of translations in the register of political essays).

It should be mentioned that in English the structure [NP country/government need to V NP], which, as mentioned above, has been excluded from the final sample, seems to be preferred. Example (91) illustrates that the marginal modal need to (cf. Biber et al. 1999, 484) modifies the phrasal verb move away, which, according to the classification of volition adopted in the present work, is a volitional rather than a non-volitional verb of feeling or thinking.
Moreover, the phrase following the main verb in this example functions as an adverbial and not as a direct object. However, from a semantic point of view, this and similar sentences also express effective stance.

(91)  **EO:** We [need to]**modal** move away from the remnants of a relationship that was one of ideological conflict and hostility with the Soviet Union, where our relations were adversarial, and our main concern was to contain the imperial tendencies of communist ideology. (EO_SPEECH_003 s20)

Coming back to the frequency distribution of the analyzed pattern, it can be observed that the construction \([NP\text{country/government} \ V\text{feel/think} \ \text{clause/NP}]\) is used more often in the German originals within the register of prepared speeches than of political essays. One explanation could be that the difference in the occurrences of the construction is not due to contrastive but rather register characteristics. In other words, the pattern could be less typical of the register of political essays in German than in the comparable register in English. Alternatively, the selection of texts for the comparable parts of this sub-corpus could have played a role. A closer look at the metadata reveals that the majority of political essays included in the English part was written by politicians, for instance by the British Prime Minister, US Secretary, Deputy Secretary or Undersecretary of State. In contrast, the authors of the German political essays are mainly journalists or economy scientists. Therefore, it could be assumed that the politicians use the construction expressing stance and potentially implying shared responsibility more often than journalists, who are expected to stay more neutral.

Having characterized the construction \([NP\text{country/government} \ V\text{feel/think} \ \text{clause/NP}]\) in the originals, we can now proceed with the analysis of its translations into German within the register of political essays. As indicated in Figure 14, the investigated pattern occurs more frequently in the data without construction shifts. However, if we consider the two variants of the construction used to communicate effective and epistemic stance, their distribution is to some extent different. The former tends to be kept in translations, whereas the latter appears to be changed more frequently. However, further studies are necessary to confirm these tendencies\(^{31}\).

\(^{31}\) Due to low values, the chi-square test has not been calculated.
Figure 14: [NP, country/government V feel/think clause/NP] - With and without construction shifts in E-G

Three instances classified as construction shifts and one from the group without shifts are characterized by changes of the original subject realized through the pronoun we. For instance, in (92), both the subject and the verb have been translated using the adverbial unserer Meinung nach ‘to our mind’. As a result, the translator has reduced the number of clauses in the target sentence while keeping the information structure of the original. Since the subject of the original subordinate clause is upgraded to the subject of the main clause, more emphasis is put on the expressed proposition rather than the personal reference and the expression of the shared responsibility.

(92) **EO:** [We] subj believe a modern, long term, and deep seated pro-European consensus in Britain about Britain's role in Europe and Europe's role in the world can and will be built. (EO_ESSAY_005 s72)

**GTrans:** [Unserer Meinung nach] adv kann und wird in Großbritannien ein moderner, dauerhafter und tief verwurzelter proeuropäischer Konsens über Großbritanniens Rolle in Europa und Europas Rolle in der Welt herbeigeführt werden.

(93) **EO:** We hope [the first two steps can be [completed] verb promptly, so that we can conclude the third before Congress adjourns this summer.] dobj (EO_ESSAY_019 s26)

**GTrans:** Wir hoffen [auf einen schnellen [Abschluß] noun der ersten beiden Schritte, so daß wir den dritten zu Ende bringen können, bevor der Kongreß in die Sommerpause geht.] probj

However, more instances of the analyzed pattern with the pronoun we are translated using a comparable German construction or involve changes of other elements, as can be seen in (93). In this example the crossing line has been created between the original direct object and the German prepositional object. By nominalizing the verb complete, the number of clauses could be again reduced. Using the SFL terminology, this is an instance of a grammatical metaphor as the clause referring to a process is compressed into a noun phrase.
Similar to the previous example and most of the construction shifts from the construction discussed in the previous chapter, the information structure remains unchanged. Moreover, apart from (92), two further construction shifts are caused by the verb believe, which is translated by the phrases der Meinung sein ‘in somebody’s opinion’ and überzeugt sein ‘be certain’, thus changing the [Subj V Dojb] into the [Subj V Compl] construction but preserving the main functional properties of the original [NP country/government V feel/think clause/NP] pattern, namely expression of stance in combination with shared responsibility.

To sum up, in a few cases the translations change the focus from the referent expressing stance to the proposition itself and, therefore, could be said to follow the tendency in German to emphasize content rather than persons (House 1997, 84). In the majority of the examples, however, the shifts should be explained through other factors, such as reduction of clauses and/or lexical preferences. In a future study different formal variants associated with similar functional features could be analyzed more closely to identify the exact conditions that favor the use of these structures.

In the German translations, 15 instances of the analyzed construction have been identified. The decision whether a German word belongs to the semantic groups of feeling and thinking was based on the senses present in GermaNet: for instance, the verb anstreben/streben ‘seek’ is classified as a hyponym of the verbs wollen/verlangen ‘require’, which are in turn hyponyms of the verbs wünschen/wollen ‘want’ (GermaNet). Anstreben is, therefore, classified as one of the verbs of feeling and is counted among the devices used to express effective stance (Marín Arrese 2011, 202). From the 15 examples of the investigated pattern, nine (60%) are translations of the corresponding English construction, whereas in one further example the original structure contains an additional grammatical function. In two other cases the pattern [Subj V Clause_{non-fin}] is present in the English originals: in both examples the original structure is changed into the construction [NP country/government V feel/think clause/NP] through a nominalization, as in (94). Moreover, the verb brauchen ‘need’ in the German construction can correspond to the marginal modal need to (see above). In the examples with construction shifts one analyzed structure embedded into a subordinate clause has been upgraded to the level of the main clause in the German translation, and one example has been introduced through various changes on the level of grammatical functions and a shift from passive to active voice. Thus, in the majority of the examples the construction
under analysis in the German translations is triggered by the corresponding English construction or a similar pattern.

(94) **EO:** We will seek \([\text{to}\ [\text{negotiate}]_{\text{verb}}\ \text{regional and bilateral agreements to open markets around the world.]}_{\text{Clause\_non-fin}}\) **EO\_ESSAY\_015 s71**

**GTrans:** Wir werden \([\text{die}\ [\text{Aushandlung}]_{\text{noun}}\ \text{regionaler und bilateraler Abkommen zur Öffnung der Märkte auf der ganzen Welt]}_{\text{NP}}\) anstreben.

No examples of the analyzed construction have been identified in the English translations belonging to the registers of political essays and popular scientific texts. In contrast, as discussed above, the construction occurs in the English translations in the register of prepared speeches.

As mentioned earlier, I assume that more instances of the construction \([\text{NP}_\text{country/government} \ \text{V}_\text{feel/think} \ \text{clause/NP}]\) could be found in the German political essays written by or on behalf of the authors, who identify themselves more with the government. More research using larger corpora containing the same register are necessary to confirm or reject this assumption. The results for the register of political essays combined with the findings from the Europarl corpus and CROCO sub-corpus SPEECH suggest that in political discourse in general the construction does not occur more frequently in English or German and, therefore, cannot be said to be more entrenched in the speakers’ minds using one of these languages. As a consequence, the translations cannot be analyzed in terms of the translation properties of shining through and/or normalization.

The present sub-chapter has shown that the difference in the occurrences of the transitive pattern \([\text{Sub}\_\text{animate} \ \text{V}_{\text{non-volition}} \ \text{Dobj}]\) in the two languages belonging to the two registers under analysis depends on the more specific construction \([\text{NP}_\text{country/government} \ \text{V}_\text{feel/think} \ \text{clause/NP}]\) associated with political discourse. The construction is used to express stance and attribute it to a group of people constituting either the Government or the people of the country. The different distribution of this construction in the originals constituting our main dataset could be explained by register characteristics or a selection of texts in the comparable sub-corpora ESSAY. The shifts from this construction in the direction English-German have been often motivated by a reduction of clauses and local lexical choices.
5.1.3 Subhuman Vvolition Dobj

As mentioned above, the transitive pattern [Subhuman Vvolition Dobj], in contrast to the patterns analyzed in the previous chapters, is positively attracted to the German originals. Therefore, we start our discussion by examining the structure in this group of texts.

In the German originals 36 instances of the analyzed pattern have been identified. Of these, 83% (30/36) are characterized by a human subject followed by a verb of speaking. This group of sentences could be classified as quotations. Literature review has shown that quotations typically involve a combination of NPhuman and V_speak followed by a clause referring to a reported proposition (Cappelen and Lepore 1997, 429-430; Brendel, Meibauer and Steinbach 2007, 6; Spronck 2012, 111). Thus, it could be assumed the pattern [NPhuman V_speak Clause], illustrated through (95) and (96), to be a prototypical representative of quotations.

(95) **GO:** ["Doch das weltweite Datennetz kann den direkten Kontakt nicht ersetzen",]Clause [sagt]V_speak [Peter Neven, Geschäftsführer Inlandsmessen beim AUMA.]NP_human (GO_ESSAY_011 s90)
(96) **GO:** [Er]NP_human [meint]V_speak [wenn die Auseinandersetzung mit dem Vater gelingt und damit die Idealisierung des Vaters aufhört, könne der Sohn, vom Vater geachtet, seinen Weg weitergehen.]Clause (GO_POPSCI_006 s72)

The individual slots include the source of the proposition, i.e. the reported speaker, a verb of speaking such as sagen/say, an optional conjunction dass/that and a reported message. The pattern itself is produced by a current speaker (or writer) (Spronck 2012, 76-77, 91). The function of this structure is to refer to a message attributing it (typically) to another speaker irrespective of the current speaker’s attitude to this proposition (Spronck 2012, 72). The reported message is sometimes indicated by quotation marks as in (95) but these are present only in certain types of quotations, and, therefore, cannot be regarded as a reliable indicator of this pattern (Brendel, Meibauer and Steinbach 2007, 6). The same could be said about a subjunctive verb form in German, such as könne ‘could’ in (96). The presence of the two indicators depends to some extent on the distinction between direct and indirect speech. Quotation marks are typical of direct speech, which is supposed to recreate the reported utterance in its original form (Quirk et al. 1985, 1021). The German subjunctive forms (Konjunktiv I, II as well as an alternative involving a combination of würde ‘would’ and an infinitive verb form) mark indirect speech (Brendel, Meibauer and Steinbach 2007, 10), which paraphrases the reported message (Quirk et al. 1985, 1021). However, the subjunctive
is often substituted by the indicative mood, especially in spoken discourse (Brendel, Meibauer and Steinbach 2007, 10). Furthermore, direct speech in both languages can be marked through punctuation: while in English the reporting clause is usually separated by a comma from the direct speech (Quirk et al. 1985, 1023, 1631), in German a colon is more typical (Brendel, Meibauer and Steinbach 2007, 9). However, the use of punctuation and quotation marks may point in different directions: punctuation typical of direct speech may be followed by a clause that is not enclosed into quotation marks. In the present study, I use the quotation marks to distinguish between direct and indirect speech.

It should also be mentioned that direct and indirect speech are just two types of the quotations distinguished in the previous research. For instance, mixed quotations are named among further variants. This pattern, shown in (97), occurs twice in our data. In this example we could observe a mixture of indirect speech characterized by the use of the subjunctive mood as its optional formal marker and quotation marks around the expression *Apotheke der Welt* 'the world’s pharmacy' attributing these words to the original speakers (Cappelen and Lepore 1997, 429). Moreover, a quotation variant illustrated through (98) is sometimes referred to as a parenthetical type. Here the reported message is interrupted through a verb of speaking and the reported speaker (Brendel, Meibauer and Steinbach 2007, 6). Brendel, Meibauer and Steinbach (2007, 10) suggest that subjunctive forms are associated also with this quotation type, thereby indicating that it could be seen as a sub-type of indirect speech. However, in our data, the three examples of such parenthetical quotations, in which the direct object is split and surrounds the main verb and the subject, contain the reported message enclosed in quotation marks, as can be seen in (99). Therefore, these examples could be interpreted as sub-type of direct speech.

(97) **GO:** Die einstige "Apotheke der Welt" rangiere nur noch im Mittelfeld, klagen [Experten]$_{NP\_human}$ (GO_ESSAY_004 s7)

(98) **GO:** [Jahre,] erzählt Reinhard Hüttl, Professor für Bodenkunde an der TU Cottbus, [habe man aber bei der Stadtverwaltung darum kämpfen müssen, daß die Universität auf Straßenrändern auftauche.]$_{dobj}$ (GO_ESSAY_015 s51)

(99) **GO:** ["Wir müssen",] sagt Dr. Wilfried Weiß, Beauftragter für Chancengleichheit bei der Deutschen Lufthansa, ["unser Kapital, die Mitarbeiter und ihre Produktivität, hegen und pflegen".]$_{dobj}$ (GO_ESSAY_023 s86)
Additionally, the present study considers the pattern [NPhuman V_speak NP/Pronoun] among quotations, even though these might not belong to the prototypical instances of this pattern. Moreover, the example in (100), similar to (97), is less typical of the category under analysis because in both cases the reported proposition is attributed not to a single definite speaker but to a group. In (97) the plural *Experten* ‘experts’ makes it less likely that the current speaker reproduces the original utterance of the reported speakers. Instead, it could be said that the current speaker summarizes the general topic of the proposition produced by different speakers. Furthermore, the pragmatic level of the reported message is reproduced through the use of the verb *klagen* ‘complain’ (Smirnova 2012, 243-244). The sentence presented in (100) reports a result of a quantitative study, and it could be assumed that the statement expressed in the subordinate clause is a pre-formulated answer possibility used in a survey. Therefore, the reported speakers, namely *58 Prozent der deutschen Unternehmer* ‘58 per cent of German entrepreneurs’, have probably not uttered the content presented in the clause but selected one of the available options. Moreover, one further example in our data deviates from a more common pattern with respect to the selected subject: rather than a typical subject of quotations in the form of the third person reference, in this example, shown in (101), the subject is the second person pronoun *Sie* ‘you’ used as a polite form of address.

Coming back to the functions of the described types of the quotations, it should be mentioned that depending on the argumentative goals of the current speaker, s/he may reproduce different linguistic levels of the reported message, such as phonetic, syntactic, lexical, semantic, pragmatic or rhetorical (Smirnova 2012, 240). For instance, Smirnova (2012, 247) suggests that the close recreation of the original semantic content, often using direct speech, is a typical feature of the quotations employed to support an argument of the current speaker. The current speaker might have subjective attitudes towards the reported message, which...
could be explicitly or implicitly present in the quotations. For instance, the subjective opinion of the current speaker could be transmitted through the choice of a particular verb signaling doubt or its modification by an adverb or a modal verb (Spronck 2012, 92). However, it should be mentioned that it is not always possible to separate the attitudes of the reported and the current speaker, in particular by analyzing the verb used in the construction (Smirnova 2012, 244).

In our data sample, the quotations occur more often in the register of political essays than in the register of popular-scientific texts. These two groups of texts contain 26 and 4 examples respectively. Moreover, the former register is characterized by a higher variation of quotation types, involving indirect and direct speech together with mixed, parenthetical and different less prototypical types of quotations discussed above, while the latter contains only instances of indirect speech as well as less prototypical instances, namely example (101) and two instances of $\text{NP}_{\text{human}} \text{V}_{\text{human}} \text{NP}$.

The use of quotations in the register belonging to political discourse could be linked to the functions of legitimization and coercion (Chilton and Schäffner 1997, 212-213), discussed in Chapter 5.1.2. Quoting others who have the same opinion as the current speaker can enhance the force of the argument (Smirnova 2012, 247, 250). This could explain why the current speaker’s attitude towards the reported message is not expressed explicitly: as stated by Spronck (2012, 84), the lack of explicit evaluation may imply the objectivity or agreement with the reported speaker. In particular the latter is the case, if the current speaker employs the quotation strategically to support her/his own agenda. The noun referring to the reported speaker is typically modified. This modification could be employed to specify the title or position of the author, thus stressing the authority of the reported speaker and further intensifying the argument. This could be observed in examples (95), (98) and (99). A similar effect is achieved in (97) through the use of the noun *Experten* ‘experts’.

Hyland (2010, 122) suggests that in popular-scientific texts it is common to cite studies of others in the form of quotations in order to increase credibility. Since in our data sample quotations realized through the abstract construction $\text{[Subj V Dobj]}$ are mainly associated with the register of political essays, it is assumed that other patterns are employed for this purpose in popular-scientific discourse. For instance, the reference to scientists or their investigations can be integrated into other grammatical functions. In (102) both the reported
speaker and the action realized through the verb of speaking *betonen* ‘stress’ are expressed through a sentence initial adverbial.

(102) **GO:** [Wie VERENA KAST zu Recht betont]_adv_ handelte es sich bei diesen Menschen nicht um unreife Personen, die sich in symbiotischen Beziehungen zu ihren Partnern aufgegeben und verloren haben, vielmehr aktualisierte sich das Selbsterleben dieser Personen im gemeinsamen Beziehungsraum, im miteinander gebildeten System. *(GO_POPSCI_004 s40)*

Could the analyzed pattern be considered a form-meaning pairing, i.e. a construction in its own right? The structure with the form [NP\text{human} V\text{speech} Clause/NP] is associated with specific functions: in the register of political essays, similar to other types of argumentative discourse, the pattern can be employed to strengthen the argument of the current speaker by attributing the reported message to an authoritative source. The NP and V slots are characterized by certain restrictions: the subject referring to one or several human beings is combined with the reporting verb. It should be mentioned that while our analysis of this construction, based on the selected categories of animacy and volition described above, considers volitional verbs, namely verbs of speaking, also verbs of thinking could be used in quotations (Quirk et al. 1985, 1024). Moreover, the subject may refer to a collective group of individuals, i.e. an organization, classified as an animate entity. Still, even taking into account these elements, which are excluded from the discussion in this part of the study, the range of lexical material that could fill two of the three slots of the pattern is limited. In particular, the V slot can be described through two semantic groups, namely verbs of speaking and thinking. Moreover, I have mentioned several additional conventional (even though in some contexts optional) indicators of the structure, namely the quotation marks and the subjunctive verb forms in German. Considering these both formal and functional features, quotation, limited in our study to its variant [NP\text{human} V\text{speech} Clause/NP], could be postulated as a more specific transitive construction. This is in line with the study by Sprock (2012, 73) that referred to the structure as a quotative construction, even though the reasons for the postulation of this construction were not explicitly discussed.

Furthermore, a previous study (Stefanowitch 2008) has suggested the existence of the independent direct speech construction. Stefanowitch discusses a number of idiosyncratic features of the pattern, such as several unmarked word orders and the syntactic behavior of the grammatical element referring to the reported message. In other words, the study
suggests that since it is only marginally acceptable to passivize this grammatical unit and because it can occur in the sentence-initial position without fulfilling the pragmatic conditions typically associated with this word order, the function of the reported message is not necessarily a direct object. Therefore, Stefanowitsch (2008, 252-257) argues against classifying this pattern, which he considers to be a construction in its own right, as a variant of the transitive structure. However, the postulation of a more specific construction with lexical restrictions on its slots and further idiosyncratic properties does not contradict its relation to the more general transitive construction. As discussed above, the constructicon is usually assumed to be characterized by certain redundancy, so that the patterns could be stored on different levels of complexity and abstraction (Goldberg 2006, 55). In a similar way, also lexically fixed idioms (which could have rather idiosyncratic syntactic and semantic features) have been considered to be variants of such abstract constructions as verb phrases (Croft and Cruse 2004, 263, see also Chapter 3.1.2). Treating the direct object construction not in isolation but in relation to other quotative patterns and the more abstract transitive construction allows us to account for their similarities and not only for properties specific for one particular structure.

Similar to the constructions discussed above, the quotative construction has been operationalized by verifying what instances of the transitive pattern [Subj human V volition Dobj] involve a verb of speaking. During this step, GermaNet or WordNet for German and English data respectively has been consulted. The examples of the analyzed constructions belonging to the German and English originals and translations in our data sample contain the following lemmas filling the V slot:

Argumentieren, behaupten, bemerken, bestätigen, bezeugen, bezweifeln, erklären, erwähnen, erzählen, feststellen, fragen, hervorheben, klagen, meinen, sagen, unterstreichen, urteilen, verlangen, versprechen, vorhersagen, zitieren

Argue, ask, complain, confirm, declare, emphasize, explain, maintain, notice, point out, postulate, present, quote, say, state, suggest, underscore

Since only the V slot can be described in terms of a set of concrete lexical items, only one collostructional analysis was performed. Due to the occurrences of the construction in both registers under analysis, the joint number of sentences in the sub-corpora ESSAY and POPSCI within the German originals have been taken into account for the calculation of the relevant
contingency tables. Similar to the collostructional analyses reported in the previous chapters, the instances of the construction in the originals but not in the translations are considered. The analysis shows that all of the verbs occurring in this construction in the German originals are significantly attracted to the construction's V slot. As can be seen in Table 17, the verb sagen ‘say’ has the highest collostructional strength: the verb is the most frequent in the two analyzed registers and occurs in 53% (16/30) of the identified instances of quotations. The verb could be assumed to be among the most neutral verbs of speaking that does not signal additional evaluation of the reported message.

<table>
<thead>
<tr>
<th>Collexeme_V</th>
<th>Freq. in the corpus</th>
<th>Freq. in the constr.</th>
<th>Coll. strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>sagen ‘say’</td>
<td>58</td>
<td>16</td>
<td>21.02</td>
</tr>
<tr>
<td>zitieren ‘cite’</td>
<td>2</td>
<td>2</td>
<td>4.10</td>
</tr>
<tr>
<td>feststellen ‘state’</td>
<td>6</td>
<td>2</td>
<td>2.94</td>
</tr>
<tr>
<td>erklären ‘explain’</td>
<td>9</td>
<td>2</td>
<td>2.56</td>
</tr>
<tr>
<td>meinen ‘say’</td>
<td>12</td>
<td>2</td>
<td>2.31</td>
</tr>
<tr>
<td>klagen ‘complain’</td>
<td>2</td>
<td>1</td>
<td>1.75</td>
</tr>
<tr>
<td>erzählen ‘tell’</td>
<td>2</td>
<td>1</td>
<td>1.75</td>
</tr>
<tr>
<td>vorhersagen ‘predict’</td>
<td>3</td>
<td>1</td>
<td>1.57</td>
</tr>
<tr>
<td>urteilen ‘judge’</td>
<td>3</td>
<td>1</td>
<td>1.57</td>
</tr>
<tr>
<td>fragen ‘ask’</td>
<td>5</td>
<td>1</td>
<td>1.35</td>
</tr>
<tr>
<td>versprechen ‘promise’</td>
<td>5</td>
<td>1</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Table 17: [NP<sub>human</sub> V<sub>speak</sub> clause/NP] - Significantly attracted collexemes for V<sub>speak</sub>

Figure 15 shows that the German originals are characterized by a high number of direct quotations, followed by indirect speech and a few examples of mixed and parenthetical quotative patterns. In contrast, the English sample not only contains significantly less instances of the construction [NP<sub>human</sub> V<sub>speak</sub> Clause/NP] ($\chi^2=18.3$, df=1, p-value=1.888e-05, see Table 18), but also other quotation types. In the English originals, it was not possible to identify any instances of direct and parenthetical quotations within the transitive pattern [Subj<sub>human</sub> V<sub>volition</sub> Dobj]. Instead, 82% (9/11) of the analyzed quotative constructions could be classified as examples of indirect speech, whereas in German this pattern accounts only for 27% (8/30) of the data. Also the distribution of the construction across registers is
somewhat different, since in English the small number of instances is almost equally divided across the two registers under analysis. But, similar to German, both registers in English also contain quotations realized through different sets of grammatical functions. In example (103) the reported speaker and the verb of speaking are both integrated into an adverbial rather than realized through the subject and the main verb of the sentence.

![Figure 15: [NP$_{human}$ V$_{speak \ clause}$/NP] - With and without construction shifts in E-G and G-E](image)

<table>
<thead>
<tr>
<th></th>
<th>NP$<em>{human}$ V$</em>{speak \ clause}$/NP</th>
<th>Subj V Dobj$_{other}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>11</td>
<td>239</td>
</tr>
<tr>
<td>German</td>
<td>30</td>
<td>142</td>
</tr>
</tbody>
</table>

Table 18: [NP$_{human}$ V$_{speak \ clause}$/NP] - Contingency table for construction & language

(103) EO: [As former president Ernesto Zedillo of Mexico said,]$_{adv}$ some supposed friends of the downtrodden "seem strangely determined to save the developing world from development." (EO_ESSAY_018 s46)

The distribution of the direct speech in English and German differs from the results reported by Stefanowitsch (2008, 257-258), where the comparison of absolute numbers of occurrences in the corpora containing one million words indicated that the structure occurs more frequently in English than in German. The difference could be potentially explained through additional limitations of the analyzed set of registers and the focus on volitional verbs in the present study.

Due to a small number of examples belonging to the English originals, collostructional analysis has not been performed. The only verbs that occur more than once in this group of texts are *argue* (3 occurrences) and *say* (2 occurrences). With respect to the subject, it should be mentioned that there is more variation in the English than in the German originals. First
of all, in the English data the majority of examples, namely 55% (6/11), contain personal pronouns as subjects, in contrast to German where 67% (20/30) of subjects are realized through a proper name, often in combination with various modifiers. This difference could be partly explained through different proportions of the quotation types found in the two languages: proper names often occur in direct speech, i.e. the quotation type, which is not present in the English originals. Additionally, in English two instances of the analyzed construction involve first person pronouns as subjects. It is questionable whether the example in (104) could still be considered a quotation because not only do the current and the reported speaker coincide, but the main verb is modified by a semi-modal *going to*, thereby indicating a future reference. Following Brendel, Meibauer and Steinbach (2007), where for instance, example (105) with the pronoun *ich* ‘I’ as the subject has been described as a quotation, I include the sentence shown in (104) in the group of quotatives. However, these instances are considered less prototypical variants of the quotative construction, at least with respect to the two registers under analysis. Based on the formal criterion of the quotation marks, example (104) has been classified as indirect speech. However, it could be argued that a combination of first person pronoun and present tense with a future reference is ambiguous in this respect, as the speaker might either repeat her/his argument or utter it for the first time. This ambiguity is not limited to the sentences with the first person pronouns functioning as subjects.

(104) **EO:** I’m going to argue that this is simply wrong. (EO_POPSCI_008 s32)
(105) Hiermit sage ich, dass die Theorie schwer zu verstehen ist. (Brendel, Meibauer and Steinbach 2007, 13)

Since the construction, similar to the one discussed in the previous chapter, occurs more frequently in the register of political essays, it is interesting to compare the results to a further register belonging to the field of political discourse. Therefore, in the next step, the CroCo sub-corpus of prepared speeches (SPEECH) has been queried for the verbs *sagen* ‘say’ as the verb with the highest collosstructional strength with relation to quotations, and its English equivalent *say*. The queries within the English and German originals returned similar amounts of hits, namely 46 and 40 sentences respectively. In the German data 29 instances contained the three required elements, namely the subject realized through a NP referring to a human, the verb *sagen* and a direct object. However, these 29 examples contained only two
that could be classified as the pattern discussed here, whereas further 79% (23/29) of examples include additional grammatical functions. Moreover, four instances involve several main clauses. In the English originals, of the total amount of 46 extracted sentences with the lemma *say*, 25 involved the relevant grammatical functions. Of these, eleven examples (44%) represent the variant under investigation, seven involved additional functions (28%) and in the remaining seven instances (28%), the construction occurs in one of the two main clauses or on a lower level integrated into another grammatical function. Comparing all instances of the pattern across the two languages, i.e. irrespective of the fact whether it involves additional functions or further main clauses, whether the pattern occurs on a lower level in the structure of the main clause or is a true hit, we could observe that in English, in contrast to German, there are more sentences with the third person reference present in subjects. Moreover, in the majority of cases the direct object corresponds to a clause.

![Figure 16: Frequency distribution in BNC and DWDS](image)

To obtain additional contrastive data not limited to particular registers, the same query for the lemmas *say* and *sagen* has been repeated in the reference corpora BNC and DWDS. The overall number of hits reached 313,865 in English and 203,650 in German. In the next step, random samples of 200 per corpus were extracted and further analyzed. Figure 16 shows that of these 200 examples, 41 and 47 in English and German data respectively did not contain all three elements under analysis and were excluded from further discussion. The remaining 159 sentences belonging to the English originals contained 84 instances (53%) of

---

32 While other contrastive analyses of additional corpora consider only the true hits, the present comparison involves all examples with the elements N_human, V_speak and Clause/NP. This is due to the fact that there are only two true hits in the German data. While the question whether the pattern with additional elements is heterogeneous and could correspond to other constructions is not investigated in more detail, the general tendency to use different subjects and direct objects in this register deserves further attention.
the analyzed construction [NP_{human} V_{speak} Clause/NP], in contrast to only 31 of the remaining 153 examples in the German originals (20%). Based on these random samples, it is possible to conclude that, in contrast to the CroCo registers of political essays and popular-scientific texts, the construction under investigation is more frequent in English than in German ($\chi^2$=33.00, df=1, p-value=9.212e-09, see Table 19). Similar to the results from the CroCo sub-corpus SPEECH, German contains a higher proportion of structures with additional grammatical elements.

<table>
<thead>
<tr>
<th></th>
<th>NP_{human} V_{speak} clause/NP</th>
<th>Other sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>84</td>
<td>116</td>
</tr>
<tr>
<td>German</td>
<td>31</td>
<td>169</td>
</tr>
</tbody>
</table>

Table 19: [NP_{human} V_{speak} clause/NP] - Contingency table for construction & language, BNC and DWDS

A closer look at the true hits reveals that 50% (42/84) of the English examples could be classified as direct speech quotations, whereas 43% (36/84) represent indirect speech. The other 7% of the data contain six instances of mixed and parenthetical quotatives, equally divided between the two groups. Also in the German data, there are more direct than indirect quotations, accounting for 42% (13/31) and 32% (10/31) of the sample respectively, but the category of parenthetical quotatives amounting to 23% (7/31) is larger than in English. As shown in

Figure 17 and Figure 18, in English the most frequent construction variant within the direct speech category is [Name V_{say}: ‘Clause’], while in German the variants [“Clause”, V_{sagen} 3rdperson Pro/NP], [“Clause”, V_{sagen} Name] and [3rdperson Pro/NP V_{sagen}: “Clause”] occur with similar frequency. In indirect speech [Name V_{say} Clause] and [3rd person Pro/NP V_{say} Clause] are particularly typical of English. However, due to low numbers, particularly in the German sample, these results should be seen as tendencies, which have to be further investigated in the future studies.
The results from the reference corpora indicate a contrastive difference between English and German with respect to the use of the construction [NPhuman V_speak clause/NP]. Based on the random samples of the lemmas *say* and *sagen*, the construction seems to appear more often in English than in German. While this result could indicate that the construction is more entrenched in the minds of the speakers of English, the different registers contained in the data should be taken into account. Our main sample consisting of the registers of political essays and popular-scientific texts reveals more quotatives in German than in English, whereas the register of prepared speeches has more instances of the construction in English. Moreover, it is interesting to observe that the 67% (56/84) of true hits in the BNC data sample belong to the register of newspaper language. In contrast, 68% (21/31) in the DWDS data correspond to the register of fiction. Thus, I suggest that the analyzed construction is
register-specific and, depending on the language, is more or less likely to occur in a certain register. Entrenchment predictions could, therefore, not be made for the whole language, independently from different registers.

Taking into account the distribution of examples in the originals belonging to the main sample of the CroCo corpus, translation shifts from the construction under analysis could be expected in the translation direction German-English. However, there are more examples without construction shifts in both translation directions. 26 sentences containing a quotation in the German originals (87%) are translated into English using the corresponding construction \([\text{NP}_{\text{human}} \ V_{\text{Speak}} \ \text{Clause/NP}]\), whereas in the two remaining examples without translation shifts the transitive construction is kept in translation but the pattern \([\text{NP}_{\text{human}} \ V_{\text{non-volition}} \ \text{Clause/NP}]\) is used.

In 27% of examples (7/26) translated by the same quotative construction, a different word order is used in the translation in comparison with the corresponding original. In (106) and three further examples, the reported clauses in the initial position are changed into final ones. However, (106) illustrates that this does not need to have any serious consequences for the text structure. While in the German original the reported speech continues the topic of the preceding sentence, in the translation the reported proposition is thematically linked to the following sentence.

(106) **GO:** Jeden Mittwoch verbringt der 44-jährige Controller mit seinen fünf, sieben und acht Jahre alten Kindern. [Der Trend der Mitarbeiter hin zu flexiblen Zeitmodellen sei unverkennbar,] reported_speech stellt Wilfried Weiß fest. Von dreißig Millionen abhängig Beschäftigten in Deutschland sind inzwischen 6,3 Teilzeitarbeiter. (GO_ESSAY_023 s92-94)

**ETrans:** The 44-year-old controller spends every Wednesday with his children aged five, seven and eight. Wilfried Weiss says [the trend the-GEN staff to flexible time.models is-SBJV unmistakable,] reported_speech says Wilfried Weiß. out of the 32,5 millions employees in Germany are meanwhile 6,3 part.time.workers.

**GO:** The 44-year-old controller spends every Wednesday with his children aged five, seven and eight. Wilfried Weiss says [the trend among staff in favour of more flexible working hours is unmistakable,] reported_speech Of the 32.5 million wage and salary earners in Germany, 6.3 million now work part-time.
Furthermore, word order changes involve two examples in which the variant [Clause $V_{\text{human}}$] is translated as [Clause $N_{\text{human}} V_{\text{human}}$], i.e. without the subject-verb inversion in English. This could be explained by a language-specific feature: inversion with a pronoun as the subject is considered archaic in English (Quirk et al. 1985, 1022). One further example involves a shift from a medial into a final position of the reporting clause, i.e. from a parenthetical direct into a direct speech construction. The attested examples of translation shifts do not involve major changes. In these cases a (parenthetical) direct speech construction is translated using additional grammatical functions, which might result in a sentence merge or make the translation more explicit, as in (107). Here the original construction is split into two main clauses introducing the direct speech clause through the prepositional phrase like this, functioning as an adverbial. This element can be seen as an additional, more explicit indicator of the following direct speech.

(107) **GO:** "Der Wirtschaft die geistige, die seelische und die materielle Ausrichtung zu geben", sagte Erhard, "das ist Sache der Politik, Angelegenheit der Gesellschaft."  
(GO_ESSAY_001 s69)  
**ETrans:** Ludwig Erhard put it like this: "It is the task of politics, a matter for society, to point the economy in the right intellectual, psychological and material direction."

It should be mentioned that of nine sentences with and without translation shifts, which are characterized by some minor changes of the analyzed construction, five occur in a translation of the same German text. Considering this together with the fact that no serious changes of the original construction have been identified in the data, it is suggested that for this translation direction the translators tend to keep the quotative construction in the target texts. If further analyses of political essays confirm that in this register the construction is more frequent in German than in English, it is possible to consider these examples as instances of shining through. Additionally, analysis of this construction in the English translations indicates that the analyzed structure is sometimes introduced in the translation without a corresponding construction in the aligned German originals. This is the case in seven examples corresponding to 22% of all occurrences of the analyzed construction in the English translations.

As mentioned above, the pattern under analysis also typically remains unchanged in the opposite translation direction. Only one example within the data without construction shifts has been classified differently, because, according to *GermaNet*, the verb *bestreiten* ‘deny'
used in the target sentence could not be classified specifically as a verb of speaking. The other eight examples in this data sample have not undergone any changes. The group of construction shifts includes only two examples of this construction. However, the changes involved are not as minor as in the translation direction German-English. In one case the analyzed transitive construction has been shifted to [Subj V Probj] in German, whereas in example (108) the quotative construction with the first person pronoun we functioning as the subject is turned into a more impersonal structure. This change of a sentence belonging to the register of popular-scientific writing can be explained through a contrastive difference between English and German mentioned above, according to which the English speakers turn to focus more on persons, whereas the German speakers are more likely to stress the facts (House 1997, 84). In this example, the adverbial wie Kontrollmessungen ergaben ‘as confirmatory measurements yielded’ does not name those responsible.

(108) **EO:** We confirmed that they had identified 85 to 90 percent of the wells correctly.  
(EO_POPSCI_002 s84)  
**GTrans:** Wie Kontrollmessungen ergaben, wurden 85 bis 90 Prozent der Wasserstellen korrekt bewertet.

Apart from the cases where the corresponding construction is also present in the originals, German target texts contain six additional instances of quotations (accounting for 43% of all instances of the construction in this group of texts), five within the category of construction shifts and one within the data without shifts from the [Subj V Dobj] construction. In these cases the analyzed construction in the translation corresponds to different grammatical patterns in the originals involving, for instance, prepositional objects. Thus, with respect to translations of the analyzed construction, the two translation directions are similar as both in translations into English and into German the pattern is often translated without any shifts. Moreover, some additional instances of the construction are introduced without the corresponding trigger in the originals. As a result, there are significant differences in the frequency of the pattern in English originals and translations ($\chi^2=12.04$, df=1, p-value=0.0005, see Table 20) as well as in German originals and translations ($\chi^2=7.66$, df=1, p-value=0.006, see Table 21).
In this chapter, it has been shown that 83% of German originals belonging to the transitive pattern [Subj human V volition Dobj] in the registers of popular-scientific discourse and political essays represent the more specific construction [NP human V speak Clause/NP]. This pattern could also be referred to as a quotation or a quotative construction. It attributes the reported proposition to another speaker, while it is often used to support the point of view of the current speaker. Its occurrences in the register of essays were linked to such general functions of political discourse as legitimization and coercion. While in our main data sample the construction is more frequent in German than in English originals, the other consulted corpora indicate that the occurrences of the construction depend on the register. In other words, the English and German registers might differ in the frequency with which they employ the construction under analysis. Both translation directions contain only a few examples of shifts from this pattern, but the word order and, as a result, information structure is changed in some examples in the direction German-English.

### 5.1.4 Subj non-ref V Dobj

As mentioned earlier, the transitive pattern [Subj non-ref V Dobj] occurs in German but not in the English originals. Therefore, it has been also selected for further analysis. The structure has nine instances in the German originals with construction shifts and five examples in the group without shifts. From these, 57% (8/14) correspond to the structure [Es gibt NP], all of which occur in the shift data. The other 43% (6/14) could be classified as [Man V Dobj].


**ETrans:** There is no plausible reason why a system that once facilitated "reform processes" should today only be capable of cementing the status quo.
As shown in (109), the structure \([Es \ gibt \ NP]\) involves the impersonal subject, realized through the pronoun \(es \ ‘it’\), the third person singular form of the verb \(geben \ ‘give’\) in present or past tense and a direct object. The object is realized through a NP, which often contains both pre- and postmodification in our data sample. For instance, in (109) the head noun Argument ‘argument’ is premodified by an indefinite pronoun together with an adjective and postmodified by a full relative clause. This type of postmodification is present in all examples of this pattern present in our data. Functionally, the structure can be characterized as one of the German existential constructions, as it can be used to state the (absence of) existence (Joseph 2000, 187; Pfenninger 2009, 294; Weinert 2013, 43), in this case of a plausible argument. Moreover, the structure could serve the pragmatic function of foregrounding information (Pfenninger 2009, 300; Weinert 2013, 43) and is sometimes argued to introduce new entities into discourse (Weinert 2013, 43), even though it is not the case for all variants of the German pattern (Pfenninger 2009, 300). Weinert (2013), among other authors, refers to the structure as a construction. Indeed, since the first two slots are lexically fixed, the pattern could be classified as a lexically-filled construction in the framework of CxG\textsuperscript{33}.

In our data, all instances of this German construction are translated into English by the structure \([There \ V_{be} \ NP]\), which has similar functions to the German \([Es \ gibt \ NP]\) construction (Lambrecht 1994, 179-181; Biber et al. 1999, 951), namely to introduce new information or to put additional emphasis on this semantic content. Lambrecht (1994, 181) suggests that an entity introduced through this English pattern does not function as a participant in an event, but is “merely made available for presentation in subsequent clauses by being raised into the addressee’s consciousness”. In English this structure is also considered to represent a partly lexically-filled construction with the fixed item there followed by a finite verb, typically the verb be, though a limited set of other verbs is also allowed in this position (Mauranen 1999, 60). Previous contrastive accounts of existential structures (Zydatiß 1981; Pfenninger 2009) have drawn certain parallels between \([Es \ gibt \ NP]\) and \([There \ V_{be} \ NP]\). Since both patterns have comparable functions and a formal similarity, namely a marker of existence in the theme

\textsuperscript{33} The verb slot does show some variation: for instance, the slot could also be filled by the verb sein ‘be’ but its use is much less frequent (Pfenninger 2009, 293). Pfenninger also suggests that the use of the verb haben ‘have’ in this function is limited to colloquial contexts and some dialects, e.g. Swiss German (2009, 296).
position\textsuperscript{34}, in the present study they are considered corresponding constructions (Leino 2010, 131). From this perspective, in all translations of the German [\textit{Es gibt NP}] construction, a corresponding English construction is selected, even though the German and English patterns display some functional dissimilarities (Pfenninger 2009, 294) and formal differences in the types of grammatical functions. The latter is the reason why these pairs of originals and translations have been classified as translation shifts. As mentioned in Footnote 34, the existential marker \textit{there} is annotated as the adverbial, while the NP following the finite verb is tagged as the subject in the CroCo corpus. Alternatively, it is possible to consider the initial element \textit{there} to function as the subject, similar to the German existential marker \textit{es}. According to this analysis, the post-verbal NP in the English construction could be classified as the complement. Therefore, in both cases the translations of the construction [\textit{Es gibt NP}] by the pattern [\textit{There V be NP}] involve crossing lines and different grammatical functions.

A number of translation and contrastive studies comparing the English existential construction with comparable structures in e.g. Norwegian (Ebeling 1998) or French (Cappelle and Loock 2013) have indicated that translators occasionally use other grammatical structures to translate one of the corresponding constructions. Both in translations into Norwegian and Finnish, these alternative structures typically preserve the information structure of the original clauses, i.e. they introduce a new discourse element after the finite verb (Ebeling 1998, Mauranen 1999). Among the possible reasons for not using a comparable pattern Ebeling (1998, 190-191) names a process of explicitation, problems finding an equivalent for a NP integrated into the original construction, complexity of a NP and preferences of individual translators. Furthermore, the corresponding constructions may be entrenched to different degrees in the languages under analysis, for instance in English and French (Cappelle and Loock 2013).

To examine the respective use of existential constructions in English and German, I consider random samples of 200 examples extracted from the reference corpora BNC and DWDS. The English query contained a lemma of the existential marker \textit{there} directly followed by the lemma \textit{be}, whereas in German the form \textit{es} followed by the lemma \textit{geben} have been searched.

\textsuperscript{34}The existential marker \textit{es} in the German construction has been described as a dummy subject (Weinert 2013, 68). A number of studies consider also the corresponding marker of existence \textit{there} as a grammatical subject (Lakoff 1987, 548; Davidse 1992; Biber et al. 1999, 944; Halliday and Matthiessen 2013, 308-309), even though a certain connection between the deictic and existential uses of \textit{there} is suggested by some scholars (e.g. Lumsden 1988, 227-229; Lambrecht 1994, 179). In the CroCo corpus the existential \textit{there} has been tagged as an adverbial of place.
for. Due to more flexible word order in German, both word order sequences have been taken into account.\(^{35}\) The fixed order of lexical items extracted from the English corpus excludes examples in the interrogative mood.

\[^{35}\text{Similar to the BNC and DWDS queries in Chapters 5.1.1 and 5.1.3, the distance between the two lemmas has been initially set to four words, meaning that up to four words could occur e.g. between there and be. However, due to technical reasons connected to the Davies interface of the BNC, it was not possible to extract a random sample of these examples. Therefore, in the present chapter the queries in both reference corpora had to be modified, so that the two elements immediately followed each other.}\]

**Table 22: [There/Es is/gibt NP] - Contingency table for construction & language, BNC and DWDS**

<table>
<thead>
<tr>
<th></th>
<th>[There/Es is/gibt NP]</th>
<th>Other sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>34</td>
<td>166</td>
</tr>
<tr>
<td>German</td>
<td>21</td>
<td>179</td>
</tr>
</tbody>
</table>

The queries described above returned 179 563 hits for English (or 3% of 6026284 sentence-units in the BNC) and 45 594 for German (or 1% of 5171328 sentence-units in the DWDS). In other words, in German both word orders of the elements es and gibt occurred three times less often than in English. While these values include false hits as well as other variants of the pattern, e.g. involving additional grammatical functions or further main clauses, it could be assumed that the construction is more frequent in English. As can be seen in Figure 19, in the random samples of 200 the number of true hits is somewhat larger in English, the p-value being slightly above the conventional significance level of 0.05 ($\chi^2=3.04$, df=1, p-value=0.08, see Table 22). In both languages, there are more examples of the analyzed items used with additional grammatical functions. This finding is in line with the previous literature, where it is suggested that the pattern is seldomly used to state mere existence (Biber et al. 1999, 947).
Instead it is more common to relate this existence to a particular location or time, or to combine it with a certain condition or personal assessment. 47% (16/34) of the instances in the BNC occur in the registers of popular-scientific and academic writings. In German, there is a relatively similar distribution of true hits between the registers of instruction manuals, academic prose, fiction and newspaper language, with only slightly more instances in the first of these groups.

Since analyses of the main data sample from the CroCo corpus concentrated on the instances of the general pattern [Subj V Dobj], the distribution of the English pattern [There V be NP] was examined only on the basis of the random sample from BNC. Therefore, it is not possible to compare the contrastive results obtained from the reference corpora to the CroCo corpus. The German data, both from CroCo and DWDS, shows that the German construction [Es gibt NP] is not predominantly used in any of the registers under analysis but is rather evenly distributed across different registers.

In the German translations included in the CroCo corpus, the pattern [Subj non-ref V Dobj] occurs 14 times, of which ten instances correspond to the construction [Es gibt NP]. Thus, both in the German originals and translations, the construction accounts for 5% (8/172 and 10/190) of the transitive pattern [Subj V Dobj]. In six cases, the construction is a translation of the corresponding English construction [There is NP]. Moreover, in four translation pairs the analyzed pattern has been triggered by different linguistic structures. For instance, in two cases of the group without construction shifts, the pattern corresponds to the combination [Subj inanimate V volition Dobj]. Here, the inanimate subject of the original sentence is turned into the direct object.

To conclude, I argue that, due to a number of formal and functional similarities, the German [Es gibt NP] and the English [There V be NP] could be considered corresponding constructions. The analysis of random samples extracted from the reference corpora has suggested that the English pattern is more frequent as it allows more variation in terms of the grammatical categories expressed by the verb. In both languages, the constructions typically occur with additional constituents.

The pattern [Man V Dobj] belongs to the impersonal patterns with generic reference. While one of the elements, namely the subject, is relatively fixed being filled by a limited group of impersonal pronouns (cf. van der Auwera et al. 2012), no tendencies with regard to the V slot
could be identified. It is interesting to note that five of six occurrences of the pattern in the German originals belong to the register of popular-scientific writing. Moreover, two occurrences of the combination \([\text{Man nehme } \text{NP}_{\text{concrete}}\]) suggest that this pattern could be common in instructional texts, such as instruction manuals and recipes. The pattern is translated into English using the impersonal pronouns \(it\) and \(one\), the pronoun \(we\) or the imperative forms. As there are only six occurrences of this pattern in our data and only the first slot appears to be relatively fixed, the pattern is not discussed here in more detail.

5.1.5 \textbf{Subj V Clause}_{\text{non-fin}}

While possible classifications of the structure \([\text{Subj V Clause}_{\text{Non-fin}}]\) have been discussed earlier (see Chapter 5.1, pp. 126-129), in this chapter I examine the translations of this heterogeneous pattern, represented through 25 examples in English and 9 examples in the German originals. In almost all of these instances, with the exception of one English sentence, the non-finite clause is an infinitive clause. In the translation direction English-German, the verb introducing the infinitive phrase, such as \textit{continue, fail or tend}, is often translated through an adverb. For instance, in (110) the verb \textit{continue} corresponds to the adverb \textit{weiterhin} ‘still’. Another translation strategy, illustrated through (111), involves a nominalization of the verb from the original infinitive clause. Here the verb \textit{seek} corresponds to the verb \textit{anstreben} ‘seek’, whereas the verb \textit{negotiate} is turned into the noun \textit{Aushandlung} ‘negotiation’ and the verb \textit{open} is translated through the noun \textit{Öffnung} ‘opening’ (see also the discussion of this example in Chapter 5.1.2). Both translation strategies are characterized by a reduction of clauses\textsuperscript{36}. This effect could also be achieved when the ST verb is either translated through a modal, such as \textit{möchten} ‘would like’ or \textit{sollen} ‘should’, or is missing in

\textsuperscript{36} As discussed above, the sentences discussed in this chapter have been grouped together because it is not clear whether the infinitive verb form and its arguments constitute one chunk on the level of the main clause. Since the rest of the present study concentrates on the main clause and its elements, the notation Clause_{\text{non-fin}} might be misleading. It is important to stress that the \([\text{Subj V Clause}_{\text{Non-fin}}]\) pattern does not necessarily consist of three units that realize separate grammatical functions. In other words, different elements following the verb might not form one unit with the function of the direct object. Therefore, the discussion of clause reduction in the context of this heterogeneous pattern should be interpreted differently from clause reduction related to structures discussed in other chapters (see e.g. Chapter 5.2.2). Here the number of clauses is determined on the basis of the number of verb phrases present in the sentence. Each clause consists of one finite or non-finite verb form (Hansen-Schirra, Neumann and Kunz 2010). Therefore, as shown in example (112), in some cases it makes sense to talk about several non-finite clauses. Because some of the sentences contain one non-finite clause, while others several ones, which could be translated through different strategies, I do not quantify the various translation alternatives employed for their translations.
the translation altogether. However, it is also possible to translate the original infinitive clause one to one by another infinitive (see (112)), or by a *dass* ‘that’-clause.

(110) **EO:** We will [continue] verb to cover Chicago’s accomplishments and its problems, its immigrant communities and its many contributions to American cultural life. (EO_ESSAY_014 s27)

**GTRANS:** Wir werden [weiterhin] Adv die Leistungen und Probleme von Chicago ansprechen, seine Einwanderercommunity und seine vielen Beiträge zum kulturellen Leben Amerikas.

(111) **EO:** We will seek to [negotiate] verb regional and bilateral agreements to [open] verb markets around the world. (EO_ESSAY_015 s71)

**GTRANS:** Wir werden die [Aushandlung] noun regionaler und bilateraler Abkommen zur [Öffnung] noun der Märkte auf der ganzen Welt anstreben.

(112) **EO:** These international partnerships help [leverage resources,] clause_non-fin [increase the knowledge base,] clause_non-fin and [expand markets for advanced energy technology,] clause_non-fin (EO_ESSAY_001 s70)

**GTRANS:** Diese internationalen Partnerschaften helfen, [Ressourcen besser zu nutzen,] clause_non-fin [das Wissen zu vergrößern] clause_non-fin und [die Märkte für moderne Energietechnologie zu erweitern.] clause_non-fin

While the reduction of clauses could be seen as a strategy to simplify sentences, especially on the part of the translator, it is important to consider that the complexity could also be moved from the structural to the lexical level or from the clausal to the phrasal level. The processing effort, which is required to process a sentence, is more likely to be language-specific (cf. Doherty 1996). Therefore, generalizations concerning the process of simplification require a more detailed contrastive analysis of possible factors having an effect on the cognitive processing of sentences.

Also in the opposite translation direction, some of these patterns are translated by a structure containing a non-finite clause, whereas in another case the original infinitive clause is turned into a finite *that*-clause. In other examples, the finite verb is either left out or translated by the marginal auxiliary *need to*. Moreover, in one case the combination *lassen hoffen*[^37] has been translated through an instance of the existential construction [*There is hope*] nominalizing the infinitive verb form.

[^37]: The German original data contains two constructions with the combination of *lassen* ‘let’ and a verb in the infinitive: these have been also considered as instances of [Subj V Clause_non-fin], since the verb *lassen* is not classified as a modal (Eisenberg 2004, 351).
Thus, depending, among other things, on the translation direction, the pattern [Subj V Clause\textsubscript{non-fin}] could be either translated by one of the strategies leading to a clause reduction or a similar structure involving a finite or a non-finite clause.

5.1.6 **Translation shifts from [Subj V Dobj]**

In the previous sub-chapters, I analyzed four rather specific patterns, which are assumed to be lower-level instantiations of the general transitive construction and function as independent constructions. These are [\textNP\text{research} V\text{show} clause/NP], [\textNP\text{country/government} V\text{feel/think clause/NP}], [\textNP\text{human} V\text{Speak} Clause/NP] and [\textit{Es gibt} Dobj] or [\textit{There be} Dobj]. The chapters discussed whether the structures could be language- or register-specific and how they are translated in both translation directions. Moreover, translations of the patterns [\textit{Man} V Dobj] and [Subj V Clause\textsubscript{non-fin}] have been briefly considered. In this chapter, we return to the general research question of construction shifts and examine in more detail possible reasons for changing the original structure in the process of translation. For this purpose, a logistic regression model with mixed effects (Baayen 2008; Winter 2013) is built, which examines the probabilities of various factors to cause translation shifts from the construction [Subj V Dobj]. Before the discussion of results, it is important to motivate the individual variables included into the model.

The discussed lower-level constructions describe a quarter of all English and German original sentences with the general [Subj V Dobj] structure in our data (25%, 104/422). These have been selected for further study based on their high frequency within the transitive patterns that are significantly attracted to English or German. Moreover, also the initial categorization of the schematic [Subj V Dobj] construction according to animacy and volition is motivated by a contrastive difference between the two languages. Therefore, the translations could be expected to reflect the frequencies of the originals in the same languages and shift from the constructions associated more with the source rather than the target languages. At the same time, analyses taking into account reference corpora have shown that the use of constructions is often linked to a specific register rather than language in general. For instance, the present study suggests that the construction [\textNP\text{research} V\text{show} clause/NP] is used more frequently in English academic and popular-scientific discourse compared to the comparable German registers, but there are some indications that in German the
construction is used in newspapers and magazines, i.e. a register of journalistic texts. However, this assumption should be further tested using a corpus of journalistic language. Comparing the results to the register of popular-scientific texts would allow identifying relative entrenchment of the construction in the two registers. Another interesting example is the construction \([\text{NP}_{\text{country/government}} \text{V}_{\text{feel/think}} \text{clause/NP}]\), which occurs only in the English original data belonging to the register of political essays. Our analysis has shown that the construction does occur in German political discourse included in the CroCo corpus, for instance in the register of prepared speeches. In fact, the use of the construction might depend on the author and topic of political texts.

The fact that these contrasts between the use of constructions in the two languages are fine-grained rather than categorical could be the reason why the presence of these patterns in the originals leads to fewer translation shifts than could be expected on the basis of the initial comparisons of originals in our main data sample. It was observed that translators often keep the original structure in their translations, even though the construction might be less entrenched in the particular target register. I assume that the translation property of shining though can be facilitated by the fact that the corresponding target constructions are used in similar types of discourse. For instance, if a target language construction is more common in political speeches than political essays, the commonalities between the two registers might have an effect on the frequency of the construction in the translations belonging to one of these registers. In other words, by translating the SL construction through a corresponding TL construction the translators might slightly violate the norms of the particular register in which they are operating but this could be preferred to potentially investing additional cognitive resources into changing the structure of the sentence. Moreover, the SL pattern is even more likely to shine through in translation, if the frequency of a construction is different depending on the author and topic rather than the register in general.

Taking into account these factors, the identified constructions can potentially explain only a small proportion of all translations shifts from the transitive patterns. In the discussion of individual examples it has been also shown that the shifts sometimes result in more impersonal structures and reduction of clauses. It is possible that high number of clauses and low degree of impersonality of the original sentences are in fact reasons for the construction shifts.
Based on these considerations, a number of predictors that could have an effect on the response variable of construction shifts were included in the statistical model. First of all, the binary nominal variables “language” (English/German) and “register” (ESSAY/POPSCI) were added to verify whether one or both of these factors increase the likelihood of shifts from [Subj V Dobj]. Secondly, I also included the ordinal variable “no.clauses.main” referring to the number of clauses on the level of the main clause. During the annotation of this variable, value 1 was assigned, if both the subject and the direct object were realized through phrases. Value 2 was assigned, if either the subject or the direct object corresponded to a finite or a non-finite clause, whereas value 3 was present when both elements were realized through clauses. The level of impersonality is a complex variable, which could depend on the linguistic context of individual sentences, the subject (i.e. whether those responsible have been named as the head of the NP, in modification or left out completely; and whether a particular study has been referred to in some way, e.g. by a research noun) and the verb (i.e. whether it is a verb of speaking, thinking, showing, etc.). Due to the lack of a coherent description of all types of impersonal structures and all effects that could play a role in this context, the variable “degree of impersonality” has not been taken into account. More research in this area is required in order to develop a reliable operationalization of this semantic feature. Moreover, since the individual constructions identified in the present study occurred very rarely or not at all in some of the categories, the analysis in this chapter does not consider the more specific constructions and, therefore, goes back to the general level of the [Subj V Dobj] construction. Apart from the fixed effects of language, register and the number of clauses, the model also includes a random intercept for texts. The response variable of construction shifts is a binary nominal variable: the values yes/no indicate whether a shift occurred or not. The analysis was performed using R (R Core Team 2014) and the package lme4 (Bates et al. 2014). The results of the fixed effects are shown in Table 23.

|                     | Estimate | Std.Error | z value | Pr(>|z|) |
|---------------------|----------|-----------|---------|----------|
| (Intercept)         | -0.62    | 0.35      | -1.76   | 0.08     |
| no.clauses.main     | 0.007    | 0.21      | 0.03    | 0.97     |
| languageGerman      | -0.68    | 0.24      | -2.84   | 0.005    |
| registerPOPSCI      | 0.42     | 0.23      | 1.85    | 0.06     |

Table 23: [Subj V Dobj] - Summary of the fixed effects on the variable of construction shift
It can be seen that the effect of the variable “language” on the category of construction shifts is significant. The negative value of the coefficient estimate indicates that for German, or, to be more precise, in the translation direction German-English the shifts from the [Subj V Dobj] construction are less probable. Moreover, the p-value for the variable “register” is close to the conventional level of significance of 0.05. Therefore, it could be assumed that also register has an effect on shifts. The register of popular scientific texts is more likely to cause shifts from the discussed argument structure construction. Finally, the variable “no.clauses.main” does not have a significant effect on the response variable under analysis.

The model was checked for multicollinearity using the R function \textit{vif.mer} \textsuperscript{38}, which can be used for logistic regression models with mixed effects. All values are below 5 (LanguageGerman 1.003, registerPOPSCI 1.0007, no.clause.main 1.002) indicating that there is no multicollinearity. To establish the goodness of fit for this model the function \textit{somers2} from the Hmisc R package (Harrell 2014) was used (cf. Baayen 2008). The values are rather low and, therefore, indicate that there are additional reasons for the construction shifts that are not accounted for by the model (C-value=0.67, Dxy=0.35). As discussed above, the level of impersonality and certainly the lower-level constructions are good candidates for additional predictors.

5.2 Subject Verb Complement

The second schematic construction that has been selected for the analysis of construction shifts is [Subject Verb Complement] \textsuperscript{39}, where complement is understood as a grammatical function rather than any obligatory element in a clause. This classification is based on the annotation used in the CroCo corpus, which, in turn, refers to Quirk et al. (1985). The same syntactic function is referred to as a predicative in Biber et al. (1999). It should also be mentioned that the analysis of the copular construction [Subj V Compl] means that only instances of the subject complement are taken into account, since the object complement follows the direct object.

In English the distinction between subject and complement is based mainly on the word order. Due to the fixed word order in English, the grammatical element preceding the copular

\textsuperscript{38} The original function \textit{variance inflation factors} (\textit{vif}) from the rms package was modified by Austin Frank (https://hlplab.wordpress.com/2011/02/24/diagnosing-collinearity-in-lme4/).

\textsuperscript{39} From now on, the construction is referred to as [Subj V Compl].
verb is considered to be the subject of the clause, whereas the second noun phrase (NP) or a clause is the complement. In contrast, in German the functions of the two NPs or clauses could be distinguished using, for instance, the *gelten als*-test: the copular verb such as *sein* is substituted by *gelten als*, so that the *als*-phrase is identified as the complement of the clause (Albert et al. 2003, 66; Duden 2009, 813). However, the test cannot be applied in all cases leaving a part of sentences ambiguous with respect to the grammatical functions of subject and complement (Eisenberg 2004, 291). Since in both languages the subjects are typically realized through NPs and clauses (Biber et al. 1999, 123; Eisenberg 2004, 284), adjective phrases (AP) and prepositional phrases (PP) occurring in the copular sentences are considered to be complements, irrespectively of the position they occur in.

Similar to the construction [Subj V Dobj] discussed above, the construction [Subj V Compl] occurs relatively frequently in the English and German originals, when compared to other constructions consisting, for instance, of four elements. As can be seen in Table 24, in the translation direction English-German it accounts for 12% (60/495 in the register of political essays, 60/488 in the register of popular-scientific texts) of all construction shifts. Moreover, in this translation direction, 41% (60/148) of all instances of the copula construction undergo a construction shift in the sub-corpus ESSAY, whereas the number of construction shifts amounts to 53% (60/113) in the register POPSCI. A comparison between the two translation directions shows similar results for the copular construction, discussed here, and the transitive construction, discussed in the preceding sub-chapter. The shifts from the construction [Subj V Compl] reach only 4% in the register ESSAY (23/657) and 3% in the register POPSCI (12/442) in relation to the overall number of shifts from German into English. The registers have similar frequencies of shifts from this construction in relation to all examples of [Subj V Compl] in the originals: the construction is shifted in 19% and 15% of cases (23/122 in ESSAY and 12/80 in POPSCI). The difference in shifts from the copular construction in the two translation directions is significant for both registers. Both in the register of political speeches ($\chi^2=13.77$, df=1, p-value=0.0002, see Table 25) and in the register of popular-scientific discourse ($\chi^2=27.46$, df=1, p-value=1.603e-07, see Table 26) there are more shifts in the translation direction English-German.
Explaining these construction shifts, it is important to take into account a contrastive difference between English and German. A previous study has shown that the languages have different preferences for grammatical functions: while in English complements are frequently used, in German the function of prepositional object is preferred (Čulo et al. 2012, 114). Also another recent study has indicated a higher number of complements (in relation to all VPs in a corpus) in English than in German (Fischer 2013, 301-302). This is in line with the contrastive analysis of our data. Table 27 shows that there are more instances of [Subj V Compl] in English rather than in the German originals ($\chi^2=21.33$, df=1, p-value=3.872e-06). Moreover, there are fewer shifts from the construction [Subj V Compl] ($\chi^2=60$, df=1, p-value=9.469e-15, see Table 28) and more shifts into this construction ($\chi^2=48.18$, df=1, p-value=3.881e-12, see Table 29) in the translations from German into English. This could be
seen as a further indicator that the function is more common in English than German. We will return to the contrastive difference with respect to the prepositional objects in Chapter 5.3.

<table>
<thead>
<tr>
<th></th>
<th>English-German</th>
<th>German-English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shifts from [Subj V Compl]</td>
<td>120</td>
<td>35</td>
</tr>
<tr>
<td>Shifts from other constructions</td>
<td>863</td>
<td>1064</td>
</tr>
</tbody>
</table>

Table 28: Shifts from [Subj V Compl] in the translation directions E-G and G-E

<table>
<thead>
<tr>
<th></th>
<th>English-German</th>
<th>German-English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shifts into [Subj V Compl]</td>
<td>26</td>
<td>114</td>
</tr>
<tr>
<td>Shifts into other constructions</td>
<td>957</td>
<td>985</td>
</tr>
</tbody>
</table>

Table 29: Shifts into [Subj V Compl] in the translation directions E-G and G-E

While the different distribution of complements in English and German could play a role in the translation shifts from the construction [Subj V Compl], it is not necessarily the only explanation for the shifts. Therefore, the analysis is extended to consider various copular patterns and potential lower-level constructions.

English copular sentences are typically subdivided into two major functional categories: the relation that holds between the subject and the subject complement could be identifying or characterizing. In other words, in its first function the pattern [Subj V Compl] may express a relation of identity between the two elements. As a result, the order of the constituents can be often reversed. In contrast, in its second function the complement is used to ascribe a certain quality to the subject, i.e. to characterize it in some way (Quirk et al. 1985; Biber et al. 1999). Another classification by Declerck (1988) distinguishes between predicational patterns, which correspond to the characterizing ones, as well as specificational structures, descriptionally-identifying sentences, identity statements, definitions and inferential types of copular structures, which could be interpreted as variants of the identifying pattern. The functions such as characterization, classification and an establishment of identity relations have been also attributed to the German [Subj V Compl] pattern (Zifonun et al. 1997, 1106; Fischer 2013, 295): due to the similarities both on the formal and functional levels, the English and German copular constructions could be considered corresponding.

However, the classification of the copular construction according to the functions is rather problematic due to the lack of a clear distinction: the sentences are often ambiguous with
respect to one of the two major functions (Huddleston and Pullum 2002, 266). Therefore, the initial classification of the structure [Subj V Compl] considers the form of the construction, i.e. the formal categories (e.g. NP or PP) through which the grammatical functions of subject and complement are realized in the two languages. The formal classification is motivated by certain interactions between formal and functional sides of the construction. For instance, copular clauses with APs as subject complements correspond to the characterizing function (Biber et al. 1999, 145-146; Fischer 2013, 295). Moreover, the following investigation of the patterns involves an identification of lower-level constructions. This more fine-grained level of analysis, therefore, considers combinations of formal and functional properties.

(113) – (116) show examples of the major copular patterns in English and German identified in the data.

(113) NP V NP
   a. **EO:** [National security]_{NP} [is]_{V} [the highest priority.]_{NP} (EO_ESSAY_021 s18)
   b. **GO:** [Die Messewirtschaft]_{NP} [ist]_{V} [eine Schlüsselbranche der deutschen Dienstleistungswirtschaft.]_{NP} (GO_ESSAY_011 s20)

(114) NP V AP
   a. **EO:** [They]_{NP} [are]_{V} [wrong.]_{AP} (EO_ESSAY_016 s9)
   b. **GO:** [Der Übergang vom Klartext zur Klartextzahl]_{NP} [ist]_{V} [einfach.]_{AP} (GO_POPSCI_008 s98)

(115) NP V Clause
   a. **EO:** [A second possibility]_{NP} [is]_{V} [that we need them all because each subtype has different properties.]_{Clause} (EO_POPSCI_010 s41)
   b. **GO:** [Das]_{NP} [heißt]_{V} [Deutschland ist hoffnungslos überbevölkert in Bezug auf die Fußabdrücke.]_{Clause} (GO_ESSAY_021 s86)

(116) Clause V NP
   a. **EO:** [Discovering markers to identify truly functional cells]_{Clause} [is]_{V} [another urgent task for the stem cell research community.]_{NP} (EO_POPSCI_006 s42)
   b. **GO:** [Ergebnis]_{NP} [ist]_{V} [daß die Versicherten künftig selber mehr investieren müssen in ihre Gesundheit (oder ihre Altersversorgung), der private Anteil steigt.]_{Clause} (GO_ESSAY_005 s41)

For this classification, all personal and impersonal pronouns (i.e. *man* and *es*) together with the pronouns of extended reference, such as *das* in (115b), are counted among NPs. In (115a) and (116a) the noun phrase and the clause are respectively the subjects because these elements occur before the finite verb, whereas in (116b) the application of the *gelten-als* test would identify the noun *Ergebnis* ‘result’ as the complement rather than the subject. Although
in example (115b) both paraphrases with *gelten als* sound rather unnatural, it is assumed that the definite element, in this case the pronoun *das*, is more likely to function as the subject of the sentence.

Similar to the transitive construction discussed in 5.1, the copula sentences may contain non-finite clauses. For instance, in the English sentences (117)–(119) the verb is followed by an infinitive clause.

(117) **EO:** Our mission at the Commerce Department is [to promote both national security and U.S. exports.]\textsubscript{Clause\_non-fin} (EO\_ESSAY\_021 s17)

(118) **EO:** Interferons - proteins produced by the body that are active against viruses and may help to suppress tumours - are proving [to be a lot more complex than early experiments suggested.]\textsubscript{Clause\_non-fin} (EO\_POPSCI\_010 s2)

(119) **EO:** This nonlinear relation between regulation and function appears [to be a feature of all integrally organized systems.]\textsubscript{Clause\_non-fin} (EO\_POPSCI\_004 s120)

This group of copular sentences is heterogeneous, as can be shown by applying the same tests used to analyze the pattern [Subj V Clause\_non-fin] in Chapter 5.1 (see pp. 126-129 for discussion). The question test proposed by Halliday and Matthiessen (2013, 567, footnote 1, 585) indicates that whereas the whole infinitive phrase following the third person singular form of the copular verb *be* in (117) could be seen as one grammatical element, the verb phrases *are proving to be* and *appears to be*, in (118) and (119) respectively, form a unit followed by an AP in (118) and a NP in (119). The corresponding questions are *what is our mission?*, *how are interferons proving to be?* and *what does this nonlinear relation appear to be?* Moreover, according to Quirk et al. (1985, 146-147), the clause in (119) could be considered a catenative verb construction, where the verb *appear* is not the main verb followed by a complement but modifies the main verb *be*. Also in German, it is not always clear whether the verb in the infinitive form constitutes a separate clause or a VP together with the main verb. In (120a) all three verbs could be interpreted as belonging to one VP with *übersehen* ‘overlook’ as its main verb describing the mental action of the subject (Duden 2009, 848). The syntactic behavior of the verb *scheinen* ‘seem’ and similar uses of certain verbs such as *pflegen, drohen* and *versprechen* is compatible with modal verbs (Eisenberg
2004, 363). Paraphrasing the main clause as a subordinate clause, the verb moves to the end of the sentence (see (120b))

(120) a. Die diuretischen Wirkungen schien man [übersehen zu haben.]\textsubscript{Clause_non-fin}
(\textit{GO\_POPSCI\_010 s29})
b. ...dass man die diuretischen Wirkungen übersehen zu haben schien.
(121) [Ebenso schwer zu realisieren]\textsubscript{Clause_non-fin} ist das Ideal, der Partner habe fähig zu sein, einen in allen seinen Gefühlen, Bedürfnissen und Handlungen restlos zu akzeptieren und zu unterstützen, einen zu verstehen und zu fördern, ohne eingrenzende Ansprüche und Erwartungen zu stellen. (\textit{GO\_POPSCI\_004 s129})

In contrast, it is observed that the verb \textit{heißen} ‘mean’ can be followed by an infinitive clause functioning as the subject complement (Duden 2009, 792). While the verb \textit{sein} ‘be’ is not explicitly listed in Duden as the one that can also introduce a non-finite clause as a complement, it is assumed that (121) could be classified as such an example. The \textit{gelten als} test indicates that the initial infinitive clause functions as the complement. However, the German data contains more examples in which a clause corresponds to the grammatical function of the subject.

Taking into account these possible classifications of the copula clauses with non-finite verb forms, that do or do not constitute independent clauses functioning as complements, the present analysis groups these instances into a separate category. Therefore, parallel to the transitive pattern discussed in Chapter 5.1, the copular verbs (e.g. \textit{be}, \textit{appear}) followed by an infinitive form the structure [\textit{Subj V Clause}\textsubscript{Non-fin}]. As above, the sentences with clauses functioning as subjects are not considered as examples of this category.

To begin with, the frequencies of copular patterns are compared between the English and German originals. Figure 20 shows an overview of different sub-types of copular constructions in the two languages, in which the category “other” comprises four minor groups of sentences, namely [\textit{Clause V Clause}], [\textit{Clause V PP}], [\textit{NP V AdvP}] and [\textit{NP V PP}].

\footnote{It should be mentioned that not all grammars agree that the verb \textit{scheinen} ‘seem’ should be included into the group of copular-(like) verbs such as \textit{heißen} and \textit{schmecken} (Eisenberg 2004, 86).}
We can observe that the patterns that occur more often in general are also the ones that display different frequencies in English and German. These are [NP V NP], [NP V AP], [NP V Clause] and [Clause V NP]. It is interesting to observe that the German originals are characterized by one dominant copular pattern, namely [NP V NP], which accounts for 64% (127/200) of all instances. In contrast, the frequencies of the three patterns [NP V NP], [NP V AP] and [NP V Clause] in the English originals cluster more closely together (42% or 99/237 instances of [NP V NP], 30% or 70/237 instances of [NP V AP] and 16% or 39/237 of [NP V Clause]). The general effect of language on the distribution of patterns is significant ($\chi^2=44.12$, df=5, p-value=2.187e-08). The mosaic plot in Figure 21 helps us determine what patterns in particular are attracted to English or German.
As can be seen in Figure 21, the copular pattern [NP V NP] is attracted to German and is less likely to occur in English, whereas the patterns [NP V AP] and [NP V Clause] are associated rather with the English language. This finding is in contrast to the previous study by Fischer (2013, 301-309), which suggested that both languages contain more APs than other elements functioning as complements, whereas the proportion of NPs is higher for English than for German. The difference could be attributed to the underlying data samples. As mentioned above, the study by Fischer is based on three texts and their respective translations belonging to the registers of fiction and political essays (Fischer 2013, 25-26), whereas the present study considers a wider range of texts both from the register of political essays and popular-scientific texts.

Based on the results reported in Figure 21, it is possible to assume that in the translation direction English-German the patterns [NP V AP] and [NP V Clause], which are more typical of English, will be shifted more often than kept in translations. To test this hypothesis, 237 examples of the English copular construction are analyzed with respect to shifts. Figure 22 shows that only one of these two structures, namely [NP V Clause], is changed more often than it is kept in this translation direction. To avoid zero values in the calculation of the chi-
squared test, for this particular analysis the minor categories [Clause V NP] and [Clause V AP] have been considered among the heterogeneous group of other patterns. The general effect of the occurrence of copular patterns on the construction shifts is significant ($\chi^2=24.8$, df=3, p-value=$1.703\times10^{-5}$). The mosaic plot in Figure 23 indicates that it is the pattern [NP V Clause] that shows significant association with the category of construction shifts. Therefore, it was possible to partly confirm the hypothesis: in the translations from English to German the pattern [NP V Clause] typically undergoes certain changes, whereas the pattern [NP V AP] has slightly more instances that remain unchanged in the target texts than the ones affected by shifts. Based on these results, in the next step the copular pattern [NP V Clause] is studied in more detail.

![Figure 22: Construction shifts from the copular patterns in the translation direction E-G](image)

![Figure 23: Copular patterns & construction shifts in E-G (mosaic plot)](image)
A closer look at the English data reveals that this structure could be subdivided into several lower-level constructions. Among those constructions discussed in the literature are *it*-clefts, pseudo-clefts (or *wh*-clefts) and demonstrative *wh*-clefts (Biber et al. 1999; Collins 1991; Oberlander and Delin 1996; Quirk et al. 1985). Below we discuss these structures, their alternative translations as well as possible reasons for the construction shifts affecting the pattern [NP V Clause]. Sub-chapters 5.2.1 and 5.2.2 take into account 39 instances of the English and 6 examples of the German [NP V Clause] structure and translations into this pattern. These numbers do not include the sentences classified as the type [Subj V ClauseNon-fin] (see the discussion above). The chapter is concluded by an analysis of different factors potentially leading to shifts from the construction [Subj V Compl] (see Chapter 5.2.3).

5.2.1 Demonstrative *wh*-cleft construction

As mentioned above, the construction shift dataset of the English originals comprises 32 examples of the copular pattern [NP V Clause]. Within these, the demonstrative *wh*-cleft, illustrated through (122), accounts for 59% (19/32), whereas only one instance of this structure is preserved in the German translations.

(122) **EO:** [That]Pro_dem-sing [is]V_be [why Congress granted and renewed fast-track authority, the forerunner of TPA, for 20 years until it last expired in 1994.]Clause (EO_ESSAY_009 s38)

The form of this cleft type has been described as [That’s *what/why/where* Pro_personal/N_proper V_cognition/communication/movement] (Calude 2009, 69). It should, however, be mentioned that this formula contains the most frequent items identified for the spoken data. While our sample is also characterized by a high frequency of the demonstrative pronoun *that* (14/20 instances or 70%) and the question pronouns *why* or *what* (present in all 20 examples), other elements of the pattern display more variation. A more general formula would be [Pro_dem-sing V_be (Pro_wh...)clause].

The demonstrative pronoun in the demonstrative cleft typically functions as an extended anaphoric reference (Collins 1991; Halliday and Hasan 1976, 66; Oberlander and Delin 1996, 215) creating a cohesive link to the preceding discourse. Moreover, rather than to introduce any new information, the main function of the pattern is to sum up a preceding argument (Biber et al. 1999, 962; Collins 1991; Oberlander and Delin 1996, 200). Using this structure,
the speaker typically refers to her/his own previous statement often with the purpose of creating relations between different parts of discourse and emphasizing the main ideas. Furthermore, in some cases the speaker could also extend the reference to the utterance of another speaker; this strategy allows for a non-threatening turn-taking (Calude 2009, 60).

The term “demonstrative (wh-) cleft” has been adopted from Biber et al. (1999) and Calude (2009). In other studies, the pattern has been referred to as the reversed wh-cleft, reversed pseudo-cleft (Collins 1991; Oberlander and Delin 1996) or simply as that is what-construction (Johansson 2001). Calude (2009) has shown that there are certain differences between the reversed and the demonstrative clefts that justify a separate analysis of the latter construction. The main difference is claimed to be the deictic link to the context, which is present only in the demonstrative clefts (Calude 2009, 59). Furthermore, previous studies have observed that the demonstrative patterns frequently cannot be reversed (Biber et al. 1999, 961; Johansson 2001, 587). This fact is related to the property of the pronoun that (Higgins 1979, 233), which, in contrast to this, can be used only anaphorically, i.e. referring to the preceding text (Halliday and Hasan 1976, 68). Since in the reversed structure the pronoun would occur in the clause-final position, it would require a cataphoric link to the following passage of text. Taking these arguments into account, the present study follows Calude (2009) in analyzing the demonstrative clefts in their own right rather than as a sub-type of reversed clefts.

Previous research has identified the demonstrative clefts as formulaic (Calude 2009). In other words, this form-meaning pairing is assumed to be a lower-level construction stored along the more schematic copular construction. The arguments that are used to support this claim are a strong link between this construction and the register of informal conversation, the relatively fixed slots of the construction, absence of pauses or discourse markers during its (phonological) production as well as a rather unspecific reference of the demonstrative pronoun introducing the pattern (Calude 2009, 63-73). Before the register and fixed slots of this pattern are discussed in more detail below, it is worth commenting on the other two indicators of formulaicity. Wray (2002, 35-37) suggests that with the acquisition of a unit-status the structures tend to be produced more fluently. While studies in this area have considered phonological features during production of certain structures in oral discourse, future research should extend the study of fluency with respect to the written production.
For this purpose the method of keystroke logging can be used to observe pauses during the production of these patterns. Apart from the fluent production of the demonstrative clefts, Calude (2009, 71-73) suggests that the extended reference of the demonstrative pronoun results in reduced salience of this construction, as it is not always clear to what text passage the pronoun refers. Following Hudson (1998, 109-110), the reduced salience is linked to the increased fixedness of the form-meaning pairing under analysis.

As mentioned above, studies have shown that the construction frequently occurs in the register of conversation (Calude 2009, 63-65; Biber et al. 1999, 961-963). This link is partly explained through the low level of formality associated with this construction, especially when the initial slot is filled with the pronoun that. Furthermore, the higher degree of repetitiveness in the register of conversation seems to be compatible with repeating and summing the individual aspects of the preceding discourse with the help of this construction (Biber et al. 1999, 963). It is in this function connected to discourse organization that the demonstrative clefts have been demonstrated to be often used in conversation (Calude 2009, 65). While the association between the register of informal conversation and the demonstrative clefts appears to be undisputable, Calude (2009, 64) also states that due to a limited number of studies that regard the demonstrative clefts in their own right, the role of this construction in written registers is still to be established. With respect to our data, we can observe that in the written registers of political essays and popular-scientific texts the construction accounts for 8% (20/237) of all copular sentences in the English originals.

The demonstrative cleft construction can be described as a partially lexically-filled construction with relatively fixed slots (Calude 2009, 65-69). Therefore, using the method of collostructional analysis, I examined what lexical items are attracted to the slots of this form-meaning pairing. The collostructional strength is calculated on the basis of all occurrences of these lexical items and the number of sentences in the two registers under analysis. As shown in Table 30, both demonstrative pronouns and the two wh-pronouns attested in our sample of English originals are significantly attracted to the investigated construction. Higher collostructional strength has been identified for the demonstrative pronoun that and the wh-pronoun why. The covarying collexeme analysis has indicated that the pairs that – why and this – what are attracted, whereas this – why and that – what are repelled. However, all of these combinations display the collostructional strength of 1.1 and are, therefore, not
significant (significance at the level of 0.05 is reached when the log-transformed p-values are above 1.3).

<table>
<thead>
<tr>
<th>Collexeme_</th>
<th>Freq. in the corpus</th>
<th>Freq. in the constr.</th>
<th>Coll. strength</th>
<th>Collexeme_</th>
<th>Freq. in the corpus</th>
<th>Freq. in the constr.</th>
<th>Coll. strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>demonstrative</td>
<td>this</td>
<td>330</td>
<td>6</td>
<td>1.72</td>
<td>why</td>
<td>52</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>that</td>
<td>939</td>
<td>14</td>
<td>3.28</td>
<td>what</td>
<td>83</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 30: [Pro\textsubscript{dem-sing} be (Pro\textsubscript{wh-})\textsubscript{clause}] - Significantly attracted collexemes for Pro\textsubscript{dem-sing} and Pro\textsubscript{wh}

Moreover, it is interesting to observe that whereas the majority of demonstrative clefts, namely 16 instances, occur in the register of political essays, all four examples from the popular-scientific discourse contain the more formal (Biber et al. 1999, 963) demonstrative pronoun this. However, these four sentences belong to the same text.

To compare the distribution of the demonstrative clefts across different registers, the construction has been queried in the BNC. The query consisted of three items directly following each other, namely the demonstrative pronoun this/that, the lemma be and one of the wh-words, i.e. why, what, how, where, when or who. Thus, it does not return instances of the construction in the interrogative mood. It identified a total of 18,653 instances of the demonstrative clefts distributed among 39 different combinations of lexical items filling the three slots. The frequency distribution confirms the results of the previous studies that the construction is much more frequent in the spoken discourse. In this group of texts, it occurs 798 times per 100,000 sentences. In contrast, in the BNC category of non-academic texts (i.e. popular-scientific discourse), the frequency of the construction amounts only to 123 times per 100,000 sentences.
Figure 24 shows the distribution of the ten most frequent variants of the demonstrative clefts across the registers included in the BNC except for the miscellaneous texts. It can be seen that whereas in the spoken data the two most frequent patterns, namely [That’s what] and [That’s why] together account already for 64% of all occurrences (5,288/8,286), in the register of non-academic texts the top six combinations reach only 61% (564/919). These six patterns are [This is what], [This is why], [That’s what], [That is why], [That is what] and [That’s why]. In other words, the form of the demonstrative clefts varies less in the spoken texts. In non-academic and academic discourse, in contrast to the four other registers, the two most frequent patterns contain the demonstrative pronoun this followed by a non-contracted form of the verb. These differences could be attributed to the higher level of formality associated with the registers of academic and popular-scientific texts (Biber et al. 1999, 963). The frequency of patterns in the register of fiction is closer to the spoken than written data potentially because certain features of spoken discourse are reflected in dialogues, which are characteristic of fictional texts (Biber et al. 1999, 16). The journalistic texts behave similar to the spoken register with respect to the most frequent structures but show more variation, similar to the groups of academic and non-academic texts. Moreover, it is worth mentioning that eleven out of 39 variants of the demonstrative clefts identified in the BNC contain the past tense form of the copula verb, one of them, namely [That was what], being among the ten most frequent. This variant plays a more important role in the register of fictional texts due to the frequent use of the past tense in narration (Biber et al. 1999, 456). Thus, the form...
of the demonstrative clefts and the level of fixedness of the individual slots seem to vary depending on the register.

A random sample of 100 demonstrative clefts from the register of popular-scientific texts shows that only about a third of instances (35 cases) form independent sentences (as in Chapter 5.1, the true hits do not include sentences containing additional clauses, constructions integrated into subordinated clauses as well as those containing further grammatical functions except for the subject, verb and complement). The category of coordinated sentences plays an important role in the sample (44 cases)\(^{41}\). In a number of these instances the passage of a text, which is being referred to by the demonstrative pronoun *this* or *that*, occurs in the same sentence as the demonstrative cleft.

A previous study on translations of demonstrative clefts from English to German (Johansson 2001) has suggested that the German pattern that is most similar to the English construction has the form \([Das \, V_{\text{sein}} \, \text{der Grund/die Dinge, warum/die...}].\) In example (123), which is characterized by this structure, the noun Grund ‘reason’ is modified by a prepositional phrase. Johansson (2001, 589) notes that “German seems to require a noun-phrase head before the subordinate clause”. This extended structure is also possible in English and is illustrated through (124). Both sentences in (123) and (124) are the only examples in the corresponding originals belonging to our main data set. To make the queries as comparable as possible, the reference corpora BNC and DWDS were searched for the combinations \([This/That \, V_{\text{be}} \, \text{the reason}]\) and \([Dies/Das \, V_{\text{sein}} \, \text{der Grund}].\) The queries returned the total of 202 hits for English and 218 hits for German. While these values are similar across the languages, they are rather low compared to 4,462 occurrences of the pattern \([This/that \, V_{\text{be}} \, \text{why}]\) in the BNC.

\((123)\) **GO:** \([\text{Das}]_{\text{NP}} \, [\text{ist}]_{\text{V}} \, \text{der Grund dafür, daß in Zukunft aller Voraussicht nach Sicherheitsschlösser entwickelt werden dürften, die aufgrund eines elektronischen Erkennungsmechanismus einzig und allein auf den Daumenabdruck der zum Öffnen befugten Person ansprechen.}]_{\text{NP}} \) (GO_POPSCI_005 s18)

\((124)\) **EO:** \([\text{That}]_{\text{NP}} \, [\text{is}]_{\text{V}} \, \text{the reason, why the British Government, along with four others, could not agree to the proposals for the EU Budget last week.}]_{\text{NP}} \) (EO_ESSAY_003 s7)

---

\(^{41}\) It should be mentioned that it is difficult to compare these results and the use of the demonstrative clefts in the spoken discourse because in the latter the borderline between clause and sentence is less clear (Biber et al. 1999, 10).
Other possible translation strategies mentioned in the previous study involve the German pattern [Das war es, was] ‘That was it, what’, which is considered to be more similar to the English it-clefts, as well as two structures with reduced number of clauses but either analogous or different semantic information mapped onto the clause-initial element (Johansson 2001).

In what follows, translations of the two more specific variants of the demonstrative clefts are discussed in more detail. The first lower-level schema [That/this is why] is used to sum up the explanations for something that has been described in the preceding stretch of text. This construction accounts for twelve instances in the construction shift data and one instance in the group without shifts. In all translations of the English construction belonging to the former category, the three initial elements characteristic of the demonstrative cleft, namely a demonstrative pronoun, the copular verb and a wh-pronoun, are replaced by the adverbials deshalb ‘therefore’ and aus diesem Grund ‘for this reason’. This element could be classified as analogous theme (Johansson 2001), i.e. the clause-initial constituent realizing the same semantic content as the parallel element in the original. In most cases it is followed by a verb corresponding to the one integrated in the original relative clause. These changes lead to a reduction in the number of clauses, as can be seen in (125).

(125) **EO:** That’s why turnout is much higher for national elections than European ones.  
**(EO_ESSAY_013 s39).**  
**GTrans:** Deshalb ist die Wahlbeteiligung bei nationalen Wahlen viel höher als bei europäischen.

This example also demonstrates that while both the original and the translation are copular constructions, the sentence pair is still classified as a construction shift due to a crossing line, in this case between why functioning as the complement in the original and deshalb in the corresponding translation, which could be classified as a connective adverbial. As the name suggests, the element deshalb has a cohesive function in German, similar to the demonstrative pronoun dieser ‘this’ in the expression aus diesem Grund ‘for this reason’. Thus, the cohesive function of the original construction is typically preserved in these translations. The same can be said about the only instance of this pattern in the data without construction shifts, which is translated by the pattern [Das ist der Grund], discussed above. Also here the initial element das ‘this’ links the sentence to the preceding discourse.
The second variant of the demonstrative clefts identified in the data is [That/this is what], which occurs seven times and is, in all cases, changed in the translation. Five of these examples are translated by a reduced clause structure. Also in these cases the analogous theme is followed by the verb from the original relative clause. Moreover, in two out of these five clauses the focusing particle genau ‘precisely’ is used, as can be seen in example (126).

(126)  **EO:** This is what has happened in gene-centred biology (EO_POPSCI_008 s23).

    **GTrans:** Genau dies geschah in der genozentrischen Biologie.

    precisely this happened in the gene-centred biology

Johansson (2001: 593-594) observes that the lexical items genau ‘precisely’ and gerade ‘exactly’ are common in the German translations of the English demonstrative clefts and suggests that they function as the “lexical reinforcement”, i.e. as a way to put an additional emphasis on the following element, in this example the demonstrative pronoun dies. The construction shift is present due to a reduction of clauses: while in gene-centred biology is part of the subordinate clause in the original functioning as the subject complement, in the translation the corresponding lexical content is mapped onto an adverbial of location on the level of the main clause.

Moreover, the construction shift data contains two examples with non-analogous theme. In (127) the original clause-initial element corresponds to the prepositional phrase in the sentence-final position. The translator alters the information structure but partly compensates this change by adding the focusing particle genau.

(127)  **EO:** This is what is meant by computing an adult organism from the information in its genes (EO_POPSCI_008 s91).

    **GTRANS:** Der Glaube, man könne einen erwachsenen Organismus aus den Informationen in seinen Genen "berechnen", bezieht sich auf genau diesen Punkt.

Considering all copular constructions, it is notable that the demonstrative clefts are significantly more frequent in the English originals than in translations ($\chi^2=8.38$, df=1, p-value=0.004, see Table 31).

<table>
<thead>
<tr>
<th></th>
<th>Demonstrative clefts</th>
<th>Subj V Compl_other</th>
</tr>
</thead>
<tbody>
<tr>
<td>English originals</td>
<td>20</td>
<td>217</td>
</tr>
<tr>
<td>English translations</td>
<td>6</td>
<td>257</td>
</tr>
</tbody>
</table>

**Table 31:** [Pro_{dem=ing, be (Pro_{wh=...})}clause] - Contingency table for construction & translation status, English
Triggers for this construction in translations are the elements *deshalb* ‘therefore’, *deswegen* ‘hence’, *so* ‘that way’, *und genau hier* ‘and exactly here’, *das ist der Grund* ‘this is the reason’ in the German originals. Thus, there are some overlaps between the German triggers of the construction in the English translations and expressions used in the German translations of the English demonstrative cleft construction. However, these structures are not always translated by the demonstrative cleft. For instance, *deshalb* is also sometimes translated as *accordingly, and precisely for this reason, therefore*.

While expressions such as *deshalb* and [*That/this is why*] are not classified as corresponding constructions due to rather limited aspects of structural similarity, this might change, if the English pattern acquires even a higher degree of fixedness. Currently the English construction is analyzed as a copular clause with the general structure [*NP V Clause*]. However, it shows signs of structural reduction, as the verb is often cliticized. If this process continues, the clause might get reduced to an expression typically mapped onto the grammatical function of a linking adverbial, in parallel to the German *deshalb*.

The form of the demonstrative cleft construction displays more variation in translations than in originals included in our dataset. Whereas all 20 examples in the originals correspond to the two variants [*That/This is why*] and [*That/This is what*], the seven patterns in the target texts have the form [*That/This is/was why/how/where*]. In translations there is not only variation between the two clause-initial demonstrative pronouns, but also in tense as well as in elements introducing a relative clause. For instance, the variant [*That was how*], which is present in a translation, is not among the ten most frequent patterns shown in Figure 24 above. This finding is somewhat unexpected, since previous studies suggest that the translations are typically more conservative than the originals (Baker 1996, 2004). In other words, it is assumed that the translations tend to reflect the prototypical patterns found in the originals, rather than show more variation. In our case increased variation in the translations could be explained by taking into account the source texts. For instance, in (128) the cleft in general and the element *how* in particular have been triggered by the lexical item *so* ‘that way’. Moreover, the past form of the copular in the translation corresponds to the past tense of the original sentence. Therefore, the use of the demonstrative clefs in translations could be said to reflect the translation property of *shining through*, even though,
as discussed above, German does not have any corresponding constructions that are likely to be translated into English by a demonstrative cleft.

(128) **GO:** So lief es in den Zeiten, als es der Konjunktur in Deutschland gut ging.

(GO_ESSAY_007 s98)

**ETrans:** That was how things worked when Germany’s economy was still going strong.

The analysis in this chapter has shown that, depending on the register, the demonstrative clefs in English vary in their degree of fixedness and the lexical items filling the individual slots. The analysis of the translation direction English-German has confirmed the findings by Johansson (2001) showing that the German translations are often characterized by a clause reduction. Moreover, our analysis has also considered the opposite translation direction. This analysis has indicated that the demonstrative clefs also appear in the English translations and, depending on the triggers in the originals, might result in variants that are less frequent in the English originals. Further research in this area should take into account larger datasets of the construction both in originals and translations, as well as trace a historical development between the correspondences of the lexical items such as deshalb and genau dies in German and the patterns [This/that is why] and [This/that is what] in English.

### 5.2.2 Further cleft, cleft-like and similar patterns

Apart from the demonstrative clefs discussed in 5.2.1, our data also contains several instances of *it*-clefs (*Spaltsätze*) and *wh*-clefs or pseudo-clefs (*unechter Spaltsatz, Sperrsatz*) (Prince 1978; Quirk et al. 1985; Collins 1991; Biber et al. 1999; Altmann 2009; Duden 2009), which are classified as containing the formal elements [NP V NP], [NP V PP] and [Clause V Clause].

An *it*-cleft could be illustrated through example (129). The present study follows the constructional analysis (Patten 2012) of *it*-clefs (cf. an overview of existing approaches in Hedberg 2000, 907-908; Patten 2012, 5-12, 71-119). This approach, similar to the extrapolation analysis, links this type of cleft to the more general copular construction. The initial pronoun *it* and the clause in the final position are assumed to form one unit, namely a discontinuous definite description realized through the formal category of a NP, in which the
relative *that*-clause postmodifies the pronoun *it* (Patten 2012, 72-78). In other words, (129) could be analyzed as containing the structure [NP V NP]42.

(129) **EO**: But [it]NP1 [was]V [Titan]NP2 [that arguably provided the most exciting discoveries.]NP1 (EO_POPSCI_001 s28)

One of the important functions of this construction is to draw the hearer’s/reader’s attention to the focused element (the NP *Titan* in example (129)) (Ahlemeyer and Kohlhof 1999, 17-18; Patten 2012, 79). At the same time, the construction can be used to present the information expressed through the cleft’s relative clause as presupposed (Ahlemeyer and Kohlhof 1999, 18; Schmid 1999, 132-134; Patten 2012, 80-82). This function can be used in academic discourse to achieve a higher level of objectivity (Biber et al. 1999, 963). However, the discussion by Collins (1991, 98-116) indicates that the notions of presupposition or givenness are not straightforward. Collins argues for a gradable concept of discourse givenness and distinguishes between different types of clefts depending on various combinations of the focused element and the relative clause with respect to the degrees of recoverability of information mapped onto these parts of the construction.

Furthermore, the cleft’s focused element is assumed to present an exhaustive list fulfilling the conditions described by the relative clause (Patten 2012, 82-85; Schmid 1999, 135-137). The construction could also have a contrastive function differentiating the clefted constituent from other groups of people or objects that are not characterized by the semantic properties ascribed to this focal element (Patten 2012, 85). From the constructional perspective, Patten (2012) shows how these functions of the *it*-cleft are systematically related to the more general copular construction.

Similar to *it*-clefts, *wh*-clefts are characterized by relative clauses that contain the presupposed information, which is typically not directly recoverable from the context but “has to be [...] ‘assumable’ as being there” (Prince 1978, 889). Moreover, the pattern can also be used to express the relation of contrast (Prince 1978, 890). The construction occurs only once in our sample of the English originals (see example (130)).

(130) **EO**: [What it implies]Clause [is]V [that knowing the molecular composition of an organism and how it changes in time is sufficient to compute, its form, since all that genes do is synthesize molecules in a particular order.]Clause (EO_POPSCI_008 s92)

---

42 Other recent analyses of *it*-clefts recognize a close semantic but not syntactic link between the pronoun and the relative clause (Hedberg 2000; Reeve 2011).
Previous research has suggested that the *it*-cleft is more frequent in English than in German (Fischer 2009). Moreover, while no claims on the overall frequency distributions are made with respect to the *wh*-clefs in the two languages, Gast and Wiechmann (2012) indicate that the English construction has a wider range of functions. However, it should be mentioned that the contrastive analysis reported by the authors is based mainly on translated texts. Since our sample contains only very few examples of *it*- and *wh*-clefs, these patterns are not considered in more detail here. The reader is referred to the previous studies on the translations of *it*-clefs in the direction English-German (Ahlemeyer and Kohlhof 1999; Doherty 1999).

Because *wh*-clefs are rather restrictive allowing only certain types of *wh*-clauses, these types of clefts are often substituted by patterns with noun phrases in the fronted position (Quirk et al. 1985, 1388-1389), as can be shown through (131). Since the *why*-clause cannot occur clause-initially, the noun *reason* postmodified through a finite clause is used instead.

(131) *[Why we decided to return]*_clause_ was that he was ill.
   [The reason we decided to return]*_NP_ was that/because he was ill (Quirk et al. 1985, 1389)

These alternatives to *wh*-clefs and similar structures with rather abstract nouns as subjects have been referred to as cleft-like sentences (Doherty 2001). The pattern consists of an abstract noun, which is typically pre- and/or postmodified, a copular verb and a clause. Since the construction is considered to be an alternative to a *wh*-cleft, it can be assumed that the two constructions have similar functions.

The original list of items occurring in this pattern includes such nouns as *person, reason, way, place* or *time* (Quirk et al. 1985, 1389), which could be used as alternatives to the *wh*-words, i.e. *who, why, how, where* or *when*. Doherty (2001) extends the NP slot to include other abstract nouns but does not specify whether there are any restrictions on the nouns allowed in this pattern. For instance, example (132), listed among cleft-like structures, contains *possibility* as the head noun of the sentence-initial NP functioning as the subject of the sentence. Other nouns occurring in this position include e.g. *purpose, strategy, approach, theory* or *surprise* (Doherty 2001). Moreover, in contrast to Quirk et al. (1985), Doherty considers only sentences with a finite or a non-finite clause as the subject complement.
(Doherty 2001, 608). However, it is not clearly specified what criteria have been used to operationalize this pattern.

(132) **EO:** [A second [possibility]_{head,noun}NP [is]V [that we need them all because each subtype has different properties.]_{Clause} (EO_POPSCI_010 s41)

It appears that the structure cannot be classified as a lower-level construction in its own right: taking into account that the subject and complement could be realized through different types of abstract nouns with pre- or postmodification as well as various clauses (or nouns) the pattern has a rather low degree of fixedness.

For the translation direction English-German, Doherty (2001) distinguishes three translation strategies applicable to the cleft-like structures. In the first two cases a verb from either the original subject or the original subject complement is used as the main verb of the target sentence instead of the copula. Both strategies are often associated with a reduction in the number of clauses. The third possibility is to translate the original copular sentence with a similar pattern in German (Doherty 2001, 610). While the present study does not draw a line between cleft-like and non-cleft-like sentences and, therefore, cannot further analyze these strategies with respect to the group of sentences discussed by Doherty, it does examine whether the strategies could be generalized to all copular sentences containing a clause.

Thus, below we consider the remaining instances of [NP V Clause] as well as the categories [Clause V NP], [Clause V AP] and [Clause V Clause] in both translation directions. The analysis does not take into account any types of clefts, i.e. demonstrative, *it*- and *wh*-clefts, as these could be analyzed as distinct constructions.

For the translation direction English-German 14 sentences with and 23 cases without construction shifts (all containing at least one clause in the originals) have been analyzed according to the number of clauses in the corresponding translations. For this analysis, the clauses on the level of the main clause have been considered. In other words, a formal structure of a copular construction involves a subordinated clause, when e.g. the subject is realized by a finite or a non-finite clause, whereas a noun modified by a clause is considered a NP. This methodological decision is due to the general focus of the present study on the main clause. Since the original copular sentences with the form e.g. [NP V_{be} Compl_{+Clause}] belong to the category [NP V PP], translations of an original clause through a PP, which has the internal structure of a pronominal adverb e.g. *darin* ‘therein’ followed by a clause are
classified as cases of clause reduction. However, the discussion of the number of clauses should take into account that in these cases it is even more questionable whether the level of complexity changes in any way. In three examples, an English copular verb has been translated through the prepositional verb bestehen in ‘consist of’. As mentioned above, prepositional objects have been shown to occur more frequently in German than in English (Čulo et al. 2012, 114), so that the translators comply with the preferences of the TL by shifting from a complement to a prepositional object. In this respect, both sentences contain argument-structure patterns typical of the corresponding languages and are both characterized by finite clauses, even though integrated into different formal units.

The analysis has shown that the number of clauses has been reduced in 41% of instances (15/37). Apart from the three sentences with the verbs bestehen ‘consist’ and sehen ‘see’ followed by prepositional phrases with integrated clauses, this group contains two examples in which the clause functioning on its own as the subject or the complement in the original has been attached to a noun in the translation as its postmodification. Moreover, in one example the clause is translated through a pronoun. 33% (5/15) of cases classified as containing clause reduction correspond to the first two translation strategies described for the cleft-like structures. In these examples, the main verb of the translation is extracted from a structure functioning as the subject or the complement in the original (cf. Doherty 2001: 610). For instance, in (133) the subject of the source sentence corresponds to an adverbial, even though the meaning is slightly changed, whereas the complement clause is upgraded to the main clause. In contrast, (134) illustrates that the original subject could be split. In this example, the semantic information corresponding to the subject of the original (the U.S. experience) is translated as the subject (die Vereinigten Staaten ‘the U.S.’) and the direct object (die Erfahrung, dass... ‘the experience, that...’), whereas the complement of the ST structure is turned into a postmodification of the noun Erfahrung. The main verb in the translation, namely machen ‘make’ is a collocate of Erfahrung. Therefore, to some extent the verb in the target sentence also corresponds to the subject in the original sentence.

(133) **EO:** [A corollary]_{Subj} has been [that proteins, in addition to their structural and enzymatic roles in cells, must be the primary agents for regulating the expression, or activation, of genes.]_{Compl} (EO_POPSCI_004 s9)
While the majority of sentences discussed above contain a finite clause in the original, the second major category of cases affected by clause reduction involves a nominalized *ing*-clause followed by a copular verb and a complement. The strategy is attested in 33% of instances (5/15), shown here in (135). These changes can be explained through the lack of the *ing*-participle form in German. Since the subject of the sentence is typically filled by a NP (Biber et al. 1999, 123) and the *ing*-participle can, in some grammatical contexts, be attributed to the categories of verbs or nouns (Biber et al. 1999, 67), the NP in German is certainly a close functional correspondent of the English nominal clause. Example (135) indicates that the nominalization of the *ing*-clause need not lead to an increase in the level of grammatical complexity. Though the level of explicitness may depend on the realization of semantic information through verbal or nominal structures (Halliday and Matthiessen 1999, 230-232), in this and similar examples the information is equally densely packed both in originals and the corresponding translations. This could be explained through the fact that the *ing*-clause, being non-finite, does not contain a subject. Therefore, the agent of the action, which could disappear in the course of a nominalization leading to a decrease in explicitness, is absent already in the original clause.

Another alternative translation of the *ing*-participles is the use of the German infinitive clause, discussed in more detail below. These translation pairs belong to a group without a change in the number of clauses. Moreover, as is illustrated through (136), the clause number can be counterbalanced if the main verb in the target sentence corresponds to an element
from the subject or complement in the original (thus reducing the number of clauses by one), whereas another element is translated through an adverbial realized through a clause (thus increasing the number of clauses by one). In the example shown here in (136), the change of the copula towards the prepositional object construction may have been due to the contrastive differences between the involved grammatical functions. However, the majority of the examples in this group, namely 73% (16/22) contain the [Subj V Compl] construction both in the original and the translation, in which one of the grammatical functions is realized by a finite clause.

(136) **EO:** [The lesson of my research]Subj is [that developments in each of these spheres may well rest on assumptions that are deeply mistaken.]Compl (EO_POPSCI_005 s153)  
**GTrans:** [Wie aus meinen Forschungen hervorgeht,]Adv beruhen die Entwicklungen in all diesen Bereichen wahrscheinlich [auf zutiefst falschen Annahmen.]Probj

In the opposite translation direction, of 29 German original sentences containing clauses, 22 examples belong to the category without construction shifts. In all of them together with four instances affected by shifts (90%, 26/29) the number of clauses was not reduced in the translations. In this translation direction, the majority of the patterns under analysis have been translated through a copular construction.

(137) **EO:** [Es] istV [richtig,]AP [daß [sic] sich auch in den Industriestaaten die Kluft zwischen denen, die von der Globalisierung profitieren, und denen, die die notwendigen Anpassungen nicht schaffen, erheblich vergrößert hat.]Clause (GO_ESSAY_009 s22)  
**GTrans:** [It] isV [true]AP [that in the industrialized countries, too, the gap has significantly increased between those who profit from globalization and those who do not succeed in making necessary adjustments.]Clause

For instance, in (137) both ST and TT sentences are examples of the pattern [Clause V AP]. In this particular sentence pair the clause is extraposed in the original and the translation, but this need not be the case. It could be assumed that no shifts with respect to the grammatical functions are introduced because the grammatical function of complement is typical of the target language, i.e. English. Moreover, it appears that English sentences are more likely to contain clauses filling the slots of grammatical functions on the level of the main clause. Fischer (2013, 249) suggests that this type of clausal complexity in English is likely to correspond to non-verbal phrases in German. This is in line with the finding that the
translation direction English-German contains more shifts towards nouns (Čulo et al. 2008, 50). We have seen that there are typically more cases of clause reduction in the translation direction English-German, and less in the opposite direction. This tendency can also be observed, if we have a closer look at the shifts into the patterns containing clauses. Of all these structures in the translated dataset (not considering different types of clefts) in 36% (15/42) of sentences there are more clauses in the English translations than in the corresponding German originals, compared to 12% (3/25) in the German translations. In other words, the translators seem to comply with the norms of the originals regarding the number of clauses on the level of the main clause.

Another interesting observation concerns the infinitive clauses in the German originals. In 33% of instances (4/12), the infinitive clause is translated through an ing-clause, as illustrated in (138). This translation strategy may involve minor changes in word order within the non-finite clauses, which are due to the contrastive differences between the languages: while in German the subordinate clauses are characterized by a verb-final position, in English the verb precedes the direct object (König and Gast 2009, 181). As mentioned earlier, the correspondence between the infinitive and the ing-participle clauses has been also detected in the opposite translation direction, even though in our data sample the English ing-clause corresponds more often to a NP in the German translations. This second strategy is also visible in the translations into English, in which we could find several examples of German nouns translated through an ing-clause (7%, 3/41). Therefore, at least with respect to the ing-clauses occurring on the level of the schematic argument structure construction [Subj V Compl], the two correspondences of the ing-clause indicated by Fischer (2013, 350) are also identified in our data. These are nominalizations and infinitive clauses.

(138) **GO:** [Seine Firma zu führen]Clause ist kein Job auf Zeit, sondern ein Lebenswerk, genauer: sein Lebenswerk. (GO_ESSAY_007 s41)

**ETrans:** [Heading a company]Clause is a life’s work or, to put it more precisely, their life’s work.

Based on the results reported in this chapter, it is possible to conclude that the translation strategies described with respect to the cleft-like structures are rather general and could be applied to various copular patterns containing clauses. Their frequency varies depending on the translation direction and partially on the type of clause in the original.
5.2.3 Translation shifts from [Subj V Compl]

The analysis in the preceding sub-chapters has considered a classification of the copular patterns based on their form. It was possible to identify three structures, namely [NP V NP], [NP V AP] and [NP V Clause] that are associated with either English or German. Only one of these, namely [NP V Clause] has a significant effect on the variable of construction shifts in the translation direction English-German. No patterns have shown a similar effect in the opposite direction. Therefore, the category of [NP V Clause] has been studied in more detail, more specifically the demonstrative clefts, classified as a lexically-filled construction. It could be shown that the slots of the construction are filled differently depending on the register. Moreover, two variants of the demonstrative clefts, namely [This/that is why] and [This/that is what] have been identified in our dataset to have different prototypical translation correspondents. The chapter has also briefly considered other types of clefts and concentrated on the remaining instances of the pattern [NP V Clause], generalizing their typical translations to all structures that contain clauses, both in the originals and translations. It has been suggested that the subject or subject complement are more likely to be filled by a clause in English than in German. Typically, translators observe this contrastive difference. They either introduce additional clauses when translating into English or reduce the number of clauses in the translations into German. However, it should be kept in mind that the clauses are then in some cases attached to other elements, most commonly to a preposition or a noun. These strategies, in addition to reducing the number of bare clauses in the sentence, could also result in crossing lines between e.g. a complement in the original and a prepositional object in the translation. Moreover, there are a number of crossing lines between some elements of the original clause (e.g. a subject) and an initial adverbial in the translation.

In the next step a logistic mixed-effect regression model (Baayen 2008, Winter 2013) was built to predict the effect of different factors on the binary nominal response variable of construction shifts from [Subj V Compl]. The results summarized above indicate that the variables of lower-level constructions and the number of clauses can be included as two predictors. However, similar to Chapter 5.1.6, not all of the lower-level constructions were found in all categories, e.g. the demonstrative wh-cleft does not have a corresponding construction in German. To avoid zero values, the specific constructions were not taken into
account in this chapter. The ordinal variable “no.clauses.main” has three levels (see Chapter 5.1.6 for details). Moreover, the binary nominal variables of language and register were considered in order to account for contrastive differences and possible variation between political essays and popular-scientific discourse. A random intercept for texts was included. In other words, the model examines the effect of the predictor variables introduced in Chapter 5.1.6 on the shifts from the abstract construction [Subj V Compl]. The results for the fixed effects are summarized in Table 32.

|                  | Estimate | Std.Error | z value | Pr(>|z|) |
|------------------|----------|-----------|---------|----------|
| (Intercept)      | -1.02    | 0.38      | -2.66   | 0.008    |
| languageGerman   | -1.55    | 0.30      | -5.14   | 2.81e-07 |
| registerPOPSCI   | 0.20     | 0.30      | 0.68    | 0.50     |
| no.clauses.main  | 0.59     | 0.26      | 2.30    | 0.02     |

Table 32: [Subj V Compl] - Summary of the fixed effects on the variable of construction shift

Thus, according to the model, the variables of language and number of clauses in the originals show significant effects on the variable of construction shifts. Also with this argument structure construction we could observe that in the translation direction German-English the probability of shifts decreases. In contrast, the probability of shifts correlates positively with the number of clauses in the [Subj V Compl] construction present in the originals. In other words, instances of the [Subj V Compl] construction containing more clauses are more likely to be translated through another structure. Since several cleft constructions discussed in this chapter realize one of the grammatical functions through a clause, the presence of clefts in our data sample is likely to contribute to these results. In contrast to Chapter 5.1.6, the p-value for the variable “register” is above the conventional levels of significance. It could be concluded that the lower-level constructions, both that have been identified and any additional patterns on different levels of abstractness present in the data, are less register-specific than the ones found in the [Subj V Dobj] construction.

The function vif.mer did not reveal any multicollinearity (languageGerman 1.0007, registerPOPSCI 1.001, no.clauses.main 1.002). Compared to the model reported in 5.1.6, the model accounts for more variation, as is shown by the goodness of fit test (C-value= 0.8, Dxy=0.6).
5.3 Subject Verb Prepositional Object

The third schematic argument structure construction analyzed in the present study is [Subject Verb Prepositional Object]. Prepositional objects refer to participants in a process, rather than provide information on place or time (Rice 1987, 93) and therefore, should be distinguished from adverbials (Fischer 2013, 270), which are also often realized through the formal category of PP. Even though the boundary between the two grammatical functions is rather fuzzy (Breindl 1989, 266), a clear set of criteria is required to identify the cases of the analyzed construction involving prepositional objects. For the purposes of our study, the category of prepositional objects has been limited to the following prepositions:

- English: about, after, at, for, in, of, on, to and with (cf. Fischer 2013, 268)
- German: an, auf, aus, für, gegen, in, mit, nach, über, um, unter, von, vor, zu and zwischen (cf. Breindl 1989, 11; Albert et al. 2003, 57; Fischer 2013, 268)

In addition, in accordance with the underlying definition applied in the CroCo corpus (Hansen-Schirra, Neumann and Kunz 2010), the prepositional objects also include the PPs introduced through by or von, which refer to the agents of the actions in the sentences written in the passive voice.

Moreover, certain PPs are not counted among the prepositional objects if their interpretation is closely linked to the typical meanings of the adverbials, such as local or instrumental. Breindl (1989, 28-35) suggests that a prepositional verb is typically associated with one specific preposition; when multiple prepositions are possible, there is usually no systematic meaning relation between the resulting prepositional objects. Therefore, if, for instance, the pattern [für ‘for’ NP] can be easily substituted through its opposite [gegen ‘against’ NP], the sentence it occurs in is not included into our data set (cf. Breindl 1989, 32; Albert et al. 2003, 57-60). Finally, a number of German verbs were excluded that had been explicitly mentioned as those that typically do not belong to the category of prepositional objects (cf. Albert et al. 2003, Appendix C2).

In English, the prepositional verbs are additionally differentiated from the phrasal verbs, i.e. cases where combinations of verbs and specific prepositions are considered to be completely lexicalized. While lexicalization is certainly a gradual process, I follow the annotation guidelines of the CroCo corpus (Hansen-Schirra, Neumann and Kunz 2010) in basing the

---

43 From now on, the construction is referred to as [Subj V Probl].
classification on the dictionary entries of phrasal verbs. In other words, if a pattern [V P] such as [take after] is listed as a phrasal verb in a dictionary, the following NP is annotated as a direct rather than a prepositional object. Fischer (2013, 274) suggests that a corresponding German category is a group of particle verbs.

With respect to the prepositional object construction, Rice (1987, 93) stresses semantic links presupposed between the individual clause elements: “Not only is the action directed towards this NP [i.e. the NP following the preposition], but potential or successful contact between it and the subject is implied”. A functional interaction between the verb and the preposition is supported by the fact that the choice of the preposition largely depends on the semantic category of the verb (Heringer 1988, 356-359; Fischer 2013, 269). Fischer (2013, 270) suggests that through its use in the prepositional object construction, a preposition obtains secondary meaning(s): for instance, the preposition über, when used as part of an adverbial of place typically means ‘above, over, on’, whereas it has the more abstract meaning ‘about’, when it is included into a prepositional object. However, Breindl (1989, 40-42, see also Rostilla 2007, 121) notes that while prepositions occurring in this construction do have certain semantic content, it is difficult, if not impossible, to reliably define this very abstract meaning, for instance because one and the same preposition typically occurs with various verb sub-classes. From the CxG perspective, it is also preferable to assign some meaning to larger constructions, in our case the abstract [Subj V Probj] construction, rather than identify a number of independent senses of prepositions (cf. Goldberg 1995, 9-10). Below we come back to this question during the discussion of more specific realizations of the general [Subj V Probj] construction.

Based on the shared set of grammatical functions, the English and German [Subj V Probj] constructions are considered corresponding. Both patterns contain prepositional objects, which are treated as participants in the referred events. These are introduced by fixed prepositions that are associated with specific verbs. However, the groups of sentences categorized as [Subj V Probj] differ across languages, as some combinations of verbs and prepositions are classified as phrasal verbs in English but not in German.

As can be seen in Table 33, in the translation direction English-German the construction under investigation is shifted in 17% (3/18) of sentences in the register of political essays compared to 67% (12/18) in popular-scientific texts. Looking at the proportion of shifts from
this construction in relation to all shifts, we could observe that these reach only 1% (3/495) and 2% (12/488) in the two registers and are, therefore, much lower than for the two argument structure constructions discussed in Chapters 5.1 and 5.2. In addition, in English the analyzed registers contain a smaller number of sentences corresponding to this argument structure construction: 1% (18/1293) in ESSAY and 2% (18/1048) in POPSCI. The frequencies are only slightly higher in German, where the construction accounts for 3% (40/1574 and 33/1167) in both registers. Nevertheless, the construction has been selected because of the reported differences in frequency distributions of the grammatical functions of complements and prepositional objects in English and German (see Chapter 5.2).

In the translation direction German-English, the construction accounts for more shifts both in political essays and popular-scientific texts. In the former register it is shifted in 78% (31/40) of all occurrences corresponding to 5% (31/657) of all construction shifts, and in the latter, 58% (19/33) of instances are changed to another structure accounting for 4% (19/442) of all shifts within this register. A comparison of shifts from the construction in relation to all occurrences of this particular structure in the register of political essays also indicates that the construction is less often changed in the translations into German ($\chi^2=16.51$, df=1, p-value=4.829e-05, see Table 34), whereas in POPSCI the difference is not significant ($\chi^2=0.11$, df=1, p-value=0.74, see Table 35).

<table>
<thead>
<tr>
<th></th>
<th>[Subj V Probj] shifts</th>
<th>[Subj V Probj] no shifts</th>
<th>All constructions shifts</th>
<th>All constructions no shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-G_ESSAY</td>
<td>3</td>
<td>15</td>
<td>495</td>
<td>798</td>
</tr>
<tr>
<td>E-G_POPSCI</td>
<td>12</td>
<td>6</td>
<td>488</td>
<td>560</td>
</tr>
<tr>
<td>G-E_ESSAY</td>
<td>31</td>
<td>9</td>
<td>657</td>
<td>917</td>
</tr>
<tr>
<td>G-E_POPSCI</td>
<td>19</td>
<td>14</td>
<td>442</td>
<td>725</td>
</tr>
</tbody>
</table>

Table 33: [Subj V Prob] and other clause-level constructions with and without shifts

<table>
<thead>
<tr>
<th></th>
<th>Shifts</th>
<th>No shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>EO-GTRANS_ESSAY</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>GO-ETRANS_ESSAY</td>
<td>31</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 34: [Subj V Prob] - Contingency table for construction shift & translation direction, ESSAY
<table>
<thead>
<tr>
<th></th>
<th>Shifts</th>
<th>No shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>EO-GTRANS_POPSCI</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>GO-GTRANS_POPSCI</td>
<td>19</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 35: [Subj V Probj] - Contingency table for construction shift & translation direction, POPSCI

<table>
<thead>
<tr>
<th></th>
<th>[Subj V Probj]</th>
<th>Other constructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EO</td>
<td>36</td>
<td>2305</td>
</tr>
<tr>
<td>GO</td>
<td>73</td>
<td>2668</td>
</tr>
</tbody>
</table>

Table 36: [Subj V Probj] - Contingency table for construction & language

It could be assumed that in the translations of political essays from German the translators tend to avoid the function that is more typical of the source language. Similarly, in the translations from English the translators could keep the element that is compatible with the target language norms. These results are supported by the frequency distributions in English and German originals. In our data sample the pattern [Subj V Probj] occurs significantly more often in the German texts ($\chi^2=7.09$, df=1, p-value=0.008, see Table 36). This, in turn, is in line with evidence from previous studies (Čulo et al. 2012, 114; Fischer 2013, 275).

The suggested difference in the use of the prepositional objects partly results from the presence of two groups of structures characteristic of one of the languages, namely the category of phrasal verbs in English and pronominal adverbs in German. Since it is possible that also some of the German prepositional verbs have reached a degree of lexicalization comparable to the one of phrasal verbs in English, the frequency distributions can be affected through this potential imbalance between the categorization of verbs in the two languages.

Moreover, I suggest that the prepositional objects occur more frequently in German because here the prepositional object can be realized not only through a PP. In 19% (14/73) of cases it corresponds to a category of pronominal adverbs, such as darin ‘therein’ or darauf ‘thereon’, which have only a very limited set of correspondences in English. Pronominal adverbs, similar to prepositional phrases in the function of prepositional objects, are linked to specific verbs and are thus lexically fixed. Their occurrences are associated with two functions. First of all, these elements are often used cohesively to make the links between the sentences more explicit (Kunz and Steiner 2013, 225). For instance, in (139) the pronominal

---

44 In the study by Fischer (2013) English phrasal verbs are included among the prepositional verbs.
adverb *hierin* ‘herein’ occurs in the sentence-initial position and creates a connection to the preceding text. Pronominal adverbs can also function as placeholders in complex sentences: whereas the prepositional verb *sorgen für* ‘care for’ would normally be complemented by a PP, the element *dafür* in (140) is added to integrate a subordinate clause (cf. Becher 2010b, 1332). In a way, also in this function of clause integration, the pronominal adverb is used cohesively adding a cataphoric link to the following clause. From a language comprehension perspective, Becher (2010b, 1332) describes the pronominal adverb in this function as a “device for managing the attention of the reader during the processing of the sentence”.

(139) **GO:** Derzeitige Umfragen zur Bewusstseinslage der Ostdeutschen verweisen auf einen Widerspruch: Während die sozial-ökonomischen Fortschritte durchaus anerkannt werden, wird die Lebensweise zu DDR-Zeiten wieder deutlich positiver eingeschätzt. [Hierin]Subj offenbaren sich ernst zunehmende Ängste vor Sozialrisiken und Ungewissheiten. (GO_ESSAY_003 s74-76)

**ETrans:** Current surveys of attitudes in eastern Germany reveal an interesting contradiction: although the socio-economic advances are fully acknowledged, people’s way of life during the GDR period is again being viewed in a far more positive light. [This]Subj exposes people's intense fears of social risks and uncertainties.

(140) **GO:** Wer mit Vigenere chiffriert, muß [sic] dafür sorgen, [daß [sic] der Empfänger das Schlüsselwort kennt.][Probj] (GO_POPSCI_008 s11)

**ETrans:** Anyone encoding with Vigenere has to make sure the receiver knows the keyword.

As shown in (139) and (140), in both types of cases the [Subj V Probj] construction is typically changed in translations into English. In fact, 13 out of 14 instances of pronominal adverbs present in the analyzed construction are affected by shifts. Seven of the 13 (54%) sentences shifted in the translation process exemplify the cohesive function. In five of them the initial pronominal adverb is replaced by another element realizing the link to the preceding stretch of text, such as a demonstrative pronoun or a conjunction. Another strategy is to merge the two sentences. Finally, the translator may also decide to simply omit any type of cohesive device. In its clause integration function, the prepositional object was translated through another function, such as a direct object, a complement or an adverbial in all six examples of construction shifts.
Coming back to the general contrastive difference in the frequency of prepositional objects in the two languages, we could observe that there are not only significantly more shifts from the [Subj V Probj] in the direction German-English ($\chi^2=14.7$, df=1, p-value=0.0001, see Table 37), but also more shifts into this construction for the opposite translation direction ($\chi^2=10.46$, df=1, p-value=0.001, see Table 38). In other words, the translators tend to introduce the element typical of the target language. This could also be demonstrated using the examples of pronominal adverbs in the German translations. In this text category there are 22 instances of [Subj V Probj] involving a pronominal adverb. With respect to the frequency distribution of pronominal adverbs, the translations are, therefore, characterized by a translation property of normalization. 19 of these sentences (86%) belong to the category of construction shifts: here the majority (63% or 12/19) corresponds to the clause integration function, where in seven cases the grammatical function of complement is shifted to a prepositional object. In the remaining examples the crossing lines involve the functions of subject, direct object and adverbials. Apart from pronominal adverbs in the clause integration and cohesive functions, the translations also contain one example of a question word corresponding to a pronominal adverb, namely *wozu* ‘why’. As can be seen in (141), the trigger for this element is the question word *why*, which functions as an adverbial in the original. The example shows that it might be insufficient to analyze individual items of the translation pairs in isolation, such as a crossing line between an adverbial and a prepositional object realized by a pronominal adverb. This local translation shift could be seen as a consequence from changing the subject of the sentence from the personal pronoun *we* to the NP *alle diese verschiedenen Alpha-Interferone* ‘all these different alpha-interferons’, which, in
turn, has to be combined with a different verb, in this case a prepositional verb requiring a
dependent object. Thus, alternatively this series of changes could be described from a CxG
perspective as a shift to the abstract argument structure construction [Subj V Probj]. While
one of the reasons for this shift might be a particular preference for the construction in the
translation, it is also possible that the reason behind this shift was the SL construction
involving a combination of the subject with human referents and a non-volitional verb. This
would be in line with the characteristic feature of German to focus on content rather than
persons (House 1997, 84).

(141) **EO:** Why do we need all these different interferons? (EO_POPSCI_010 s36)

**GTrans:** Wozu dienen alle diese verschiedenen Alpha-Interferone?

The relatively low number of examples representing the [Subj V Probj] structure in our data
set makes it difficult to further categorize the pattern, as the resulting groups would be too
small to allow meaningful analysis on a lower level of abstraction. A larger sample could be
classified based, for instance, on the semantic (sub-)groups of the prepositional verbs,
because these, as mentioned earlier, have been reported to govern the choice of the
preposition (Heringer 1988, 356-359; Breindl 1989, 40-42; Fischer 2013, 269; Rostilla 2007,
130). Rostilla (2007, 166) suggests that, compared to the transitive construction, the [Subj V
Probj] construction is less schematic. Rather than postulating an abstract [Subj V Probj]
construction with three abstract elements, he argues for a group of partially lexically filled
constructions with fixed prepositions. The meaning of each construction from this set is
assumed to represent a generalization over a set of semantically related verbs, which are
combined with the same preposition (Rostilla 2007, 168).

An example of such a construction would be a subclass of verbs taking the same preposition
such as [NP V lächeln/lachen/sich freuen über NP] ‘[NP V smile/laugh/be glad about NP]’ with the function of
“emotional reaction to a certain stimulus”. Rather than attempting to define different
meanings of the preposition über ‘about’, depending on its co-occurrence with different verbs
(cf. Fischer 2013, 269), the constructional interpretation stresses the fact that it is the

---

45 As shown in Chapter 5.1.2, the pattern [We need Dobj] occurs more frequently in the English part of our data set. However,
the discussion in 5.1.2 could not be directly related to the sentence pair in (141). First of all, the latter involves an additional
grammatical function, which has not been taken into account in the data samples analyzed in the present study. Secondly,
the pattern could well have different functions in the register of political essays investigated in 5.1.2 and the popular-
scientific texts.
combination of elements that creates meaning and is assumed to be stored as a separate form-meaning pairing.

In the present chapter, it was shown that the argument structure construction [Subj V Probj] is more typical of German than English. This contrastive difference is evident not only in the comparison between the English and German originals, but also in the number of shifts both from and into the analyzed abstract pattern. Since the frequency of construction shifts differs depending on the translation direction, it is assumed that translators tend to follow the preferences of the target language. For instance, in the translation direction English-German the shifts into [Subj V Probj] increase, whereas in the opposite direction the shifts from this construction are higher. This contrastive difference could be partly explained through the category of phrasal verbs distinguished only in English and the category of pronominal adverbs typical of German. Finally, following Rostilla (2007), I suggest that the construction under investigation could also be studied on a more concrete level identifying partially lexically filled patterns with restrictions on the verbal and prepositional slots.

5.4 Shifts between the analyzed argument structure constructions

As mentioned above, the three constructions selected for the present study are basic argument structures. As such, these patterns are closely linked. First of all, different classification decisions adopted in the course of the study have a direct impact on the categorization of certain sentences in our data sets. For instance, if a combination of a verb and a following preposition is listed as a phrasal verb, the sentence is classified as [Subj V Dobj] rather than [Subj V Probj]. Similarly, depending on a list of verbs that are categorized as copula ones, the sentence could be counted among the examples of [Subj V Compl] or [Subj V Dobj]. Moreover, if a sentence contains a form of the verb be followed by a word ending in –ed it could be classified as [Subj V Compl], [Subj V Probj] or [Subj V Adv], depending on whether the word is used as an adjective or a past participle (see Chapter 4.2.2). Secondly, as was shown in the preceding sub-chapters, the translation pairs often involve shifts between the three argument structure constructions under analysis. Focusing on the [Subj V Probj] construction and the contrastive difference involving this construction and the structure [Subj V Compl], it could be observed that among all shifts from [Subj V Probj] in the translations from German into English there are 38% (19/50) of changes into [Subj V Compl]. In the opposite translation direction, 32% (14/44) of construction shifts into the German
construction [Subj V Probj] in both registers correspond to the construction [Subj V Compl] in the English originals (see also Čulo et al. 2012, 114-115 for the discussion of shifts between complements and prepositional objects in three other CroCo registers). Similarly, the data contains examples of shifts between other pairs of the analyzed patterns. While one of the motivations for these changes may be simply the preference of a particular abstract pattern in one of the languages, other reasons are visible only at a lower level of abstraction taking into account more fine-grained semantic and/or syntactic classifications of examples.

The present corpus-based study of construction shifts has indicated that certain corresponding partially lexically filled constructions have different frequencies in English and German originals and translations. While the discussed constructions have certain restrictions on the individual slots, they are not completely fixed. It is thus possible to go one step further on the cline of abstractness and consider certain shifts on the level of lexically fixed constructions. As an example, the correspondences between \([V_{\text{be right/wrong}}]\) and \([V_{\text{haben Recht/Unrecht}}]\) \('[V_{\text{have right/wrong}}]'\) could be analyzed, illustrated through the second sentence in (142). As can be shown in this translation pair, the English and German patterns correspond semantically. Both mean that somebody has said or done something that is assessed by the speaker as right or wrong. While the English expression has the form [Subj V Compl], in German this meaning is realized through [Subj V Dobj]. Therefore, the presence of these lexically filled patterns in the originals is very likely to lead to construction shifts.

(142) **EO:** Some of our trading partners have expressed concern that this heralds a protectionist turn in U.S. policy. [They] \([\text{are}]_{\text{V Compl}} \text{[wrong.]}\) (EO_ESSAY_016 s8-9)

**GTrans:** Einige unserer Handelspartner haben sich besorgt geäußert, dies kündige eine protektionistische Wende der US-Politik an. [Sie] \([\text{haben}]_{\text{V Dobj}} \text{[Unrecht.]}\)

The queries for the lemmas *right/wrong* and *Recht/Unrecht* in all eight registers of the CroCo corpus have identified ten instances of the pattern in the English originals and three examples in the German originals. These sentences either form the analyzed [Subj V Dobj] or [Subj V Compl] constructions, occur with an additional adverbial or as part of the subordinated clause or coordinated sentence. From the extracted sentence pairs containing one of the expressions in the English and German originals, 85% of examples (11/13) show correspondences between the English and German expressions \([V_{\text{be right/wrong}}]\) and \([V_{\text{haben Recht/Unrecht}}]\). Moreover, the queries helped detect similar lower-level patterns, namely
[V bekommen Recht] and [V behalten Recht], each represented by single examples in the corpus. These German transitive patterns are also turned into copular ones in the English translations, as is shown in (143), in this case with an adverbial as an additional grammatical function in both the source and the target sentences.

(143)  **GO:** Skeptiker haben leider Recht behalten. (GO_SHARE_006 s175)

  **ETrans:** Unfortunately, the skeptics were proven right.

Another lexically-filled pair of patterns is shown in (144). The English original belonging to the schematic construction [Subj V Compl] includes the lower-level pattern [V be part of], which is translated through the expression [V gehören zu] ‘[V belong to]’ resulting in the clause-level construction [Subj V Probl]. In this example the preposition zu ‘to’ is part of the pronominal adverb dazu ‘to it’, which establishes an anaphoric link to the preceding sentence (see also Chapter 5.3). Both English and German expressions mean that something or somebody belongs to a group.

(144)  **EO:** The European budget cannot be separate from this debate. [It] subj [must be] v [part of it.] compl (EO_ESSAY_003 s37-38)

  **GTrans:** Diese Debatte kann nicht losgelöst vom EU-Haushalt geführt werden. [Er] subj [gehört] v [dazu.] prob

Originals belonging to all registers of the CroCo corpus have been queried for the lemmas part and gehören. Of theses only the examples with occurrences of the expression part of in the function of subject complement and the prepositional verb gehören zu have been analyzed further. The final sample includes 29 sentences in the English originals and 87 instances in the German originals. Among these, there are also examples with additional grammatical functions and variants of the pattern integrated into the subordinate clause or occurring in a coordinated sentence. In the translation direction English-German only three instances (10%) of the pattern [V be part of] correspond to the structure [V gehören zu]. In addition, two examples are translated through the particle verbs dazugehören ‘belong to it’ and angehören ‘belong to’. In the majority of examples in this translation direction the grammatical function of complement is translated without crossing lines. The expression [V be (Adj) part of] corresponds to [V sein/bleiben/werden (Adj) Teil/Bestandteil NP GEN/von] ‘[V be remain/become (Adj) part NP GEN/of]’ in 59% (17/29) of instances. Also in the translation direction German-English the analyzed structures, namely [V be part of] and [V gehören zu],

232
correspond only in 9% (8/87) of cases. In addition, there are two examples of the idiomatic expression \(V_{\text{be is part and parcel of}}\) in the English translation belonging to two different texts in the register of prepared speeches. In this translation direction, however, a number of other translation correspondences lead to translation shifts, because they are realizations of the English copular construction. These include \(V_{\text{be among}}, V_{\text{be one of}}, V_{\text{be an aspect of}}, V_{\text{be a member of}}\) and \(V_{\text{be features of}}\) and account for 23% (20/87) of the translation pairs. Other frequent translation strategies comprise the prepositional verb \(\text{belong to}\) (14%, 12/87), i.e. the immediate equivalent of \(\text{gehören zu}\), and the verb \(\text{include}\) (11%, 10/87).

Taking into account that the verb \(\text{include}\) is followed by the direct object, also these 11% involve crossing lines between the words and the grammatical functions. Therefore, the English translations of the analyzed pattern \(V_{\text{gehören zu}}\), which is integrated into the variants of the construction \([\text{Subj V Probj}]\) in the German originals, are frequently characterized by translation shifts as the prepositional object often corresponds to the complement or the direct object.

These examples indicate that while some possible reasons for construction shifts could be found on a relatively high level of abstractness, others require an analysis of the lexically filled expressions. Such lower-level idiomatic schemas are assumed to be even more language-specific than the partially filled argument structure patterns discussed in the present study. Therefore, these should also be taken into account in the analysis of potential reasons for shifts between the argument structure constructions.

5.5 Conclusion

The present chapter has considered three general argument structure constructions, namely \([\text{Subj V Dobj}], [\text{Subj V Compl}]\) and \([\text{Subj V Probj}]\). It has suggested that potential correlations between the SL construction and translation shifts could be identified mainly on the lower levels of abstraction. Therefore, the abstract constructions have been classified further. The initial classification has been based either on a contrastive difference (in the case of the \([\text{Subj V Dobj}]\) construction) or on the formal categories that realize the analyzed grammatical functions (in the case of the \([\text{Subj V Compl}]\) construction). Despite certain restrictions that have been imposed by the boundaries of the individual categories (see Chapter 7), this approach has allowed us to identify a number of more specific partially lexically filled constructions. As can be seen in Figure 25, the more specific constructions of the general
transitive construction are $[\text{NP}_{\text{research}} \ V_{\text{show}} \ \text{Clause/NP}]$, $[\text{NP}_{\text{government}} \ V_{\text{feel/think}} \ \text{clause/NP}]$, $[\text{NP}_{\text{human}} \ V_{\text{speak}} \ \text{Clause/NP}]$ together with even more fixed constructions $[\text{Es gibt} \ \text{NP}]$ and $[\text{Subj} \ V_{\text{hhaben}} \ \text{Recht/Unrecht}]$. With respect to the copular construction, the demonstrative cleft construction, represented by the formula $[\text{Pro}_{\text{dem-sing be}} \ (\text{Pro}_{\text{wh...}})\text{clause}]$, and the more specific constructions $[\text{Subj} \ V_{\text{be right/wrong}}]$ and $[\text{Subj} \ V_{\text{be part of NP}}]$ could be identified in our data sample. The set of patterns related to the general copular construction also includes the it- and wh-cLEFTs, which have not been discussed in detail here. Finally, due to its low numbers, the instances of the prepositional object construction have not been systematically classified. However, based on the previous literature, the partially lexically filled construction $[\text{Subj} \ V_{\text{lächeln/lachen}} \ \text{über NP}]$ has been mentioned. Moreover, the data contains a number of shifts from and to the more lexically fixed construction $[\text{Subj} \ V_{\text{gehören zu NP}}]$.

On the general level of abstract argument structures, it could be observed that the translation direction plays an important role in the occurrence of shifts both from $[\text{Subj} \ V \ \text{Dobj}]$ and $[\text{Subj} \ V \ \text{Compl}]$. In addition, the variable of register has an effect on the former, and the number of clauses in the originals plays an important role in the latter group of shifts. The study has confirmed the contrastive difference discussed in the previous literature according to which English has a higher frequency of complements, whereas the prepositional objects are more typical of German. With respect to the prepositional objects, this difference could be partly
explained through the group of phrasal verbs specific for English and the group of
pronominal adverbs typical of German.

The discussion of individual lower-level constructions and their effects on translation shifts
has shown that most of the patterns that are associated with one of the languages in our data
set have not resulted in as many shifts as could be expected considering the initial
assumption of the present study, according to which the translators are likely to select
constructions that are more entrenched in the target language to make the translations sound
more natural. Instead, the corpus-based analysis reported in this chapter indicates that in
terms of translation properties, there is a complex interplay of both normalization and
shining through of the SL (cf. also Teich 2003). In other words, some instances of the
constructions, which are more typical of the source than the target language, are indeed
replaced by other patterns, while in other cases the translator selects a corresponding
construction in the process of translation. This could be explained by taking into account
additional factors, such as register. It has been shown that the same register in English and
German may be characterized through different constructions, for instance because the
authors of the texts in the register of political essays differ with respect to their involvement
into the political affairs. This feature might be either a consequence of the particular selection
of texts for the register of political essays in the CroCo corpus or a more general tendency in
the corresponding registers within the two languages.

Finally, the study has briefly touched upon a number of lexically fixed expressions that
function as specific instances of the general argument structure constructions. If such a
construction has a very fixed translation, which involves a crossing line between words and
grammatical functions, the construction shift is likely to be a consequence of these translation
correspondences rather than other factors.

Among further possible reasons for translation shifts and/or their effects I also considered
changes in the level of impersonality and potential changes in the information structure. The
majority of the discussed lower-level constructions have been shown to be more typical of
English and, therefore, the translation direction English-German has been considered in
more detail. In the English originals the constructions are typically realized through the
canonical SVO order, while their corresponding translations into German try to keep the
information structure stable. This process results in changes of the order of grammatical
functions and remapping of the semantic content onto different functions. In contrast, in the opposite translation direction, as has been shown with respect to the quotative construction, more word order variation present in the originals results in more changes of the information structure during the process of translation.

To complement this corpus-based analysis of construction shifts, the next chapter reports on a pilot translation experiment. Its analysis concentrates on the cognitive aspects connected to the translation of constructions.
6. Cognitive processing during the process of translation

The previous chapter discussed corpus-based evidence for the contrastive differences in the use of the corresponding constructions in English and German, as well as the tendencies to change or keep certain constructions in the translation process. While the frequency distributions obtained through these analyses could be used as the basis for hypotheses on the cognitive entrenchment of these structures, it should be considered that this is only indirect evidence. Therefore, in order to get more reliable insights into the nature of the analyzed phenomena it is necessary to combine different sources of data. Gilquin and Gries (2009, 11) suggest that corpus and experimental work could be used to verify assumptions formulated based on one of these sources of data. For instance, a corpus study functions as a validatee, if its results are subsequently validated through an experiment. The two data sources could provide converging evidence and thus strengthen the research conclusions (cf. Gries, Hampe and Schönefeld 2005, 2010). The corpus and experimental methodology could also be combined to answer related questions that require different approaches. In this case the methods have equal footing as they complement each other. Finally, the corpus data can also be used for stimulus composition to ensure the ecological validity of the experiments (Gilquin and Gries 2009, 11).

In translation studies, data triangulation typically refers to the combination of different methods used in process-based studies (Alves 2003). Such methods as eye-tracking, keystroke logging and retrospective protocols (see below for more details) are applied to obtain a variety of measures on the complex abstract feature of cognitive processing and are in a way validating each other. Corpus-based findings can be added to complement the picture by providing information on contrastive and register features that might also play a role in the processing of information during the translation process (Hansen 2003; Alves et al. 2010).

Therefore, based on the results of our corpus-based study, research questions that should be further studied from an experimental perspective and our hypotheses are specified in Chapter 6.1. The experiment design is described in Chapter 6.2, followed by a report on the results of this pilot experiment in Chapter 6.3. Finally, Chapter 6.4 provides some concluding
remarks and discusses how the evidence from the corpus-based and experimental studies fit together.

6.1 Research questions and hypotheses

The present chapter concentrates on one partially lexically-filled construction identified among the transitive constructions typical of English, namely [NP\text{research} V\text{show} that-clause/NP]. The corpus-based investigation reported in Chapter 5.1.1 has shown that this construction occurs more frequently in the English than in the German originals. This difference has been observed not only in the main data sample taken from the CroCo corpus, but also in the reference corpora BNC and DWDS. Moreover, the construction is one of the few that is shifted to another structure slightly more often than is kept in translations. Therefore, it is suggested that the construction deserves a more detailed analysis within a translation experiment.

English and German have corresponding constructions combining the form [NP\text{research} V\text{show} that-clause/NP] with the function of reporting the results of the studies. Based on the number of occurrences in the originals belonging to the two languages, it is assumed that the German construction is less entrenched in the speakers’ minds. As mentioned in Chapter 3, previous research on translations of the English resultative construction into Spanish has suggested that since the source language construction does not have a corresponding construction in Spanish, its translation requires additional cognitive processing reflected in the eye-tracking measures used in the study (Rojo and Valenzuela 2013). Similarly, an investigation by Vandepitte and Hartsuiker (2011) involving translation experiments from English into Dutch has established that combinations of subjects mapped onto inanimate referents and verbs that typically occur with human agents require additional cognitive effort during their translation. This effect is explained through a contrastive difference between English and Dutch. Similar to German, there are more restrictions on the mapping of semantic roles onto the subject in Dutch than in English. In other words, the analyzed structure, which roughly corresponds to the general transitive pattern [Subj\text{animate} V\text{volition} Dobj] identified in Chapter 5, is claimed to be more typical of English than of Dutch. Therefore, translations of this pattern into Dutch seem to pose additional problems to students who participated in the experiment as indicated by longer translation times (Vandepitte and Hartsuiker 2011).

Taking into account the results of these studies, the present study assumes that the
translation of the English construction [NPresearch Vshow that-clause/NP] into German is cognitively more effortful than the translation of the pattern [NPhuman Vshow that-clause/NP], in which the verbs of showing are combined with human agents\textsuperscript{46}. Moreover, based on the results of the corpus-based study reported in Chapter 5.1.1, the analyzed construction is expected to be more often shifted to another structure than the control items.

Another research question that could be analyzed with the help of the experimental methods concerns the role of constructions in the process of translation. It is assumed that the process-related concept of the translation unit (introduced in Chapter 2) could be used to investigate whether translators operate with partially-lexically fixed expressions, which are more specific realizations of complex argument structure constructions.

As discussed in Chapter 2, previous studies have shown that during the translation process and text production in general the length of pauses is higher between larger grammatical units: for instance, pauses between phrases are longer than between words, followed by even longer pauses between clauses and sentences (Schilperoord 2001: 83, Immonen 2006: 332, Immonen and Mäkisalo 2010: 57-58). Moreover, based on Bolinger (1965), Schilperoord suggests that clauses are characterized by the so-called temporal modification: the number of options during ongoing text production seems to decrease within clauses leading to a decreased level of uncertainty, as reflected by a decrease in mean pause times as well as pause time variances and their increase at the clause boundaries. He concludes that the clause could be seen as a basic processing unit (2001: 86-87), at least for monolingual text production.

This observation could be linked to the translation of the argument-structure construction [NPresearch Vshow that-clause/NP], more specifically to the part of the main clause [NPresearch Vshow that], i.e. a segment prior to the subordinate clause boundary. On the one hand, since the pattern [NPresearch VPshow that-clause/NP] is considered as a construction in its own right, stored and processed as a unit, the whole construction could be expected to form a

\textsuperscript{46} It should be stressed that the present study is certainly related to the investigation by Vandepitte and Hartsuiker (2011) in terms of the underlying assumption based on different realizations of the grammatical function of subject in English, on the one hand, and Dutch and German, on the other. However, our study does not deal with all instances of [Subjanimate Vvolition Dobj] but focuses on its specific variant [NPresearch Vshow that-clause/NP]. While the former is interpreted as one of the rather general transitive patterns, the latter is considered to be a lower-level construction in its own right (see Chapter 5.1.1 for a more detailed discussion). In this sense the present study is more comparable to the research by Rojo and Valenzuela (2013) as the difficulties that the translators may encounter with translations of these structures are more readily interpretable as instances of constructional resistance (cf. Szymańska 2011) or constructional mismatch (cf. Rojo and Valenzuela 2013).
translation unit more often than the control structure [NP\textit{researcher} V\textit{show} \textit{that}]. On the other hand, as the construction [NP\textit{research} V\textit{show} \textit{that}-clause/NP] is less typical of the target language and its translation is likely to cause additional cognitive processing, it is also possible that the critical items are more often split into smaller processing chunks during the process of translation than the control items. Therefore, no directed hypothesis can be formulated at this point of analysis. A comparison of translation units formed during the translation of each type of stimuli is required.

6.2 Experiment design

The pilot experiment carried out in the context of the present study combines the methods of eye-tracking, keystroke logging and retrospective protocols. This particular constellation of methods has been used in previous studies (e.g. Alves, Pagano and da Silva 2010; Alves et al. 2010) as the individual methods are argued to complement each other. For instance, it has been suggested that the pauses in text production, used as a keystroke logging measure related to the abstract variable of cognitive processing, can be interpreted in a more meaningful way, if they are combined with eye-tracking measures (e.g. Jakobsen 2011, 37-38). The values for fixation duration, fixation count, regressions or pupil dilation (e.g. Pavlović and Jensen 2009; Rojo and Valenzuela 2013) are often analyzed as concrete indicators of cognitive effort. Furthermore, the interpretation of results obtained from these online indirect observations could benefit from retrospective protocols, which belong to offline direct methods. In this case information on the translation process is gathered after the final version of the translation is produced (Spelman Miller 2006, 24). It relies on the introspection of the experiment participants. Therefore, while this procedure does not interfere with the translation process itself, it can provide important information on certain aspects of the translator's line of thinking not available in other types of data.

As mentioned above, the present chapter reports on a pilot experiment conducted to test the experiment design and indicate tendencies, which should be verified in a larger study with professional translations. At this initial research stage two students of English linguistics took part in the experiment. Their translation experience was limited to a lecture on contrastive linguistics, which involves some theoretical foundations of translation studies, or a seminar on translation studies. Therefore, the results do certainly reflect a lack of professional translation training and translation experience, as these factors have been shown to have an
effect on the translation process and product (e.g. Englund Dimitrova 1995; Carl and Kay 2011). In Chapter 6.3 we return to the discussion of how the results might change when professional translators participate in this experiment. However, both participants are native speakers of German, so that the experiment is expected to reflect the lower level of entrenchment of the German construction despite the limitations of the study mentioned above.

For the experiment the remote eye tracker Tobii TX300 running with the software Tobii Studio (Tobii Technology 2012) and the keystroke logging software Translog II (Carl 2012) were used. The material involved 16 short texts (see Rojo and Valenzuela 2013, Vandepitte, van Assche and Hartsuiker 2014 who used similar experiment design), each consisting of three sentences. In six of these texts, the second sentence corresponded to a critical item, i.e. the analyzed construction [NP\_research V\_show that-clause] (see example (145)), whereas in six other cases the second sentence was a control item, i.e. the structure [NP\_human V\_show that-clause] (see example (146)).

(145) It should be admitted that some animals possess some power of imagination. [Recent studies]NP\_research [have shown]VP\_show [that at least the higher primates probably have dreams when they sleep.]that-clause As to what their, and our, dreams mean – that is a matter of guesswork.

(146) Chimpanzees are very intelligent animals. [Some researchers]NP\_human [have shown]VP\_show [that chimps can create artworks and understand the concept of language.]that-clause For instance, humans and chimps can use sign languages to communicate with each other to a reasonably advanced level.

Moreover, four partially modified stimuli designed in the project Pro-Eyes\textsuperscript{47} were used as fillers. The twelve main texts were based on authentic corpus examples extracted from the CroCo and the BNC registers of popular-scientific texts. However, the stimuli were modified to ensure better control over confounding factors. The second sentences corresponding to critical or control items have comparable length both in terms of characters and words. However, since it was not possible to keep these values completely stable across twelve texts, the eye-tracking measures are normalized per ten characters\textsuperscript{48}. Moreover, the structure of

\textsuperscript{47} Stella Neumann and Piritta Pyykkönen. Universität des Saarlandes Anschubfinanzierung.

\textsuperscript{48} Normalization per single character often resulted in values below one, which are problematic when looking at fixation count. Since the analyzed stimuli consisted of 23-35 characters, the stimuli were normalized per ten characters.
the initial NP is similar across all instances of the analyzed patterns. The noun referring to research or a human agent is premodified through such elements as *some, several, recent* or *preliminary*, which are among the frequent premodifiers identified for the construction under investigation, namely \[\text{NP}_{\text{research}} \ V_{\text{show}} \ \text{clause/NP}\], in the corpus study reported in Chapter 5.1.1. Moreover, in all stimuli the verb of showing is followed by a *that*-clause. To ensure that all stimuli sound natural, the modified versions of texts were controlled by a native speaker of English. The instances of the patterns used in the experiment are shown in Table 39. During the experiment the stimuli (critical and control items) and fillers were presented in a random order, which was kept constant across the participants (see Appendix II for the complete texts in the order used in the experiment).

<table>
<thead>
<tr>
<th>[\text{NP}<em>{\text{research}} \ V</em>{\text{show}} \ \text{clause/NP}]</th>
<th>[\text{NP}<em>{\text{researchers}} \ V</em>{\text{show}} \ \text{clause/NP}]</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Many studies have shown that</em></td>
<td><em>Some researchers have shown that</em></td>
</tr>
<tr>
<td><em>Recent studies have shown that</em></td>
<td><em>Some researchers have suggested that</em></td>
</tr>
<tr>
<td><em>Recent studies have shown that</em></td>
<td><em>Several researchers suggest that</em></td>
</tr>
<tr>
<td><em>Preliminary studies have shown that</em></td>
<td><em>But social scientists reveal that</em></td>
</tr>
<tr>
<td><em>Previous studies suggest that</em></td>
<td><em>Demographic historians have shown that</em></td>
</tr>
<tr>
<td><em>New evidence suggests that</em></td>
<td><em>Several scientists have shown that</em></td>
</tr>
</tbody>
</table>

Table 39: \[\text{NP}_{\text{research}} \ V_{\text{show}} \ \text{clause/NP}\] & \[\text{NP}_{\text{researchers}} \ V_{\text{show}} \ \text{clause/NP}\] used as stimuli

The experiment consisted of three parts. After the meta-information on the participants was collected, the first task was to re-type a short German text presented in the upper part of the *Translog II* window. This copy test is designed to assess the average typing speed of the participants and to let them get used to the experiment setup. In the second step, the participants were asked to translate the short texts described above from English into German (into their native language). The translation brief contained an instruction to translate the texts for a popular-scientific magazine. Only one text was visible at a time. Each text was presented in the upper half of the *Translog II* window in Arial Standard, font size 20 with double spacing, whereas the lower half was used to type in the translation. Care was taken to present the analyzed part of stimulus in one line because line breaks are problematic for the eye-tracking analysis of the data. However, as a result, in some cases a new paragraph had to be introduced after the first sentence making the text appear more fragmented. This was reflected in several corresponding translations where participants started new
paragraphs after the first and the second sentences. After the translation was completed, the next text became visible. The participants were allowed to use the bilingual online dictionary leo⁴⁹ as the only translation resource. During the third part of the experiment the participants were asked several questions concerning certain highlighted regions in a number of selected texts used in the experiment. The regions corresponded either to a (part of a) stimulus or a filler. This guided retrospective protocol aimed at identifying further potential comprehension problems or translation difficulties related to the critical and control items (see Appendix III for a complete list of questions and the corresponding highlighted regions).

6.3 Results and discussion of the pilot experiment

In the first step the eye-tracking data is used to determine the cognitive effort required for the translation of the construction [NPresearch Vshow that-clause/NP] and the related pattern [NPhuman Vshow that-clause/NP]. The analysis of the eye-tracking measures concentrates on the areas presented in Table 39 rather than the whole sentences, as the subsequent clauses would introduce more lexical and structural variation into the analysis. The analyzed regions were selected as our areas of interest (AOIs) in the software Tobii Studio. The amount of cognitive effort was operationalized through the fixation duration and fixation count within the AOIs. These values were extracted from Tobii Studio after the time segments with the active leo window had been excluded from the calculations. This step is necessary to prevent the fixations within the leo window adding up to the fixations within the AOIs. The eye-tracking measures have been normalized per ten characters⁵⁰. In the first analysis, we examine the effect of the pattern used in the English originals (critical vs. control items) on the behavioral measures of the fixation duration and fixation count, which are then linked to the level of cognition. This link is based on the eye-mind assumption, according to which the linguistic structures are processed while they are fixated (Just and Carpenter 1980), even though this might not always be the case (e.g. Carl and Jakobsen 2009). Therefore, ideally the eye-tracking evidence should be compared to data from other sources, such as keystroke logging.

⁴⁹ Available at http://dict.leo.org/ende/index_de.html.
⁵⁰ The values corresponding to the total number of fixations and the fixation duration on a particular AOI have been divided through the number of characters comprising this AOI and multiplied by ten. As a result, the normalized values are directly comparable irrespective of the minor variations in the number of characters within each AOI.
The interaction between the variables of fixation duration/count and constructions is investigated using the t-test for independent samples (Gries 2009). Since the test’s assumptions of normal distribution and homogeneity of variance are not met, the dependent variables have been log-transformed. The distribution of values for both measures over the total of 24 observations, corresponding to twelve critical and twelve control items, is shown in Figure 26. As expected, the means for the number of fixations and the total fixation duration are higher for the critical than control items, even though only the p-value for the effect of constructions on the fixation count is below the level of significance (fixation count: mean critical=2.2, mean control=1.6, t=-2.4, d=21.21, p-value=0.03; total fixation duration: mean critical=0.837, mean control=0.236, t=-1.98, df=21.99, p-value=0.06). Therefore, the eye-tracking data supports the hypothesis that the translation of the construction [NP research V show that-clause/NP], which is more typical of the source language, involves additional cognitive effort.

As mentioned above, to strengthen the validity of the analysis (Grésillon and Perrin 2014), the results of the eye-tracking investigation are further tested with the help of the keystroke logging data collected during the experiment. In this part of the study the cognitive effort is operationalized through the time required for the production of the translations of the source language structures under analysis (Vandepitte and Hartsuiker 2011). Similar to the eye-tracking analysis, the time is normalized for the production of ten keystrokes. In other words, to account for the fact that the target language structures are of different length and the
translation might involve a number of revisions, the average time per ten keystrokes (cf. Rautenberg 2011, 17, 26) is analyzed as the dependent variable. This measure is dependent on pauses occurring during text production, which are also often associated with the cognitive effort (e.g. Dragsted 2005).

To determine the average time per ten keystrokes, the XML files generated by the software *Translog II* were imported into *Excel*. The overall time required to produce the target text structures is calculated for the stretch of the process text between the full stop or the last letter press/backspace preceding the targeted patterns and the press of the second *s* in the word *dass*, which corresponds to *that* in the original. For instance, in the first translated text produced by the first participant the critical item *Many studies have shown that* corresponds to the structure *Viele Studien haben gezeigt, dass*. Thus, the time when the full stop preceding *Viele* *'many'* was typed is subtracted from the time corresponding to the last letter *s* and divided by the total number of keystrokes including deletions and navigation keys as well as the mouse movements that were used to produce this text string. The time which had been used to consult *leo* was considered as well, since also these pauses reflect the increased cognitive effort during translation. Similarly, the average time per ten keystrokes in each copy test was determined by subtracting the time of the first key from the time of the last key and divided by the total number of events, i.e. all text production or revision keystrokes as well as mouse movements. Only the initial and final events (start recording and the initial mouse movements/stop recording) were not taken into account, since the pauses between these events and the key presses are used, among other things, for the initial and final orientation (e.g. to search for the “stop recording” button).

The boxplot in Figure 27 shows the distribution of the data. As above, the t-test for independent samples is applied to test whether the occurrences of the analyzed partially lexically filled construction have an effect on the average time per keystroke. In contrast to the eye-tracking data, no significant difference can be observed between these two variables.

---

51 The calculations considered only the keys and mouse movements related to the production of the analyzed linguistic structures during the first draft. In other words, if the participant stopped working on the segment and returned to it at a later point after working on other parts of the target text (even before producing a complete version of the translated text), the time required for the revision of the segment was not added to the time of the first draft. This was done to ensure the comparability of the measure across all stimuli: the total average time for the translation of the stimulus affected by later revisions would be decreased, even though it is not clear whether the translator was in fact thinking about the segment of interest during the production of subsequent stretches of text. Therefore, the time associated with the production of the revised version may reflect not the entire cognitive effort invested in the process.
even though the mean of the log-transformed values for the critical items is slightly higher (mean critical=1.691 seconds, mean control=1.365 seconds, t=-1.68, d=21.91, p-value=0.11). From the perspective of the individual participants, the results for the first participant are characterized by a significant interaction between types of stimuli and average time per keystroke. In this case the mean value for the control items, namely 1.102 seconds, is slightly lower than the time observed in the copy test of this participant, corresponding to 1.172 seconds\textsuperscript{52}, whereas the mean for critical items is higher reaching 1.753 seconds. In contrast, the second participant produced both types of stimuli with the same speed (1.628 seconds), which was higher than in the copy test (0.663 seconds). This more detailed comparison indicates certain variation across the participants. A subsequent experiment with a larger group of participants is required to further investigate the time needed to produce the translations of the different stimuli compared to the average typing speed of each participant. Additional translator-related variables might help explain this variation.

![Boxplot for the effect of stimulus type on average time per keystroke](image)

**Figure 27: Boxplot for the effect of stimulus type on average time per keystroke**

Based on the results related to the eye-tracking and keystroke logging measures of cognitive processing, the first assumption could be only partially confirmed. While in the analyses discussed above the values for the critical items are to some extent higher than for the control items, this difference is not significant in the case of the average time required to produce the analyzed structures. However, this result has been shown to vary depending on the

\textsuperscript{52} To make the numbers more comparable, the value for the average time per ten keystrokes obtained from the copy test has been also log-transformed.
participant. All behavioral variables related to the first participant indicate that the translation of the construction [NP\textsubscript{research} V\textsubscript{show} that-clause/NP] is connected to increased cognitive effort. This result is further supported by the dictionary look-ups of this participant. The verb \textit{suggest} has been looked up in examples (148) and (150), but not in (147) and (149). Whereas the former are both instances of the construction [NP\textsubscript{research} V\textsubscript{show} that-clause/NP], the latter correspond to control items. This difference could not be explained through the order of stimuli. For instance, example (147) was shown before (148), but the dictionary was consulted only in the second case. This indicates that the verb \textit{suggest}, similar to \textit{show}, has different senses depending on its co-occurrence with the subjects referring to human agents or to research process and result (cf. Doherty 1996, discussed in Chapter 5.1.1 above). The first participant translated the verb \textit{suggest} both in (147) and (149) through the verb \textit{vorschlagen} (‘suggest’), but decided against using the same target verb to translate instances (148) and (150). This suggests that the verb \textit{vorschlagen} is more likely to be integrated in the transitive pattern [Subj\textsubscript{human} V\textsubscript{volition} Dobj]. This assumption should be verified in a larger data sample.

(147) \textbf{EO:} Some researchers have suggested that crabs in undersized shells were more likely to be eaten by predators. (Text 6)
\begin{itemize}
  \item \textbf{GTrans\textsubscript{1}:} Einige Forscher haben vorgeschlagen, dass Krabben in zu kleinen Schalen größere Gefahr laufen von Raubtieren gefressen zu werden.
  \item \textbf{GTrans\textsubscript{2}:} Einige Forscher meinen, dass Krebse mit zu kleinen Panzern eher von Raubfischen gefressen werden.
\end{itemize}

(148) \textbf{EO:} Previous studies suggest that you must raise your heart rate for at least 20 minutes per session. (Text 10)
\begin{itemize}
  \item \textbf{GTrans\textsubscript{1}:} Vorherige Studien empfehlen, dass man seinen Herzschlag für mindestens 20 Minuten pro Sitzung erhöhen muss.
  \item \textbf{GTrans\textsubscript{2}:} Bisherige Studien sagen, dass man seine Herzfrequenz dabei mindestens 20 Minuten pro Trainingseinheit hochhalten muss.
\end{itemize}

(149) \textbf{EO:} Several researchers suggest that cohesive groups may display higher levels of group productivity. (Text 11)
\begin{itemize}
  \item \textbf{GTrans\textsubscript{1}:} Einige Wissenschaftler schlagen vor, dass der Gruppenzusammenhalt einen höheren [sic] Grad an Gruppenproduktivität zeigen kann.
  \item \textbf{GTrans\textsubscript{2}:} Einige Forscher behaupten, dass zusammenhaltende Gruppen ein größeres Level an Produktivität zeigen.
\end{itemize}

(150) \textbf{EO:} New evidence suggests [that computers are now being used more often for a variety of planning purposes.]
\textbf{Clause} (Text 13)
GTrans_1: Neue Beweise erwecken [die Vorstellung, dass Computer mittlerweile öfter für eine Reihe von Planungabsichten [sic] benutzt werden.] NP

GTrans_2: Neue Ergebnisse zeigen, dass Computer heutzutage öfter für eine Vielzahl von Planungszwecken gebraucht werden.

To test the second assumption, according to which the instances of the construction [NP_{research V_{show that-clause/NP}] are expected to shift more often to another structure than kept unchanged in the process of translation, the final translations have been compared to the linguistic structures in the originals. As could be seen in examples (147) – (150), both types of stimuli are translated without changes of the transitive pattern. Also the remaining 16 instances of stimuli are characterized by rather literal translations with only minor changes. For instance, in (150) the translation by the first participant contains the direct object realized by a NP with the head noun *Vorstellung* ‘conception’, postmodified by a finite relative clause corresponding to the clause in the original. Moreover, in one example, shown in (151), the pattern [NP_{researchers V_{show that-clause/NP}] is translated through another transitive pattern, namely [Sub]_{human V_{non-volition Dobj]}.

(151) EO: But social scientists reveal that increased choice has been accompanied by decreased well-being.

GTrans_2: Sozialwissenschaftler haben allerdings rausgefunden, dass eine größere Auswahl mit einem schlechteren Selbstbefinden einhergeht.

Contrary to the assumption based on the contrastive difference in the use of the analyzed partially lexically filled constructions and construction shifts observed in the CroCo corpus, the pattern is mainly left unchanged in the pilot translation experiment. The reason for that might be the group of the experiment participants. I assume that the students of English linguistics with little or no experience in translation tend to retain the original linguistic structures. Professional translators with a high level of expertise are expected to be more aware of the contrastive difference between English and German, and, more importantly, have probably developed certain entrenched translation strategies that let them deal with this case of constructional resistance investing less cognitive effort than unexperienced translators. Taking the group of professional translators as subjects could thus also have an implication for the first hypothesis: the more often the translators encounter the construction [NP_{research V_{show that-clause/NP}] in English originals and have to find an adequate translation, the less cognitive effort could be required. Therefore, a future study
should repeat the experiment not only with a larger group of participants, but also with professional translators characterized by different years of experience.

The third research question concerns the correspondence between the translation units that could be identified within the translation process and the argument structure constructions, in particular the partially lexically filled construction discussed in the present chapter. For this purpose, the total of 24 Excel tables with the recorded keys and mouse movements corresponding to the translation of twelve texts by two participants have been further processed to determine the individual pauses between any of the two events. This calculation subtracted the time of one keystroke logging event from the time of the following event. Similar to the previous analysis, the pauses between the initial mouse movements and the first keystroke, as well as the last keystroke and the stop of recording are not taken into account.

As argued in Chapter 2, it is better to base the pause threshold on the typing speed of the individual translators rather than assume a general cutoff point of e.g. one second for the whole translation experiment. Since the average time per keystroke obtained from the copy test is the average typing speed of the participants and corresponds to an average pause between any of the two keystrokes, I do not consider it suitable as a pause threshold for the division of the process texts into larger production chunks, interpreted as translation units. Instead, the present study takes into account all pauses between individual events and compares the segments between the longest 10% and 5% of pauses per log file. The length of pauses is, therefore, adjusted not only to account for the different participants but also for the various production rates per stimulus, thus making the segments more comparable across all texts irrespective of e.g. lexical variation. In the next step the resulting TUs within the stimuli are analyzed in more detail. During the calculation of the number of TUs per stimulus, the segments involving only punctuation or space characters are not considered as separate TUs. In the first step, the 10% and 5% thresholds were compared to determine the number of cases with TUs, which consist only of fragments of words, as e.g. the TU *ien* in (152). Here such TUs as *haben gezeigt, dass* pat 'have shown that pat', which occurs both in (152) and (153), are not considered incomplete, since they contain several complete words. Therefore, the stimulus in (153) has not been counted among the stimuli with TUs consisting of word fragments. The analysis shows that while the 10% threshold results in 58% (14/24)
of stimuli with word fragments, this number goes down to 13% (3/24), if the threshold of 5% is adopted. Even though this pause length also means that some stimuli involve chunks going beyond the clause and sentence boundaries, the 5% threshold is adopted for the further analysis, as it involves more meaningful TUs.

(152) 10% threshold: [pause 4.664 sec.] Neue Stud [pause 0.593 sec.] ien [pause 0.687 sec.] haben gezeigt, dass pat [pause 1.825 sec.] (2nd participant, Text 4)
(153) 5% threshold: [pause 4.664 sec.] Neue Studien haben gezeigt, dass pat [pause 1.825 sec.] (2nd participant, Text 4)

Using this threshold, ten stimuli (42%, 10/24) consisting of only one TU could be identified. This TU corresponds either exactly to an instance of the analyzed construction or involves additional elements. For instance, example (153) contains one translation unit containing the analyzed part of the construction and a fragment of the next word. Within these ten stimuli, there is an equal distribution of critical and control items. The majority of these clause-level TUs have been produced by the second participant, while the text production of the first participant is characterized by smaller TUs. Therefore, also this analysis indicates variation among participants and should be repeated with a larger population sample. The results of this pilot experiment indicate that with respect to the size of TUs there is no difference between the instances of the analyzed partially lexically filled construction and the control items. The size of TUs does not seem to depend on the type of pattern realized through a clause. This assumption should be further tested in the future experiment.

6.4 Conclusion

The pilot experiment has indicated that the translation of the construction \([\text{NP research V show} that\text{-clause/NP}]\) could be cognitively more effortful than the control structure \([\text{NP researchers V show} that\text{-clause/NP}]\). This conclusion is supported by the eye-tracking measures, in particular by the findings related to the fixation count, which show a significant difference between the two conditions. The keystroke logging data has been shown to be more dependent on individual participants. Apart from small changes, the English construction has been translated by a similar German structure, possibly because of the limited translation experience of the experiment participants. Finally, no difference has been observed between the critical and control items in terms of translation units. The two types of stimuli have equally often formed one processing unit during the translation process.
Thus, from the corpus-based perspective our study has shown that the partially lexically filled construction under analysis is more frequent in English than in German originals and, therefore, tends to be changed in the translations from English to German. Based on these results, it is suggested that the construction is more entrenched in English than in German. A subsequent pilot experiment has partially provided additional evidence for this hypothesis indicating that the translation of this construction might require additional cognitive processing. In the next step, an experiment involving professional translators should be carried out to further test these hypotheses. Moreover, the keystroke logging and eye-tracking data can be aligned to examine the fixation patterns corresponding to individual pauses during translation. This fine-grained analysis might give us further insights into the cognitive processing during the translation process.
7. General discussion and evaluation

The two preceding chapters reported on the investigation of a number of selected constructions. The related research questions included the characteristic features of the analyzed abstract and partially lexically filled patterns, their role in translations in terms of translation shifts and translation properties as well as other factors that influence the translation of argument structure constructions. The corpus-based findings were also complemented by a brief outlook on the analysis of cognitive processing during the translation of one specific construction. The present chapter abstracts from the level of individual constructions to evaluate the application of the CxG framework to the study of translation shifts for the language pair English-German. For this purpose, we return to the initial motivations on using the CxG approach outlined in Chapter 3.2, namely the use of constructions as cross-linguistic units of analyses, operationalization of the translation phenomenon of naturalness by linking it to the usage-based analyses performed within CxG (among other theories) as well as the link between the cognitive features of CxG and the cognitive processing during translation. The chapter discusses the results and the methodological challenges associated with each of these aspects.

First of all, in Chapter 3.2 it is indicated that constructions represent suitable units of cross-linguistic analysis, which range in the degree of complexity and abstractness. In the course of the study, a number of corresponding constructions in English and German were considered, namely three abstract argument structure constructions and several more specific constructions. The latter have been argued to be instances of the three general constructions under analysis and at the same time to be stored separately as constructions in their own right. Therefore, as shown in Chapter 5.5, the set of discussed patterns is organized in a network. It is argued that the study of various interconnections between the constructions on different levels of abstractness helps us interpret the use of the abstract patterns in originals and their potential shifts in translations. For instance, the study has discussed four partially lexically filled constructions on the clause level, which correspond to the abstract [Subj V Doj] argument structure. This set of constructions accounts for certain amount of variation within the use of the abstract pattern, even though the study is limited to the most frequent patterns within individual categories and thus does not aim at providing a full
inventory of all possible constructions present in the data samples. The identification of these lower-level constructions improves our understanding of the more specific functions realized through the abstract [Subj V Dobj] construction and the lexical restrictions on the particular slots, e.g. on the grammatical function of subject. These, in turn, provide us with different perspectives on the analysis of translations, which could consider the contrastive differences on different levels of abstractness, e.g. both between the abstract [Subj V Compl] and [Subj V Probj] constructions and the language-specific preferences for certain lexically filled patterns, such as [V be part of] and [V gehören zu]. Therefore, taking into account the different levels of abstractness we gain new insights into the potential reasons for construction shifts.

It is thus important to keep in mind that the same sentence is simultaneously an instance of several constructions of different degrees of abstractness, which can be more typical of one of the languages or associated with a particular register. All of these constructions with their sets of characteristic features could potentially contribute to a translation shift or a translation through a corresponding construction.

Moreover, every sentence also consists of several constructions of different sizes (e.g. VPs, non-finite clauses and specific lexical items), which also are more or less likely to be changed in the translation process. As discussed in Chapter 5.3, the identification of the construction that triggers a translation shift is not straightforward. Changes could be caused by the wish to shift from a clause-level pattern occurring in the source sentence (e.g. a combination of the pronoun we and the non-volitional verb need followed by a direct object) as well as to shift to an abstract argument structure construction or a specific category typical of the target language (e.g. to the [Subj V Probj] in general or, more specifically, to the pronominal adverb).

The flexible nature of the unit’s size is considered to be one of the advantages of the selected approach as the researcher can analyze the same structure from multiple perspectives taking into account not only different levels of abstractness but also different components of the structure and, most importantly, their combinations. For instance, the study by Doherty (1996) shows that the translation of the verb show from English to German depends on its co-occurrences with different types of subjects, because different combinations of subjects and verbs are more or less typical of English and German. Similarly, in our study it is the complex construction [NP,research V show clause/NP] that is more frequent in English than in
German originals, rather than the verbs of showing in isolation. While the cross-linguistic use of the corresponding constructions depends on the general frequencies of non-agentive subjects within the two languages, it could be indicated that the limited set of non-agentive subjects referring to research in combination with the verbs of showing have a specific function characteristic of the registers of academic and popular-scientific discourse. CxG provides a coherent framework for analysis of interconnections between units of various sizes allowing us to account for the role of non-agentive subjects but also their more specific functions and frequencies in the context of the larger clause-level constructions within certain registers.

The cross-linguistic comparability of constructions as units of analysis is ensured by the focus on corresponding constructions with similarities both on the formal and functional levels. However, the operationalization of particular constructions has posed a number of challenges. As discussed in Chapter 4, some of these challenges have been encountered already on the abstract level of grammatical functions because certain categories are present only in one of the languages, such as phrasal verbs in English and Zustandspassiv in German. For instance, sentences are classified either as [Subj V Dobj] or [Subj V Compl] depending on whether a verb is listed as a phrasal verb or not. The subsequent classification into the groups with and without construction shifts has been considered problematic when the crossing lines were the result of different initial categorizations rather than genuine differences in mapping of semantic information onto grammatical functions in originals and their translations that led to a shift in perspective or changes in the information structure (see Chapter 4.2.2).

Moreover, in accordance with the methodological decisions of this study, all instances of the abstract constructions have been first classified into several groups, motivated through formal or functional criteria. The more specific constructions have been identified in the next step based on the frequent patterns within the defined groups, which displayed contrastive differences and/or were frequently affected by translation shifts. The first step was, therefore, necessary to sort the data, so that potentially interesting patterns could be identified. However, the borders of the individual categories, such as [Subj inanimate V volition Dobj], are potentially too rigid for a further investigation of the selected constructions. For instance, examples such as the one shown in (154), are not considered among the more
specific construction \([\text{NP}_{\text{research}} \text{ V}_{\text{show}} \text{ clause/NP}]\), because the verb *find* is classified as a non-volitional verb, whereas the analyzed construction has been identified within the transitive pattern \([\text{Subj}_{\text{inanimate}} \text{ V}_{\text{volition}} \text{ Dobj}]\).

(154) **EO:** [A recent study by BRAC, in collaboration with the World Bank, found that villagers are even willing to pay part of the cost for installation.] (EO_POPSCI_002 s47)

**GTrans:** [Wie eine von uns gemeinsam mit der Weltbank durchgeführte Studie zeigt, wäre die Landbevölkerung sogar bereit, einen Teil der Installationskosten selbst zu tragen.

We could observe that the translation of this sentence represents a typical translation strategy associated with the construction \([\text{NP}_{\text{research}} \text{ V}_{\text{show}} \text{ clause/NP}]\). The subject and the finite verb of the original correspond to an adverbial in the translation, whereas the subordinate clause is upgraded. While the translation strategy alone is not a reliable indicator of the classification of the original sentence, it could be argued that the sentence in (154) has the same function as the examples discussed in Chapter 5.1.1 making it possible for this sentence to be an instance of the same construction. However, a reliable investigation of particular constructions does require clear criteria for the operationalization of its slots, so that the decision on the inclusion of sentences into the study of a particular construction is based not on the intuition of individual researchers. Therefore, the description of a partially lexically fixed construction requires definitions of the semantic groups that fill the individual slots. Here the challenge is to find a balance between categories that are too restrictive and too fuzzy. In the former case the analysis covers only a part of the instances that could be associated with the construction, whereas in the latter case the study cannot be replicated by a different group of researchers using a different data sample. For instance, due to a lack of a clear definition, it was not possible to study in more detail the category of cleft-like structures identified by Doherty (2001).

However, though the definition of very precise criteria reduces the amount of subjective decisions of the individual researcher and allows for the replication of the study under different conditions, the issue of cross-linguistic comparability still remains. In this study I used the categorization of verbs available in *WordNet* and *GermaNet*. However, it is difficult to assess to what extent the English and German semantic groups created using these tools are comparable. Staying with the example of the construction \([\text{NP}_{\text{research}} \text{ V}_{\text{show}} \text{ clause/NP}]\), we
could note that a very low frequency of the construction within the German data could depend on the overall number of the German verbs that are linked to certain semantic groups in GermaNet, such as the groups of showing and saying. If the number of verbs annotated as instances of these categories could be shown to differ between the two tools, it is still impossible to say with absolute certainty whether this is a difference in the annotation or a contrastive difference between English and German. Example (155) could be used to illustrate the described difference in the classification. On the level of the abstract argument structure patterns, the translation pair in (155) has been classified within the category of no construction shifts because both sentences consist of the same three grammatical functions with no crossing lines. However, on the level of transitive patterns, the structures used in this translation pair belong to the different groups. Based on WordNet the verb suggest has a volitional sense of speaking, whereas, according to GermaNet, there is only a non-volitional sense of perception that could be identified for the verb belegen ‘inform’. Therefore, the sentence in the German translation is not considered to be an example of the partially lexically filled [NP_{research} V_{show} clause/NP] construction (see Chapter 5.1.1), even though this pattern has certainly a similar function. Taking into account examples (154) and (155), future research needs to further refine the criteria for the operationalization of the constructions.

(155) **EO:** The research that my colleagues and I have done [suggests]_{V, volition} that maximizers are prime candidates for depression. (EO_POPSCI_005 s127)

**GTrans:** Unsere Forschung [belegt,]_{V, non-volition} daß [sic] Maximierer besonders anfällig für Depression sind.

Following Szymańska (2011), it was also suggested that the concept of naturalness in translation studies could be operationalized through the frequency with which a particular construction occurs in the target language. To produce a natural translation, translators were assumed to change linguistic structures to constructions that are typical of the target language. Therefore, before investigating translation shifts, the frequency distributions of the corresponding constructions in the originals have been compared. This allowed the investigation of the role of constructional resistance or mismatch proposed in previous studies. For instance, the lower-level construction [NP_{research} V_{show} that-clause/NP] has been shown to be more typical of English than of German. This contrastive difference may lead to a certain amount of constructional resistance reflected in construction shifts in the
translation direction English-German. However, as mentioned in Chapter 5.5, it was not possible to establish a clear correlation between an occurrence of a construction in the original and a construction shift. Instead, translations show a combination of two translation properties, namely normalization and shining through. These results are in line with the previous study by Teich (2003), where these properties have been operationalized through different linguistic features. The analysis in the present study indicates that, in addition to linguistic features investigated by Teich, also clause-level constructions can be used as one way to operationalize the abstract translation properties, thus also going beyond shallow statistics as the measures of their analysis.

Previous studies have also suggested that in the case of some constructions we could observe an increase in the frequency with which a construction is used in the originals of one language, due to the general influence of translations and other types of language contact situations (König and Gast 2009, 108; Hansen-Schirra and Steiner 2012, 274, see Chapter 5.1.1). The construction once borrowed from another language (Serbina 2013, 187) gradually becomes more natural, at least within certain registers.

While the present study started with a relatively large number of instances corresponding to the three argument structure constructions under analysis, the size of the samples reduced significantly once the investigation concentrated on the more specific constructions. The lower values for the latter patterns could be expected, since the original data sets are classified into several groups zooming in onto several partially lexically-fixed constructions. Due to their specific functions, these constructions account only for a small proportion of the overall variation within the abstract patterns. Thus, while, as discussed above, the investigation of the same structures on different levels of abstractness is the strength of the study, a small set of examples on the lower level is certainly problematic. Even though the study conducts a number of statistical analyses in an attempt to generalize beyond the analyzed sample, these generalizations are only tentative due to a small number of instances.

In this study, I operationalize construction shifts using a parallel English-German corpus annotated for grammatical functions and aligned on both the word and phrasal levels. In this respect, the CroCo corpus has proven to be a useful resource, as it contains all the types of data enrichment, which are required for this type of research. However, the size of the corpus or at least of the two selected registers appears to be insufficient for the study of the more
specific patterns. While the extension of the investigation to the remaining registers included in the corpus might increase the overall number of examples on the level of abstract constructions, I do not expect to find a large amount of additional instances of the analyzed lower-level structures, since they are often specific to particular registers. In this respect, only a more detailed study of the register of prepared speeches seems promising as this group of texts is related to the political essays.

On the one hand, a small corpus size has the advantage that the research does not have to be limited to a random set of e.g. copular constructions, but can consider the whole range of variation within this category, annotate various formal and functional categories and analyze these both quantitatively and qualitatively. Chapter 4.1 also observes that a larger corpus size might increase the chances of type I error, falsely rejecting the null hypothesis. On the other hand, if the sample size corresponding to the specific constructions is very small, “[i]t will typically be difficult to reject H0 […], unless the true proportion is very far away from the null hypothesis” (Baroni and Evert 2009, 783).

Therefore, a solution might be to use a smaller corpus for an initial heuristic identification of potentially interesting patterns, i.e. constructions, which are characterized through different frequency distributions in the two languages under analysis. Once such lower-level constructions are identified, the next step requires their systematic extraction from a larger parallel bidirectional corpus containing the same registers and ideally annotated for grammatical functions. This would allow a further verification of the assumptions related to contrastive differences, shifts from and to specific constructions and the translation properties of normalization and shining through, operationalized through the occurrences of language-specific patterns.

The additional corpora queried within the present study met only some of these requirements. For instance, the Europarl corpus is a parallel corpus, which consists of parliamentary debates in a range of languages including English and German (Koehn 2005). Therefore, the corpus could be theoretically used as a larger basis for the study of English and German originals and translations belonging to the political discourse. However, the corpus search does not distinguish between translated and non-translated data belonging to the same language, even though the interface used for the queries suggests the opposite (see Chapter 5.1.2 for more details). This distinction is, however, necessary for the specific
research questions investigated in the present study. Due to their large size of 100 million words, the reference corpora, queried to investigate the intralingual use of the analyzed constructions, include a sufficient number of selected lower-level constructions (the analyses within our study have been limited to random samples of 200 examples). However, both BNC and DWDS are not annotated for grammatical functions, so that the extraction of the argument structure constructions has been based on individual lexical items (see Chapter 4.3). As a result, it was not possible to assess the full variation of lexical material that can occur in the slots of the construction. Moreover, due to the limitation of the performed queries to individual lexical items, the degree of association between certain words and constructions could not be established using the data from the reference corpora. The collostructional analyses have been carried out using the samples identified in the CroCo corpus. Furthermore, these reference corpora cover only one of the registers under analysis, namely popular-scientific prose. This is problematic for the study of lower-level constructions associated with the register of political discourse.

Finally, Chapter 3.2 indicated that the application of a cognitive theory to the study of translations is compatible with the analysis of cognitive processing during the process of translation, investigated in particular in the process-based branch of translation research. Based on the quantitative data, the present study could, therefore, formulate a hypothesis on different degrees of entrenchment related to the corresponding constructions with the form \([\text{NP}\text{research } V_{\text{show clause}}/\text{NP}]\) by speakers of English and German. In Chapter 5.1.1, it is suggested that the construction is more entrenched in English, whereas German speakers are likely to use other patterns to express the same function, namely to refer to results of the studies. Within the theoretical framework of CxG, the entrenchment assumption links information on the frequency of occurrences, observed in the corpora, to the abstract level of cognition. This link allows us to use the results of the corpus-based study as the basis for a hypothesis on cognitive processing during the translation process, which is, in turn, also operationalized through concrete indicators, such as number and total duration of fixations as well as average production time of the construction \([\text{NP}\text{research } V_{\text{show clause}}/\text{NP}]\). In Chapter 6, it was assumed that the translation of the pattern with a higher degree of entrenchment in the source than in the target language involves additional cognitive processing, compared to the translation of another pattern with a similar function.
The eye-tracking data collected within the pilot translation experiment provides initial support to this hypothesis. We can thus observe that assumptions about the level of cognition serve as common ground between the theory of CxG and the area of translation studies, which can be investigated through data triangulation. In future studies, a combination of different data sources can also help us test assumptions with respect to independent mental representations of the individual constructions. This step should ensure that the investigated patterns are psychologically real categories rather than a product of the analysis (Tomasello 2003, 98).

Once the translation experiment reported in this study is repeated to obtain a larger data sample from translations performed by professionals, it is important to further consider different types of operationalizations associated with the level of cognition. In the present pilot experiment the eye-tracking data provided more support for the assumption on the amount of cognitive processing than the average production time calculated using the keystroke logging data. The future experiment should consider the variation both in terms of individual subjects and stimuli by performing mixed effect regression analyses. Moreover, as mentioned in Chapter 6, the temporal alignment of the pauses occurring during the text production and the place of fixations will provide further insights into the translation process. This alignment requires temporal synchronization of eye-tracking and keystroke logging files as well as development of querying techniques that would allow extracting data from both streams\textsuperscript{53}.

To conclude, this study has indicated a number of ways in which the analysis of translations in general and of translation shifts in particular can benefit from the Construction Grammar approach. It is important to stress that I do not aim to present the selected theory as the only or the best framework to the investigation of translations. Instead, it is argued that the concept of constructions in combination with usage-based and cognitive aspects of the theory presents a promising alternative providing new perspectives on the analysis of translation shifts. The study has also shown that the application of CxG could be combined with certain concepts from the SFL. A comparable role of linguistic functions and the usage-based nature of both theories allow for the theories to complement each other. For instance, one of the

\textsuperscript{53} See work within the interdisciplinary project \textit{e-cosmos} at RWTH Aachen University, \url{http://www.humtec.rwth-aachen.de/index.php?article_id=1026&clang=1}
challenges of the present research, namely operationalization of the abstract categories, has been addressed using the SFL notions. Such concepts as the logical subject and grammatical metaphor integrated into the CxG framework have helped us develop a set of criteria distinguishing between volitional and non-volitional uses of verbs. This integrated approach is not restricted to the area of translation studies and appears promising to an investigation of semantic features associated with constructions of various sizes. Moreover, the identification of methodological issues discussed in the present chapter could give new impulses for empirical linguistic studies. In particular, I have stressed the importance but also the difficulty of precise operationalization of the discussed categories. The suggested definitions for a number of criteria associated with such categories as construction shifts and volition leading to the delimitation of the lower-levels constructions provide the basis for further methodological work within the empirical usage-based studies in linguistics.
8. Conclusion

The aim of the present study has been to investigate the phenomenon of translation shifts from the perspective of Construction Grammar. For this purpose, based on the frequency of construction shifts on the level of the main clause and relevant contrastive differences, the study has concentrated on the three argument structure constructions, namely [Subj V Dobj], [Subj V Compl] and [Subj V Probj]. The constructions [Subj V Dobj] and [Subj V Compl] have been selected because they are frequent in the English and German originals and at the same time are often changed to other structures in the process of translation. Due to the contrastive difference in the use of complements and prepositional objects in the two languages under analysis (Čulo et al. 2012, 114), I have also included the [Subj V Probj] construction in the set of analyzed structures, despite its lower frequencies.

In order to make sure the analyzed units are comparable across languages, the study identifies corresponding constructions in English and German, i.e. constructions sharing certain features both on the formal and functional levels (Leino 2010, 131). Thus, even though constructions are considered to be language-specific, the analyzed structures still have a number of formal and functional similarities in English and German. It has been shown that the constructions are associated with different frequencies of use in the two languages. Our data sample indicates that the structures [Subj V Dobj] and [Subj V Compl] occur more often in the English originals, whereas [Subj V Probj] is more frequent in the German originals. Moreover, further classification of the sentences with the general structure [Subj V Dobj] and [Subj V Compl] has suggested that there are further functional and formal differences related to the English and German corresponding constructions.

The [Subj V Dobj] construction has been classified according to the semantic criteria of animacy and volition. This classification was motivated by the reported contrastive difference between the mapping of semantic roles of Time, Location and Instrument onto the grammatical function of subject (Hawkins 1986, 57-61; König and Gast 2009, 108-109) and our assumption on the different acceptability of the combinations of these subjects with the volitional verbs (cf. also Doherty 1996; 2002, 76-78) in the two languages. The cross-linguistic comparison of the main data sample has shown that the patterns [Subj\text{inanimate} V_{\text{volition}} Dobj] and [Subj\text{animate} V_{\text{non-volition}} Dobj] are positively associated with the English
originals, whereas the pattern \([\text{Subject, human} \ V\ volition \ Dobj]\) is attracted to the German originals. Moreover, the structure \([\text{Subject, non-ref} \ V \ Dobj]\) has been identified only in the latter group of texts. These patterns have thus been analyzed in more detail.

It has been shown that it is the structure \([\text{NP, research} \ V\ show \ clause/\text{NP}]\) that accounts for over a half of shifts from the transitive pattern \([\text{Subj, animate} \ V\ volition \ Dobj]\). This structure is considered to be a construction in its own right. It is used to report the results of the studies particularly in the registers of popular-scientific and academic writings. The two slots of the construction can be filled by a limited set of lexical items, namely nouns of research and verbs of showing. To be more precise, the data extracted from the CroCo corpus indicates that the lexemes \textit{study} and \textit{show} are likely to co-occur in this construction. According to the comparison of the instances \([\text{NP, study} \ V\ show \ clause/\text{NP}]\) and \([\text{NP, Ergebnis} \ V\ zeigen \ clause/\text{NP}]\) in the BNC and DWDS, the construction occurs more frequently in English than in German. Therefore, it could be assumed that the construction is more entrenched in the minds of the English speakers. We have also observed that this lower-level construction is often shifted to another structure in the German translations, but at the same time there are a number of translations using the corresponding German construction, even though it is less typical of this language. Therefore, with respect to this pattern, the German translations are characterized by a combination of two translation properties, namely normalization and shining through. The property of normalization is also present in the English translations, as this construction, which is more characteristic of English, is sometimes introduced in the translations without a corresponding construction functioning as a trigger. The information structure of the original is typically preserved. In the cases of normalization the level of impersonality is likely to increase in the German translations.

Within the transitive pattern \([\text{Subj, animate} \ V\ non-volition \ Dobj]\) we have concentrated on the construction with the form \([\text{NP, country/government} \ V\ feel/think \ clause/\text{NP}]\) used to realize the functions of political discourse of legitimization and coercion, more specifically communication of stance and implication of shared opinions. This pattern has been attributed to a large share of sentences with the general structure \([\text{Subj, animate} \ V\ non-volition \ Dobj]\) in the English originals, in contrast to only one example in the German originals from the register of political essays. However, the register of prepared speeches, which could also be classified as belonging to political discourse, shows the opposite cross-linguistic frequency
distributions. A closer investigation of the meta-information indicates that the construction is more likely to occur if the texts are written by politicians rather than journalists, as the latter try to stay more objective. Therefore, the difference in the use of the construction, observed in the political essays, either depends on the particular selection of texts included in the CroCo corpus, or reflects a general difference in the register characteristics in English and German. The original English construction is often kept unchanged in the corresponding translations. The cases of translation shifts have been explained through the reduction of clauses and local lexical choices.

The third transitive pattern considered in our study, namely [Subjhuman Vvolition Dobj], is characterized by the more specific structure [NPhuman Vspeak Clause/NP], i.e. a quotative construction. Also this pattern is considered a construction in its own right. Characterized by lexical restrictions particularly on the verb slot, it has a distinct function, as it is used to refer to the message expressed by another speaker and thereby often to support the argument of the current speaker. In our data, it occurs more frequently in the register of political essays, where it could also be linked to the functions of legitimization and coercion. The construction has been further classified into different quotation types. While in the main data sample the construction is more frequent in German than English originals, the pattern occurs more frequently in the English originals belonging to the register of prepared speeches and registers included into the corpora BNC and DWDS. Therefore, it is assumed that the use of quotations depends on the register, rather than the language in general. It has been also observed that due to language-specific structural restrictions the information structure is often changed in the translation direction German-English: the translation of German quotation through the corresponding English construction often requires changes in word order, for instance by reversing the order of subject and verb.

The German structure [Es gibt NP], which was identified within the transitive pattern [Subjnon-ref V Dobj], is assumed to correspond to the English structure [There Vbe NP]. Both English and German patterns are existential constructions, which introduce new information or emphasize the semantic content expressed by the final NP. The constructions can express the same functions and have a formal similarity, as both contain the marker of existence, i.e. there or es. The random samples from BNC and DWDS indicate a slightly larger number of instances of this construction in the English originals. However, it has also been shown that
the construction often contains additional grammatical functions, which have not been considered in the present study.

The analysis of shifts from the [Subj V Dobj] construction has been concluded by building a logistic regression model with mixed effects, which takes into account a number of more general predictors, namely the source language, register and the number of clauses in the original sentence (on the main level of analysis). These function as fixed effects, whereas the variable of text is included as a random effect. The model has indicated that the number of construction shifts varies depending on the source language and, to some extent, also register.

In the next step, the [Subj V Compl] and the [Subj V Probl] constructions have been compared in terms of their frequencies and the number of shifts from and into these constructions. As shown in previous studies, the former pattern is more typical of English, while the latter structure occurs more frequently in German. The [Subj V Compl] construction has been classified further based on its formal properties. This analysis showed that while the German [Subj V Compl] construction is mainly realized by the pattern [NP V NP], the English corresponding construction has three frequent patterns, namely [NP V NP], [NP V AP] and [NP V Clause]. From these, only the latter has a significant effect on the category of construction shifts in the translation direction English-German.

Therefore, this pattern has been selected for a more detailed analysis. It was indicated that more than a half of the sentences with the structure [NP V Clause] in the English originals could be characterized as the demonstrative wh-cleft construction typical of English. According to the analyses based on the data from the BNC, the formal properties of the construction, namely the specific combinations of lexical items filling individual slots and the degree of fixedness, appear to vary across registers. In the future studies, a closer look at the random samples of this construction belonging to different registers might also reveal functional differences.

The translations of the demonstrative wh-clefts for the direction English-German are characterized by clause reduction. The construction also occurs in the English translations, where it displays more variation with respect to its individual slots. This can be at least partially explained taking into account the aligned data. Specific linguistic features present in the German originals, which might trigger the introduction of the demonstrative wh-clefts in
the corresponding English translations, have an effect on the construction’s variant selected by the translator.

The present study has also shown that copular constructions containing clauses corresponding either to the grammatical function of subject or subject complement are characterized by the three general translation strategies suggested by Doherty (2001) for a more specific category of cleft-like sentences. The number of clauses can be reduced in translation by using the verb from the original subject or complement as the main verb of the target sentence. In addition, a similar pattern containing a finite or a non-finite clause can be used in the translation. The analysis indicated that clause reduction is more typical of the translation direction English-German. In general, English contains more sentences, which are characterized by clauses filling one of the slots of the abstract construction [Subj V Compl]. Moreover, clause reduction may depend on the type of clause present in the original. As shown in Fischer (2013), the ing-clause, which exists only in English, corresponds to either noun phrases or infinitive clauses in the German sentences.

The multivariate analysis of translation shifts from the [Subj V Compl] construction considered the same variables as the investigation of shifts from [Subj V Dobj]. It identified the source language and the number of clauses in the originals as the two significant predictors. Since some types of clefts, which are included in the data, namely the wh-clefts and the demonstrative wh-clefts, have the formal structure [Clause V Clause] or [NP V Clause], the presence of clefts in the English originals can help explain these results. Moreover, as mentioned above, it is less typical of German to fill the main clause constituents by finite or non-finite clauses.

Moreover, the study has also discussed the construction [Subj V Probj]. However, due to the limited number of instances corresponding to this pattern, no further classification of the data was possible. It is suggested that the construction occurs more frequently in German due to the category of phrasal verbs, which is distinguished only for English and the group of pronominal adverbs typical of German. The chapter also discussed how this construction could be analyzed on the lower levels of abstraction given a larger data sample. It is assumed that certain translation shifts could be explained taking into account specific collocational patterns and idiomatic expressions, such as \[V_{be \text{ right/wrong}}\] and \[V_{haben \text{ Recht/Unrecht}}\].
On a more general note, the thesis has shown that our understanding of translation shifts from the abstract argument structure constructions could be deepened considering various interconnections between more specific constructions of varying degrees of abstractness and complexity. Each translation pair contains a number of both abstract and more specific constructions of different sizes in the original and the corresponding translation. All of these structures are characterized by a set of formal and functional features, which include e.g. the associated level of impersonality or the set of registers it typically occurs in. Potentially, each of these language-specific patterns could lead to translation shifts. Therefore, the multivariate statistical approach to the analysis of translation shifts appears promising and should be pursued further by extending the range of considered constructions. For instance, our study has mainly focused on clause-level structures but smaller idiomatic chunks and (register-specific) frequency distributions of individual words could also be taken into account. For this purpose, a larger data collection is required, which might lead not only to more data points and occurrences of additional constructions, but also to refinement of the analyzed categories.

In the present work, the corpus-based results on one more specific construction, namely [NP\text{research} V\text{show} clause/NP], have been triangulated with the data from a pilot translation experiment using the methods of eye-tracking and keystroke logging. Based on the higher frequency of the construction in the English originals, I have assumed its higher level of cognitive entrenchment in English than in German. This, in turn, led to the hypothesis related to an increased cognitive effort during the translation of the English construction into German. The assumption was tested by an investigation of concrete behavioral measures, namely fixation count, total fixation duration and total production time. In particular, the eye-tracking measure of fixation count has provided support for the hypothesis. The keystroke logging data indicated that the production time could be subject to more variation among the participants. Moreover, in contrast to our hypothesis, the English construction was translated with the corresponding German construction, with only minor changes. This finding was explained taking into account the group of participants: professional translators are expected to change the structures more frequently than the students who took part in the experiment. No differences between critical and control items could be identified with respect to the size of translation units formed during the translation of the analyzed structures. However, it is
important to stress that only tentative conclusions are possible as this pilot experiment was conducted with only two students of English linguistics to test the experiment design and show how the experimental methods can complement our corpus-based study of translation shifts. In the next step, the translation experiment should be repeated using a larger sample of participants with professional background in translation.

Future research can also look at construction shifts from a slightly different perspective by identifying a set of constructions that can be linked to the same register-specific function. Determining the conditions that lead to shifts from and to these constructions will help identify how the use of these alternative structures might affect the more global level of popular-scientific or political discourse both in original and translated texts.

The thesis has shown that the study of translation shifts gains greater explanatory power by applying the concepts of Construction Grammar. More specifically, the analysis of constructions of different sizes and levels of abstractness in both the originals and the translations has the potential of accounting for a range of factors that lead to changes in translations. Using a combination of corpus-based and experimental evidence as the basis for these analyses is a promising direction of research in this area.
List of References


Alves, Fabio, and Daniel Couto Vale. 2009. “Probing the unit of translation in time: Aspects of the design and development of a web application for storing, annotating, and querying translation process data.” Across Languages and Cultures 10 (2): 251–73.


Davies, Mark. 2004-. *BYU-BNC: (Based on the British National Corpus from Oxford University Press).* http://corpus.byu.edu/glowbe/.


Finthammer, Marc. GermaNet-Explorer. Adapted to the JAVA API 2.0 by Verena Henrich.


http://www.natcorp.ox.ac.uk/


### Appendix I: Covarying collexeme analysis for the English construction (C)  
\[ \text{NP}_{\text{research}} \text{V}_{\text{show clause/NP}} \]

<table>
<thead>
<tr>
<th>Collexeme_NP</th>
<th>Collexeme_V</th>
<th>NP freq. in C</th>
<th>V freq. in C</th>
<th>Obs. freq. of NP and V in C</th>
<th>Expect. freq. of NP and V in C</th>
<th>Relation</th>
<th>Coll. strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>study</td>
<td>show</td>
<td>6</td>
<td>11</td>
<td>5</td>
<td>2</td>
<td>attraction</td>
<td>2.02</td>
</tr>
<tr>
<td>dogma</td>
<td>state</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.03</td>
<td>attraction</td>
<td>1.52</td>
</tr>
<tr>
<td>projections</td>
<td>paint</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.03</td>
<td>attraction</td>
<td>1.52</td>
</tr>
<tr>
<td>theory</td>
<td>dictate</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.03</td>
<td>attraction</td>
<td>1.52</td>
</tr>
<tr>
<td>evidence</td>
<td>suggest</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>1.09</td>
<td>attraction</td>
<td>1.28</td>
</tr>
<tr>
<td>account</td>
<td>imply</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.06</td>
<td>attraction</td>
<td>1.22</td>
</tr>
<tr>
<td>rule</td>
<td>imply</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.06</td>
<td>attraction</td>
<td>1.22</td>
</tr>
<tr>
<td>measurement</td>
<td>indicate</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.09</td>
<td>attraction</td>
<td>1.04</td>
</tr>
<tr>
<td>work</td>
<td>show</td>
<td>2</td>
<td>11</td>
<td>2</td>
<td>0.67</td>
<td>attraction</td>
<td>0.98</td>
</tr>
<tr>
<td>treatment</td>
<td>reveal</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0.12</td>
<td>attraction</td>
<td>0.92</td>
</tr>
<tr>
<td>evidence</td>
<td>show</td>
<td>4</td>
<td>11</td>
<td>0</td>
<td>1.33</td>
<td>repulsion</td>
<td>0.75</td>
</tr>
<tr>
<td>assessment</td>
<td>suggest</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0.82</td>
<td>repulsion</td>
<td>0.43</td>
</tr>
<tr>
<td>study</td>
<td>reveal</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>0.73</td>
<td>repulsion</td>
<td>0.37</td>
</tr>
<tr>
<td>research</td>
<td>show</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>0.67</td>
<td>repulsion</td>
<td>0.36</td>
</tr>
<tr>
<td>result</td>
<td>show</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>0.67</td>
<td>repulsion</td>
<td>0.36</td>
</tr>
<tr>
<td>study</td>
<td>suggest</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td>1.64</td>
<td>repulsion</td>
<td>0.33</td>
</tr>
<tr>
<td>experiment</td>
<td>suggest</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>0.55</td>
<td>repulsion</td>
<td>0.28</td>
</tr>
<tr>
<td>result</td>
<td>suggest</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>0.55</td>
<td>repulsion</td>
<td>0.28</td>
</tr>
<tr>
<td>work</td>
<td>suggest</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>0.55</td>
<td>repulsion</td>
<td>0.28</td>
</tr>
<tr>
<td>study</td>
<td>indicate</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>0.55</td>
<td>repulsion</td>
<td>0.27</td>
</tr>
</tbody>
</table>
Appendix II: Stimuli used in the translation experiment in the order presented to the participants

One-parent families form a substantial, and growing, proportion of all families with dependent children. Many studies have shown that these families tend to rely mainly on state benefits for their incomes. But the general income support available to all families with children may not be relevant to the particular needs of single-parent families.

Chimpanzees are very intelligent animals. Some researchers have shown that chimps can create artworks and understand the concept of language. For instance, humans and chimps can use sign languages to communicate with each other to a reasonably advanced level.

Something very curious has happened to biology in recent years. Organisms that we see all about us have disappeared as the fundamental units of life. In their place we now have genes, which have taken over all the basic properties that used to characterize living organisms.

Now that sufficient quantities of interferons are available, we must establish whether human beings can tolerate higher doses. Recent studies have shown that patients will tolerate 20 times more interferon than assumed. But we do not know how patients would respond after taking interferon for a long time.

In many sunny parts of the world, solar cooking is a viable option. The recent development of solar cookers is an important achievement, as these devices save electricity. Portable solar cookers are easy to use.

It is important for crabs to occupy shells of the right size. Some researchers have suggested that crabs in undersized shells were more likely to be eaten by predators. Crabs can use a variety of factors to rapidly assess the size of a particular shell.

It should be admitted that some animals possess some power of imagination. Recent studies have shown that at least the higher primates probably have dreams when they sleep. As to what their, and our, dreams mean – that is a matter of guesswork.

Scientists want to create the world’s first quantum computer. Such a machine would manipulate individual atoms so that many calculations could be performed at once. However, some people argue that quantum computers could never work.

The project examines the effects of a geographical move on the individual, concentrating on students who leave home to take a place at university. Preliminary studies have shown that a high proportion of first year students experience homesickness. This may have an effect on their daily efficiency and concentration.

Regular, enjoyable exercise is the key to fitness improvement. Previous studies suggest that you must raise your heart rate for at least 20 minutes per session. Initially some individuals will find it hard to keep working out because their joints and muscles are not accustomed to exercise.
Group cohesion can be described as the sum of all the factors that influence members to remain in the group. Several researchers suggest that cohesive groups may display higher levels of group productivity. Moreover, members experience less work-related stress and are better adjusted to life in the organization.

It is only logical to think that more choice is always better. But social scientists reveal that increased choice has been accompanied by decreased well-being. Recently conducted research offers insight into why many people end up unhappy rather than pleased when their options expand.

As computing costs fall and equipment becomes easier to use, the application and influence of computers is likely to increase. New evidence suggests that computers are now being used more often for a variety of planning purposes. Further research is required to investigate computer applications in UK local authority planning departments.

Titan is the biggest of Saturn’s moons. Apart from the discovery that Titan has an atmosphere, little else is known about it. But many scientists hope that its further exploration could reveal how life arose on Earth.

We now have detailed information on the households in England in early modern times. Demographic historians have shown that only 5 per cent of the households consisted of three generations. But these figures do not tell us whether the grandparent lived just a few houses away.

Reality Monitoring is the ability to distinguish between memories of events that actually occurred and memories of events that were only planned. Several scientists have shown that people over seventy may confuse these two kinds of memory. Further experiments will be carried out to determine the underlying reasons for this kind of confusion.
Appendix III: Questions asked during the retrospective interview and the corresponding stimuli with highlighted regions

1. Was wird in dem hervorgehobenen Satz beschrieben?
   ‘What is described in the highlighted sentence?’

One-parent families form a substantial, and growing, proportion of all families with dependent children. Many studies have shown that these families tend to rely mainly on state benefits for their incomes. But the general income support available to all families with children may not be relevant to the particular needs of single-parent families.

2. Warum nehmen die Wissenschaftler an, dass die Schimpansen intelligent sind?
   ‘Why do the scientists assume that chimpanzees are intelligent animals?’

Chimpanzees are very intelligent animals. Some researchers have shown that chimps can create artworks and understand the concept of language. For instance, humans and chimps can use sign languages to communicate with each other to a reasonably advanced level.

3. Woher wissen wir, dass die Patienten mehr Interferon vertragen?
   ‘Where do we know from, that patients can tolerate more interferon?’

Now that sufficient quantities of interferons are available, we must establish whether human beings can tolerate higher doses. Recent studies have shown that patients will tolerate 20 times more interferon than assumed. But we do not know how patients would respond after taking interferon for a long time.

4. Warum ist die Entwicklung des Solarkochers so wichtig?
   ‘Why is the development of a solar cooker so important?’

In many sunny parts of the world, solar cooking is a viable option. The recent development of solar cookers is an important achievement, as these devices save electricity. Portable solar cookers are easy to use.

5. Was haben die Forscher gemacht?
   ‘What did the scientists do?’

It is important for crabs to occupy shells of the right size. Some researchers have suggested that crabs in undersized shells were more likely to be eaten by predators. Crabs can use a variety of factors to rapidly assess the size of a particular shell.

6. Was wird in dem hervorgehobenen Satz beschrieben?
   ‘What is described in the highlighted sentence?’

It should be admitted that some animals possess some power of imagination. Recent studies have shown that at least the higher primates probably have dreams when they sleep. As to what their, and our, dreams mean – that is a matter of guesswork.

7. Worauf bezieht sich der hervorgehobenen Ausdruck?
   ‘What does the highlighted expression refer to?’
Titan is the biggest of Saturn’s moons. Apart from the discovery that Titan has an atmosphere, little else is known about it. But many scientists hope that its further exploration could reveal how life arose on Earth.

8. Was hat Auswirkungen auf die Konzentration der Studierenden?
   ‘What has consequences for the concentration of students?’

The project examines the effects of a major geographical move on the individual, concentrating on students who leave home to take a place at university. Preliminary studies have shown that a high proportion of first year students experience homesickness. This may have an effect on their daily efficiency and concentration.

9. Was war das Ergebnis der vorherigen Studien?
   ‘What was the result of the previous studies?’

Regular, enjoyable exercise is the key to fitness improvement. Previous studies suggest that you must raise your heart rate for at least 20 minutes per session. Initially some individuals will find it hard to keep working out because their joints and muscles are not accustomed to exercise.

10. Was macht die Leute unglücklich?
   ‘What makes people unhappy?’

It is only logical to think that more choice is always better. But social scientists reveal that increased choice has been accompanied by decreased well-being. Recently conducted research offers insight into why many people end up unhappy rather than pleased when their options expand.

11. Was haben die Forscher gezeigt?
   ‘What have the researchers shown?’

Reality Monitoring is the ability to distinguish between memories of events that actually occurred and memories of events that were only planned. Several scientists have shown that people over seventy may confuse these two kinds of memory. Further experiments will be carried out to determine the underlying reasons for this kind of confusion.