The Nature of Translated Text

—

An Interdisciplinary Methodology for the Investigation of the Specific Properties of Translations

Dissertation zur Erlangung des
Grades eines Doktors der Philosophie der
Philosophischen Fakultäten der
Universität des Saarlandes

vorgelegt von Silvia Hansen

Saarbrücken, im Dezember 2002

Tag der letzten Promotionsleistung: 3. Dezember 2002
Dekan der Philosophischen Fakultät II: Prof. Dr. Erich Steiner
Berichterstatter: Prof. Dr. Erich Steiner
Berichterstatterin: PD Dr. Elke Teich

Preface

Wenn du eine Stunde glücklich sein willst, schlafe!
Wenn du einen Tag glücklich sein willst, geh fischen!
Wenn du ein Jahr glücklich sein willst, habe ein Vermögen!
Wenn du ein Leben glücklich sein willst, liebe deine Arbeit!

Chinesische Weisheit[1]

Als ich dieses Sprichwort zum ersten Mal las, habe ich mir Gedanken darüber gemacht, wie sehr mir das, was ich seit einiger Zeit als meine Arbeit bezeichne, wirklich gefällt. Zuerst schoss mir dabei der Gedanke durch den Kopf, dass ich meine Arbeit während der Promotion mehr gehasst als gemocht habe. Wahrscheinlich wissen alle, die etwas ähnliches durchgestanden haben, was ich damit meine. Genauso wahrscheinlich ist, dass allen Promovierenden das schlechte Gewissen bekannt ist, das sich zu jeder Tages- und Nachtzeit sowie unabhängig von Jahreszeiten einstellt, wenn man sich mit irgendetwas beschäftigt, das nicht mit der Dissertation zu tun hat. Dabei spielt natürlich der Zeitfaktor eine besondere Rolle: Aus eigener Erfahrung kann ich mittlerweile schlussfolgern, dass Wörter wie *Freizeit* oder *Urlaub* ihre Bedeutung verlieren, wenn man sich einbildet, man könnte während der Promotion, die nach Möglichkeit innerhalb von drei Jahren zu absolvieren ist, einer Ganztags-Beschäftigung nachgehen. Nichtsdestotrotz

[1] With those in mind who most supported me during my PhD project, this preface was written in German so that they, too, can understand it. I apologize for any inconvenience English speaking readers might have, but I promise that the preface is the only chapter written in German. — Dieses Vorwort ist in deutscher Sprache verfasst, damit auch die Menschen, bei denen ich während meines Promotionsvorhabens die größte Unterstützung fand, dazu in der Lage sind, es voll und ganz zu verstehen. Ich bitte die Englisch sprechenden Leser um Verständnis dafür, verspreche aber gleichzeitig, dass das Vorwort das einzige auf Deutsch verfasste Kapitel dieser Arbeit ist.

habe ich während dieser Zeit unglaublich viel gelernt – vielleicht sogar gerade wegen der vielen Probleme und Schwierigkeiten, die sich vor allem in der letzten Phase der Dissertation zu häufen schienen und durch die mir mehr als klar wurde, wo meine persönlichen Grenzen liegen. Auf diese Weise habe ich zumindest gelernt, wie man Strategien zur Problemlösung entwickelt und wie man gewisse Grenzen überschreiten kann. Letzten Endes habe ich es sogar irgendwie geschafft, die Dissertation zu Ende zu bringen. Als Fazit würde ich also sagen, dass während meiner Promotion etwas zustande gekommen ist, auf das ich stolz bin und das mir im nachhinein wahrscheinlich mehr Spaß gemacht hat, als ich es während der zweieinhalb Jahre des Analysierens und Schreibens zugegeben hätte – und dessen Beendigung mein Leben um sehr viel mehr glücklicher macht ;-)

Aus diesem Grund möchte ich dieses Vorwort dazu nutzen, ein paar Dankesworte auszusprechen. Mein besonderer Dank gilt:

- dem Deutschen Akademischen Austauschdienst (DAAD), der die Anfangsphase meines Promotionsvorhabens durch ein Stipendium zur wissenschaftlichen Aus- und Fortbildung im Ausland unterstützte (Dieses Stipendium nutzte ich zur Analyse des Translational English Corpus (TEC) und des British National Corpus (BNC) an der University of Manchester Institute of Science and Technology (UMIST));

- Erich Steiner, der mir zur Beantragung des Stipendiums und zum Manchester-Aufenthalt geraten hat, mich bei der Finanzierung von Konferenzteilnahmen unterstützte, mir zeitweise seine halbe Bibliothek ausgeliehen hatte und mir jederzeit mit Rat und Tat zur Seite stand;

- Elke Teich, die meine Arbeit oft mit fachlichen Ratschlägen und teilweise äußerst kontroversen Diskussionen (meistens am Fenster) vorwärts trieb;

- Hans Uszkoreit, der mir die Teilnahme an Konferenzen ermöglichte und die Beendigung der Dissertation unterstützte;

- Mona Baker und Maeve Olohan, die mir in Manchester die nötige organisatorische und fachliche Hilfestellung bieten konnten;

- Christoph Scheepers, der mir meine psycholinguistischen Fragen beantworten konnte;
- Alexandre Ndeffo, der als französischer Muttersprachler das Experiment betreute;
- Frank Klein, der mir bei der technischen Realisierung des Experiments half;
- den beiden anonymen Übersetzern, die sehr kooperativ waren und ohne die das Experiment nie hätte stattfinden können;
- Mary Klaumann, die als englische Muttersprachlerin einen Korrekturlese-Job abgeliefert hat, der besser nicht hätte aussehen können;
- Sabine Brants, die nicht nur für fachliche Fragen und nicht nur während des Mensa-Essens ein offenes Ohr für mich hatte;
- Stella Neumann, die in ständiger Bereitschaft bei diversen Kaffee-Pausen nicht nur Probleme mit mir diskutiert, sondern auch gelöst hat;
- meinen Eltern, die mir die Möglichkeit zur Promotion erst eröffnet haben und ohne die ich, trotz aller Linguistik, bis heute noch nicht wüsste, was ein "Phrasen-Schwein" ist;
- meinen Schwestern, die trotz noch so großer Bockigkeit immer an mich geglaubt haben;
- und last but not least Stephan Schirra, der mich nicht nur in der schwersten Phase der Promotion ertragen musste und sowohl mich als auch meine Häschen im wahrsten Sinne des Wortes gefüttert hat, sondern der mich auch immer aufheitern und auf andere Gedanken bringen konnte, wenn ich glaubte, es wüchse mir alles über den Kopf – und der mir beibrachte, dass ich arbeite, um zu leben und nicht lebe, um zu arbeiten.

Des weiteren gilt ein herzlicher Dank allen fleißigen Helfern, die mich und meinen Gips-Arm beim Einarbeiten der sprachlichen Korrekturen entlastet haben und ohne die die fristgerechte Abgabe dieser Dissertation gefährdet gewesen wäre.

Zu guter Letzt möchte ich noch einmal betonen, dass trotz all der Hilfe, die mir durch die oben genannten Menschen zugekommen ist, alle Fehler und Schwächen dieser Arbeit einzig und allein mir zuzuschreiben sind.

Abstract

The aim of the present thesis is to define a methodology for investigating the specific properties of translations, on the one hand, and investigating the nature of translated text, on the other. The analysis of translated text as a special kind of language variation is the issue in focus here. This kind of research starts from the assumption that translations differ both from their source language texts and from comparable texts in the target language. As a result, they exhibit specific properties which cannot be found in non-translated text and which cause a register shift, as compared to originally produced text in the source language as well as the target language.

In order to investigate translated text as a special kind of text type or register, an interdisciplinary methodology is proposed in this thesis, which exploits advantages of empirical, cross-linguistic and psycholinguistic methods. The combination of computational techniques provides a product- as well as process-oriented view on the nature of translated text.

The proposed methodology is applied to the empirical analysis of English translated narrative texts and English original narrative texts, to the cross-lingual analysis of English translated narrative texts and their German and French source language texts as well as to a psycholinguistic experiment which gives insight into the black box of translation, by looking at the cognitive processes taking place during the translation process. Not only does the combination of product- and process-oriented computational investigation methods allow the description of the nature of translated text, but also the explanation of possible sources of the specific properties of translations. Finally, the methodology is assessed and issues for future work and possible application contexts are discussed.

List of notational conventions

- CAPITALS are used for abbreviations of names of computational tools and techniques, standards, corpora, linguistic theories, linguistic institutes as well as linguistic categories used in examples,
- **bold face** highlights technical terms in headings and relevant structures in examples,
- *italics* label technical terms,
- "double quotes" mark quotations and examples,
- 'single quotes' indicate technical terms in quotations and examples.

List of abbreviations

- AG – Annotation Graphs
- BNC – British National Corpus
- CAT scan – Computerized Axial Tomography scan
- CDATA – Character Data
- CLAWS – Constituent Likelihood Automatic Word-tagging System
- CQP – Corpus Query Processor
- DCU – Dublin City University
- DSR scan – Dynamic Spatial Reconstructor scan
- DTD – Document Type Definition
- EEG – ElectroEncephaloGram
- EN – English
- ENPC – English-Norwegian Parallel Corpus
- EPI – Echo-Planar Imaging
- ERP – Event-Related Potential
- FR – French
- GE – German
- GEPCOLT – German-English Parallel Corpus Of Literary Texts
- GR – Grammatical Relations
- GUI – Graphical User Interface

- IMS – Institut für Maschinelle Sprachverarbeitung
- ISO – International Organization for Standardization
- KWIC – KeyWord In Context
- LOB corpus – Lancaster-Oslo/Bergen corpus
- MEG – MagnetaEncephaloGram
- MRI – Magnetic Resonance Imaging
- NP – Noun Phrase
- OMC – Oslo Multilingual Corpus
- ODAG – Ordered Directed Acyclic Graph
- PCDATA – Parsable Character Data
- PETT scan – Positron Emission Transaxial Tomography scan
- POS – Part-Of-Speech
- PP – Prepositional Phrase
- SALIS – School of Applied Language and Intercultural Studies
- SF – Systemic Functional
- SFL – Systemic Functional Linguistics
- SGML – Standard Generalized Mark-up Language
- SL – Source Language
- STTS – Stuttgart-Tübingen TagSet
- SVO word order – Subject-Verb-Object word order
- TAP – Think-Aloud Protocol
- TEC – Translational English Corpus
- TEI – Text Encoding Initiative
- TL – Target Language
- TnT – Trigrams 'n' Tags

- TSV format – Tab Separated Vector format
- TT – Translation Technique
- TU – Translation Unit
- UMIST – University of Manchester, Institute of Science and Technology
- VP – Verb Phrase
- XML – eXtensible Mark-up Language

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"If we should be worrying about anything to do with the future of English, it should be not that the various standards will drift apart but that they will grow indistinguishable. And what a sad, sad loss that would be."

[Bryson, 1991]

1.1 Goals

According to Bryson's citation, the survival of different varieties of English should be worried about since it would be a loss if they merged into one indistinguishable standard form. However, as long as people speak and researchers investigate and describe the different variations of English, their survival can be guaranteed. These *Englishes* occur in many forms of variation: regional, social and functional varieties. Another form of variation, which has recently become subject of investigation, is analyzed in this thesis: the nature of translated texts. This kind of variation is based on the assumption that translations differ from their source language (SL) texts and from comparable texts in the target language (TL), in the sense that they have specific properties which cannot be found in non-translated text. Obvious sources of these specific properties of translated text can be register- or typology-specific language use. This means that a translator is influenced by the given constraints on register use and typological differences between the languages involved. However, these constraints are relevant to every kind of text production and not limited to the translation process. Thus,

another possible source of the specific properties of translated text lies in the translation process itself: as translation-specific constraints, such as the SL text and the translation purpose, have a strong impact on the cognitive processes taking place in the translator's mind, the translation process represents a specific kind of text production, hence, it results in a specific kind of text type.

Since the nature of translated text has not yet been investigated in combination with the sources of the specific properties of translations, the main goals of this thesis are as follows:

- **The development of a methodology for the investigation of translation properties.** Since this methodology is intended both to describe and explain the specific properties of translated text, it has to meet the requirements of product- as well as process-oriented research. Thus, it has to combine techniques which allow the analysis of the relation between English translations and English comparable texts, the relation between English TL texts and their German and French SL texts as well as the sources of the differences between TL texts, SL texts and comparable texts in the TL.

- **The investigation of the nature of translated text.** This investigation looks into the following questions: What is the relationship between English translations and English comparable texts of the same text type or register? What is the relationship between English TL texts and their German and French SL texts? How can the differences and commonalities of the TL texts, SL texts and comparable texts in the TL be explained? These questions aim at describing and explaining the specific properties of translations.

The following sub-goals are of importance with a view to these main objectives:

1. In order to describe the nature of translated text, the differences between English translations and English originals are to be investigated in terms of quantitative patterns on the basis of a monolingually comparable corpus.

2. In order to describe and explain the nature of translated text, the differences between English translations and their German and French originals are to be investigated in terms of qualitative patterns on the basis of a multilingually parallel corpus.

3. In order to give insight into the translation process with a view to explaining the nature of translated text, the translation process is to be investigated on the basis of a psycholinguistic experiment.

The first sub-goal aims at a quantitative study and can be investigated on the basis of an empirical corpus-based analysis, while the second sub-goal involves a qualitative study and can be investigated with the help of a cross-linguistic corpus-based analysis. The third sub-goal, however, requires the development of a psycholinguistic experiment. Thus, the independent variables are different for each sub-goal (i.e. large comparable corpus vs. smaller parallel corpus vs. psycholinguistic experiment). This also applies to the dependent variables (i.e. the relation between English translations and English originals in view of their quantitative patterns vs. the relation between English translations and their German and French originals in view of their qualitative patterns vs. cognitive processes during the translation process). The first two sub-goals are product-oriented, whereas the last one is process-oriented. Since the three sub-goals are essential for the two main goals, the distinction made between empirical, cross-linguistic and psycholinguistic research can be seen as a leitmotif which recurs in each chapter of this thesis. Thus, the state of the art, the theoretical foundation, the methodology and the analysis reflect this threefold principle.

1.2 Motivation

Many translation scholars have described the specific properties of translations and the relation between the SL and TL text. For example, [Nida, 1964], [Newmark, 1988] and [Vinay & Darbelnet, 1995] provide detailed descriptions of *translation strategies*, *translation procedures* and the use of *equivalence*, occurring in translation. [Wilss, 1977] and [Snell-Hornby, 1988] introduced the notion of *retrospective vs. prospective translation*, whereas [Koller, 1984] coined the terms *adapting vs. transferring trans-*

lation. [Sager, 1994] speaks of *translation text types* and [Toury, 1995] formulated the *law of growing standardization vs. law of interference*. Furthermore, [Berman, 2000] discusses tendencies of translated text, such as *rationalization*, *clarification* and *expansion*, on the basis of which the foreignness of translations can be tested. The problem with these approaches to the nature of translated text is that they were developed rather intuitively and that little evidence is furnished to underpin them.
Recently, translation scholars have begun to conduct empirical research into the specific properties of translations with the aid of corpora: [Baker, 1996], [Laviosa-Braithwaite, 1996b] and [Olohan, 2001] count among those who formulated and tested the hypotheses regarding *explicitation*, *simplification*, *normalization* and *levelling out*. Within this context, other researchers also developed corpus-based hypotheses concerning the nature of translated text: e.g., Kenny's *sanitization* (cf. [Kenny, 1998b]), Teich's *shining-through* (cf. [Teich, 2001a]) and Steiner's *grammatical metaphor* (cf. [Steiner, 2001b]). House's *overt vs. covert translation* and the application of a *cultural filter* is another approach which is tested on corpora (cf. [House, 1977] and [House, 1997]). These corpus-based approaches yield evidence for testing the hypotheses, nonetheless they cannot explain all the differences and commonalities to be found in SL texts, TL texts as well as comparable texts in the SL and TL, since they are product-oriented rather than process-oriented. It is still unclear how the nature of translated text is influenced by register-specific language use, language typology or even the translation process itself.

For this reason, the goals formulated in Section 1.1 require a combination of product- and process-oriented research. This kind of combined research makes it possible to develop a methodology for investigating the nature of translated text as well as to describe and explain the specific properties of translations.
Thus, at a meta-level, this thesis integrates translation studies, corpus linguistics, contrastive linguistics, comparative linguistics, computational linguistics, psycholinguistics and linguistics in general to produce a methodology which allows the investigation of the nature of translated text and to analyze the translation properties from a product- as well as process-oriented viewpoint.

Finally, it should be mentioned that this study is devoted to the nature of English translated fiction. Therefore, hypotheses concerning other languages, registers or even translation universals are not taken into account. Thus, everything written in the following on the nature of translated text is limited to English translated fiction.

1.3 Methods

The methodology combining product- and process-oriented research has to meet the following requirements:

- The differences between English translations and English originals have to be described in terms of quantitative patterns.
- The differences between English translations and their German and French originals have to be described and explained in terms of qualitative patterns.
- The translation process has to be investigated in order to gain insight into the nature of translated text.

Since the first requirement deals with the analysis of quantitative patterns in English translations compared to English originals, a large English comparable corpus is necessary. The automatic comparison of the frequencies of certain features of the translations and the original texts requires part-of-speech tagging. As to the second requirement, a parallel German-English and French-English corpus is needed, on which qualitative investigations are to be performed. This purpose requires a corpus smaller than the comparable corpus used for the empirical analysis. The comparison of SL and TL texts moreover entails manual annotation to furnish semantic data. Both corpus-based investigations require a methodology of their own for linguistic interpretation, formal representation and corpus querying. As regards the last requirement, if the process-oriented approaches to translation studies do not provide insight into the cognitive processes of translation, recourse will have to be taken to psycholinguistic methods to develop a psycholinguistic experiment.

Thus, the investigation comprises three different approaches, namely the empirical, the cross-linguistic and the psycholinguistic approach. The empirical and the cross-linguistic approaches are corpus-based and therefore product-oriented, whereas the psycholinguistic approach is process-oriented. Since the cross-linguistic analysis involves the comparison of SL and TL texts, it serves as a basis for the explanation of the influence the SL has on the results of the empirical analysis, which compares translations to originals written in the same language as the TL. The psycholinguistic experiment, however, is intended to clarify and explain the results gained from both corpus analyses. As mentioned in Section 1.1, the dependent and independent variables vary according to the three different approaches (empirical, cross-linguistic and psycholinguistic) of the investigation of translated text (see Table 1.1).

	dependent variable	independent variable
empirical approach	English translations vs. English originals	large comparable corpus
cross-linguistic approach	English translations vs. German and French SL texts	small parallel corpus
psycholinguistic approach	cognitive processes during the translation process	psycholinguistic experiment

Table 1.1: Dependent and independent variables

The independent variables are a large comparable corpus for the empirical approach vs. a smaller parallel corpus for the cross-linguistic approach vs. a psycholinguistic experiment for the psycholinguistic approach. The relation between English translations and English originals in terms of quantitative patterns vs. the relation between English translations and their German and French originals in terms of qualitative patterns vs. the cognitive processes during the translation process are the dependent variables. Register, typology and cognitive processes are not considered separately as variables since they are interacting factors influencing the properties of translated text. However, the empirical approach focuses on register-specific language use as a large comparable corpus controlled on the basis of the notion of register is analyzed. A parallel corpus of SL and TL texts is investigated, hence, the cross-linguistic approach mainly deals with the typological differences

between the languages involved. The psycholinguistic approach is designed to reveal the cognitive processes taking place during the translation process. Since these different approaches combine methods from translation studies, corpus linguistics, contrastive linguistics, comparative linguistics, computational linguistics, psycholinguistics and linguistics in general, the posed requirements can be met on the basis of product- and process-oriented research. Thus, the nature of translated text and the sources of the specific properties of translations can be investigated with the help of the interdisciplinary methodology developed within the framework of this thesis.

1.4 Outline

The distinction between empirical, cross-linguistic and psycholinguistic research being a leitmotif of this thesis, each chapter includes three sections on empirical, cross-linguistic and psycholinguistic research, which can, of course, be split up in further subsections, as well as an introductory and a concluding section. Thus, empirical, cross-linguistic as well as psycholinguistic issues are dealt with in each chapter of this thesis. The main contents of the chapters are described in the following:

- **Chapter 2** discusses major approaches in corpus-based translation studies and corpus-based register analysis, which are used as a basis for the empirical investigation carried out in this thesis. Furthermore, it presents major approaches to cross-linguistic description and an overview of the methods employed in experimental psycholinguistics. Thus, this chapter constitutes the theoretical framework of the present thesis, introducing the basic research questions and methods of the empirical, cross-linguistic and psycholinguistic approaches of the present thesis.

- **Chapter 3** discusses the advantages and disadvantages of the empirical, cross-linguistic and psycholinguistic methods introduced in Chapter 2. On this basis, the theoretical concepts relevant to the present thesis are selected and new ones are introduced.

- **Chapter 4** discusses the empirical approach and the cross-linguistic approach developed in this thesis. Since, in connection with both ap-

proaches, corpora are investigated, issues concerning corpus annotation, corpus representation and corpus querying are approached. Furthermore, the psycholinguistic experiment is presented.

- **Chapter 5** presents the results of the empirical analysis, the cross-linguistic investigation as well as the psycholinguistic experiment. For each part of the analysis, hypotheses, results and interpretations are discussed.

- **Chapter 6** concludes the thesis with a summary of the findings gained regarding the nature of translated text. Moreover, the methodology is assessed and proposals for future work are brought up.

As can be seen, every chapter is structured according to empirical, cross-linguistic and psycholinguistic research and thus integrates product- and process-oriented research by combining corpus-based as well as psycholinguistic techniques.

Chapter 2

State of the art

2.1 Introduction

As described in Chapter 1, the investigation of the nature of translated text is a focal issue in translation studies. Within this context, the following translation-based properties have been discussed[1]:

- *translation strategies*, *translation procedures* and the use of *equivalence* (cf. [Nida, 1964], [Newmark, 1988] and [Vinay & Darbelnet, 1995]),
- *retrospective vs. prospective translation* (cf. [Wilss, 1977] and [Snell-Hornby, 1988]),
- *adapting vs. transferring translation* (cf. [Koller, 1984]),
- *translation text types* (cf. [Sager, 1994]),
- *law of growing standardization vs. law of interference* (cf. [Toury, 1995]),
- *rationalization*, *clarification*, *expansion* etc. (cf. [Berman, 2000]),
- *explicitation*, *simplification*, *normalization* and *levelling out* (cf. [Baker, 1996], [Laviosa-Braithwaite, 1996b] and [Olohan, 2001]),
- *sanitization* (cf. [Kenny, 1998b]),
- *shining-through* (cf. [Teich, 2001a]),

[1]Note that the first six approaches to the nature of translated text were developed rather intuitively, whereas the latter ones are investigated on the basis of corpus analyses.

- *grammatical metaphor* (cf. [Steiner, 2001b]),
- *overt vs. covert translation* and the application of a *cultural filter* (cf. [House, 1977] and [House, 1997]).

This chapter, however, presents the theoretical issues on the basis of which the specific properties of translations can be investigated on an empirical, a cross-linguistic as well as a psycholinguistic basis. For this reason, an overview of research questions and techniques is given for each discipline. The detailed description of the methods used for empirical, cross-linguistic and psycholinguistic research is necessary, all of them being possible candidates for the interdisciplinary approach to the investigation of the specific properties of translations (cf. Chapters 4 and 5). For this reason, the advantages and disadvantages of the different methods as well as their eligibility for this thesis are discussed and assessed in Chapter 3.

Since the empirical analysis is essential for the cross-linguistic as well as the psycholinguistic analysis, the section on empirical linguistics is presented first (cf. Section 2.2). This section is split up into a subsection about corpus-based translation studies (cf. Section 2.2.1) and a subsection dealing with corpus-based register analysis (cf. Section 2.2.2). On this basis, the theoretical concepts regarding empirical linguistics can be discussed (cf. Section 3.2), which provides a profound basis for the methodology developed in Section 4.2 and the generation of hypotheses in Section 5.2. Furthermore, Section 2.2 is important for the description of the register-related properties occurring in translated text.
Secondly, different approaches to cross-linguistic description are introduced (cf. Section 2.3). Section 2.3.3 contains an outline of Systemic Functional Linguistics (SFL) and its multilingual methods of linguistic description. These methods serve as the basis for the theoretical foundation with respect to the cross-linguistic approach (cf. Section 3.3), the methodology developed in Section 4.3 and the analysis in Section 5.3. In order to compare the SFL approach taken on in this thesis to classical approaches used in contrastive linguistics and translation studies, the English-German and English-French comparative typologies are presented in Section 2.3.1 and the translation techniques in Section 2.3.2. Additionally, Section 2.3.1 serves as a basis for the interpretation of the specific properties of translations which

result from typological differences between languages, whereas Section 2.3.2 helps to explain the properties of translated text owing to the translation process.
Thirdly, Section 2.4 introduces different experimental methods used in psycholinguistics. On this basis, methods relevant to the present thesis are selected (cf. Section 3.4) and the experiment of Section 4.4 is developed. This approach is then used for the psycholinguistic analysis in Section 5.4. Finally, Section 2.5 summarizes the product- and process-oriented methods introduced in this chapter. This overview serves as a basis for the theoretical foundation discussed in Chapter 3.

2.2 Empirical linguistics

There are many different fields of study and professions related to language in which empirical methods play a very important role. This means that linguists require empirical data in order to substantiate their theories or to apply them. These empirical data cannot be provided through introspection, but with the help of corpus analyses. This is the case, for example, in the areas of lexicography, language teaching, computational linguistics, contrastive analysis, literary studies, sociolinguistics, psycholinguistics and cultural studies.
For this thesis, two corpus-based applications are of particular importance: Baker's approach to corpus-based translation studies (presented in Section 2.2.1 (cf. [Baker, 1996]))[2] and Biber's approach to corpus-based register studies (introduced in Section 2.2.2 (cf. [Biber, 1995])). Thus, this section outlines the state of the art of the empirical research relevant to this thesis.

2.2.1 Corpus-based translation studies

In translation studies, the role of corpora was traditionally restricted to their use to the applied branch of this discipline. In particular, it has been used in the fields of terminology, translation aids (e.g., to develop translation memories or machine translation programs), translation criticism and translation training (to improve the final product with the help of corpus-based contrastive analysis and the study of *translationese*). Corpus linguistics has

[2]This section also includes approaches related to or derived from Baker's approach.

only very rarely been considered in terms of its importance to the theoretical and descriptive branches of translation studies. Researchers even tried to ban translations from corpora because translated text was regarded as inferior compared to originals and it was not considered worth investigating because it is generally constrained by the presence of a fully articulated text in another language. [Sager, 1984] was one of the first researchers who saw the need to examine translations as a special kind of text production and to look into their special characteristics. Nevertheless, he thought that the value of a translation is dependent on that of its original text (cf. [Sager, 1994]). In contrast to Sager, [Baker, 1995] goes a step further by trying to exclude the influence of the SL on a translation in order to analyze characteristic patterns of translations independent of the SL. Within this context, [Baker, 1996] developed the following hypotheses on the universal features of English translations.

- **Explicitation.** Explicitation means that translators tend to render explicit implicit contents of the SL text in their translations. As a result, translated texts tend to contain less ambiguities than originally produced texts. Evidence of explicitation may, for example, be found in the text length (number of words of the individual texts), since, in many cases, translations are longer than texts produced originally in the TL or in the SL. This kind of analysis requires a comparison of SL texts and their translations on a text-by-text basis. Explicitation can also be analyzed in view of lexis and syntax, using a monolingual corpus of translated texts and a comparable corpus of original texts produced in the same language. Translations tend to use more explanatory vocabulary (e.g., "therefore", "consequently") and optional subordinators (e.g., "that") than in originally produced texts, thereby rendering implicit contents more explicit.

- **Simplification.** Simplification describes the tendency of translators to (consciously or unconsciously) simplify texts in order to improve the readability of their translations. Evidence of this tendency may, for instance, be found in the average sentence length: the mean sentence length of translated texts tends to be lower, as translators often break up long and complex sentences into two or more sentences in their translations in an effort to make the texts easier to read. Some

linguistic features indicating simplification (e.g., the use of finite structures in English translations as opposed to non-finite structures in English originals) may also, at the same time, be a sign of explicitation. Another linguistic feature which reflects simplification is punctuation. Translators often change the punctuation from a weaker to a stronger mark, often using semicolons or periods instead of commas and periods instead of semicolons. This can be seen as an attempt to make texts easier and structure them more clearly by strengthening the punctuation. Another piece of evidence of simplification consists in the lexical density of a corpus, lexical density being the ratio of lexical vs. grammatical words. It is calculated by first subtracting the number of function words from the total number of words. The number of lexical words thus obtained is divided by the total number of words and then multiplied by 100. In translations, the lexical density tends to be lower than the lexical density of originals. This means that translations contain more function words and fewer lexical words than originals and are thus easier to read. A further method to test simplification is the type-token ratio, that is the ratio of different tokens vs. running words. This percentage is determined by dividing the number of different lemmata (types) by the total number of words (tokens) and then multiplying the result by 100. Translators tend to use fewer types in translations than authors do in originals, and thus the type-token ratio of translations is lower than the type-token ratio of originals.

- **Normalization.** Normalization (or conservatism) means that translators tend to conform to the typical patterns of the TL or even to exaggerate their use. If, however, the status of the SL has an influence on the language use of the TL (like the influence of the English language on other languages in the area of software), normalization in translations is weakened, or even counteracted by a contrary tendency. If this is the case, the typical patterns of the SL are still visible in the translations. This universal feature also includes the tendency to normalize marked and ungrammatical structures. This often occurs in simultaneous or consecutive interpreting, where interpreters tend to finish unfinished sentences and to grammaticalize ungrammatical structures.

- **Levelling out.** In a corpus which consists of a sub-corpus of translations and a sub-corpus of texts originally produced in the TL, translations are more alike in terms of features such as lexical density, type-token ratio and average sentence length than the individual texts in the comparable corpus of SL and TL originals. This means that translators tend to use centered linguistic features, moving translations away from extremes.

In order to show how Baker's and similar hypotheses can be tested, in the following, four different corpus designs and two possible analysis scenarios respectively are introduced: the comparable corpus with Laviosa-Braithwaite's as well as Olohan and Baker's hypothesis testing (cf. [Laviosa-Braithwaite, 1996b] and [Baker & Olohan, 2000]), the parallel corpus with Kenny's as well as Doherty's analyses (cf. [Kenny, 1998b] and [Doherty, 1999b]), the bilingually comparable corpus with Teich's as well as Steiner's investigations (cf. [Teich, 2001a] and [Steiner, 2001b]) and the multilingually comparable corpus with Johansson's as well as Fabricius-Hansen's tests (cf. [Johansson, 1995] and [Fabricius-Hansen, 1999b]).[3]

The comparable corpus. A comparable corpus consists of two separate collections of texts in the same language, i.e. a sub-corpus of original texts in the language in question and a sub-corpus of translations into that language from any given SLs. Such a corpus allows the identification of universal features of translations, or more precisely, of patternings specific to translations. These features are analyzed irrespectively of the SLs involved. Thus, this kind of corpora cannot directly be used in order to train translators or to improve machine translation systems, but can only be employed for basic research in the field of translation studies. However, especially this kind of pure research is essential for the survival of this discipline because it fuels the generation of hypotheses and helps to test them.

The Translational English Corpus (TEC) is such a corpus, built up and maintained at the Centre for Translation Studies at the University of Man-

[3]There are of course other researchers who work with similar corpus designs. This list is not intended to be complete, but to introduce different corpus designs, where English and/or German translations and originals are involved, in combination with possible analysis scenarios. For this reason, those are presented who built up and analyzed the corpora (or sub-corpora) for the first time.

chester Institute of Science and Technology (UMIST). TEC consists of English translations translated from different SLs and is divided into several registers (the largest part being fiction). In order to generate hypotheses concerning the universal features of English translations (cf. [Baker, 1993], [Baker, 1995], [Baker, 1996], [Baker, 1998], [Baker, 2000b], [Baker, 2000a] and [Laviosa-Braithwaite, 1998]) and to develop a possible framework for hypotheses testing (cf. [Laviosa-Braithwaite, 1997], [Laviosa, 1998a], [Laviosa-Braithwaite, 1996b], [Baker & Olohan, 2000] and [Olohan, 2001]), a subcorpus has been compiled from the British National Corpus (BNC) (cf. [Laviosa-Braithwaite, 1996a], [Laviosa, 1998b], [Laviosa, 1997] and [Baker, 1999]), which is as comparable as possible to TEC (see Figure 2.1).

On the basis of the newspaper sub-corpora of TEC and BNC, [Laviosa-Braithwaite, 1996b] tested the hypotheses of simplification and levelling out by analyzing the average sentence length, the lexical density and the type-token ratio. She found that the average sentence length of translated texts in The Guardian is significantly lower than the average sentence length of English originals printed in the same newspaper, which confirms the simplification hypothesis. Furthermore, [Laviosa-Braithwaite, 1996b] discovered that the lexical density of translated texts in The Guardian is significantly lower than in originals in the same newspaper, which again supports the simplification hypothesis. Analyzing the type-token ratio of translated texts in The Guardian and comparing it to the type-token ratio of originals, [Laviosa-Braithwaite, 1996b] found that it is lower in translations than in originals of the same newspaper. This difference, however, is not statistically significant. Thus, the translated texts in The Guardian appear simplified in terms of their mean sentence length and their lexical density, but not in terms of type-token ratio. Additionally, [Laviosa-Braithwaite, 1996b] found that similarity in terms of lexical density, type-token ratio and mean sentence length is greater among the translated texts in The Guardian than in the case of the originals. Nevertheless, since mean sentence length was the only feature with a highly significant variance, levelling out could only be attested for this feature (cf. [Laviosa-Braithwaite, 1996b, 124 ff.]).

A possible method of testing the explicitation hypothesis consists in analyzing the optional syntactic features in translations and originals of the

same language. [Baker & Olohan, 2000] applied this method comparing TEC to BNC. They analyzed the *that*-connective vs. the *zero*-connective of the reporting verbs *say* and *tell* and found that the *that*-connective occurs more often in TEC than in BNC, whereas the number of *zero*-connectives is significantly higher in BNC compared to TEC (cf. [Baker & Olohan, 2000, 145 ff.]). As the results are statistically significant, these findings can be interpreted as the universal feature of explicitation for translated English.

Figure 2.1: The comparable corpus

The parallel corpus. A parallel corpus consists of SL texts and their translations into the TL. Such a corpus is not immediately user-friendly – for the corpus to be useful, it is necessary to identify which units in the sub-corpora are translations of each other. Such a parallel corpus is known as an aligned corpus, since it establishes an explicit link between the SL sentence and the corresponding passage in the translation. For this purpose, different alignment techniques are available, which help to align the sub-corpora automatically. Parallel corpora are very useful for computer-aided translator training, bilingual lexicography, the improvement of the performance of machine translation systems and as translation memories in translation practice.

Figure 2.2: The parallel corpus

An example of a parallel corpus is the German-English Parallel Corpus of Literary Texts (GEPCOLT) built up and held at the School of Applied Language and Intercultural Studies (SALIS) at the Dublin City University (DCU) (cf. [Kenny, 1997]). This corpus is used to test the normalization hypothesis on the basis of aligned SL and TL texts (cf. [Kenny, 1998a]), but it can also serve as a basis for generating and analyzing hypotheses concerning the translator's creativity (cf. [Kenny, 1999], [Kenny, 2000b], [Kenny, 2000a] and [Kenny, 2001]). Another example of the same kind of corpus is

the English-German parallel corpus of popular-scientific texts built up at the Humboldt University in Berlin (cf. [Doherty, 1999b]). For instance, this parallel corpus provides insight into the differences in information structure in translations between English and German (cf. [Doherty, 1991], [Doherty, 1993], [Doherty, 1996], [Doherty, 1997] and [Doherty, 1998a]). The design of GEPCOLT and the English-German parallel corpus of popular-scientific texts is illustrated in Figure 2.2.

Looking at English-German parallel texts, [Kenny, 1998b] developed the sanitization hypothesis (see Examples 2.1 and 2.2 taken from [Kenny, 1998b, 518 ff.]).

(2.1) a. English: I was a mad gambler then and one day I picked up my uncle's Giro, cashed it and had it on a dog which got beat. Nothing was said, but he may have been terribly hurt knowing what I'd done.

b. German: Ich war damals ein besessener Spieler, und eines Tages klaute ich das Scheckheft meines Onkels, schrieb einen dicken Scheck aus, löste ihn ein und setzte das Geld auf einen Hund, der verlor. Mein Onkel muß schrecklich betroffen und enttäuscht gewesen sein, als sich herausstellte, was ich getan hatte.

(2.2) a. English: Bob's eyes flickered over me, a brief three-cornered glance, a check for progress. I felt like some sort of time-bomb. Thirteen now, about to go off.

b. German: Bob musterte mich aus den Augenwinkeln, ein kurzer Dreiecksblick, ein Abschätzen meiner Fortschritte. Als ob ich eine Zeitbombe wäre. Nun schon dreizehn, kurz vor der Explosion.

In Example 2.1, the English word "Giro", which is an informal word for an unemployment support check, is translated with the German word "Scheckheft", which means "checkbook". This translation is more general and neutral, lacking the negative connotation of the SL expression. This example illustrates that the TL text is somehow tamed or sanitized. The

contrary tendency of sanitization can be seen in Example 2.2, where the English expression "three-cornered glance" is translated into the German word "Dreiecksblick". Both expressions are unusual collocations in the respective languages, but the translation preserves the same degree of originality and creativity as the original version does. In a parallel corpus such as GEPCOLT, more evidence of sanitization and/or of creativity in translations can be found.

[Doherty, 1999b] uses the English-German parallel corpus of popular-scientific texts to contrast the information structure of English originals and their German translations (see Example 2.3 taken from [Doherty, 1999b, 113]).

(2.3) a. English: It behaves in this manner when the covalent bond is polarised, which means that the electron-pair of the bond is attracted away from the hydrogen toward the other atom. This can occur when hydrogen is attached to nitrogen, oxygen, fluorine and a few other elements.

b. German: Dies geschieht (zum Beispiel in Verbindungen mit Stickstoff, Sauerstoff oder Fluor) durch eine Polarisierung der kovalenten Bindung, bei der das Elektronenpaar der Bindung vom Wasserstoff weg- und zum anderen Atom hingezogen wird.

In Example 2.3, three differences concerning information structure can be found between the English SL text and its German translation: the reduction of the English temporal clause ("when the covalent bond is polarised") to the German PP ("durch eine Polarisierung der kovalenten Bindung"), the change of the English subordinate clause ("that the electron-pair") to the German relative clause ("bei der das Elektronenpaar") and the embedding of the second sentence in the English original into the first sentence using brackets. Taking position and linearization into account, it can be said that the English original is more explicit than the German translation.

The bilingually comparable and parallel corpus. A bilingually comparable and parallel corpus consists of SL texts in a language A, their TL texts in a language B, SL texts in a language B and their TL texts in a language A. The important aspect is that the term refers to sub-corpora in two different languages, built up on the basis of similar design criteria (e.g.,

the same register), and their translations into the other language, respectively. On the basis of a bilingually comparable and parallel corpus, the following relations can be analyzed: the relation between originals in two languages, the relation between translations in two languages, the relation between SL texts and their translations and the relation between originals and translations in the same language. The investigation of these relations is interesting for research in contrastive linguistics as well as translation studies.

The English-German bilingually comparable and parallel corpus of popular-scientific texts compiled and kept at the Institute for Applied Linguistics, Translation and Interpreting at the Saarland University (cf. [Teich, 2001a] and [Teich & Hansen, 2001a]) is an example of this corpus type (see Figure 2.3). This corpus is to serve the investigation of the specific properties of translated text (cf. [Hansen & Teich, 1999], [Teich & Hansen, 2001b], [Hansen & Teich, 2001], [Teich, 2001b], [Teich, to appear], [Steiner, 2001b] and [Steiner, 2001a]), grammatical metaphor in translations (cf. [Steiner, to appear]) and typological differences between English and German (cf. [Steiner & Teich, to appear]). Furthermore, different concepts of corpus representation and querying have been tested on the basis of this corpus (cf. [Teich et al., 2001]).

Figure 2.3: The bilingually comparable and parallel corpus

[Teich, 2001a] used the bilingually comparable and parallel corpus to investigate cross-linguistic variation in translations and bilingually comparable originals. For this purpose, she analyzed the features typical of the register of popular-scientific writing in English and German originals and translations. One of these register features for English is the extensive use of

passive. In German, passive is a typical feature of popular-scientific texts as well. German, however, also offers a wide range of passive alternatives, which fulfill a similar function (i.e. not specifying or underspecifying the agent of a process). [Teich, 2001a] found a significant difference in the use of passive and passive alternatives in English originals compared to German originals, English using more passives than German. The extensive use of passive is therefore more typical of English than of German with its extensive use of passive alternatives. Comparing German originals and German translations, she found, however, that German translations show a higher frequency of passive than German originals. This finding indicates that the typical language use of the SL "shines through" in the German translations (i.e. SL shining-through). At the same time, passive alternatives are significantly more frequent in the German translations compared to the German originals, which shows that the language use of the TL is normalized concerning this register feature (i.e. TL normalization) (cf. [Teich, 2001a, 160 ff.]).

[Steiner, 2001b] analyzed grammatical metaphor in translations from English into German on the basis of examples taken from the bilingually comparable corpus (see Example 2.4 taken from [Steiner, 2001b, 20]).

(2.4) a. English: The suspicion that volcanic eruptions are the primary source of aerosols in the upper atmosphere has been around for many years.

b. German: Seit vielen Jahren vermutet man schon, daß die Aerosole in den höheren Schichten der Atmosphäre vor allem aus Vulkanausbrüchen stammen.

According to [Steiner, 2001b, 20 ff.], the de-metaphorization hypothesis (i.e. the encoding of the same ideational meaning by means of different phrasal categories, which express more explicit and congruent variants) can be confirmed since there are more verbal-type phrases than nominal-type phrases in the translation (e.g., the English NP "The suspicion" is de-metaphorized into the German VP "vermutet man"). More evidence of grammatical metaphorization or de-metaphorization in German-English and English-German translations can be found in the bilingually comparable and parallel corpus.

The multilingually comparable and parallel corpus. A multilingually comparable and parallel corpus consists of SL texts in a language A, their TL texts in the languages B and C, SL texts in a language B, their TL texts in the languages A and C, SL texts in a language C and their TL texts in the languages A and B. The important thing is that the term refers to sub-corpora in more than two different languages, built up on the basis of similar design criteria (e.g., the same register), and their translations into the other languages, respectively. On the basis of a multilingually comparable and parallel corpus, the following relations can be analyzed: the relation between originals in three languages, the relation between translations in three languages, the relation between SL texts and their translations into two languages and the relation between originals and translations in the same language as the TL. Similarly to the relations represented in a bilingually comparable and parallel corpus, the investigation of these relations is of interest in connection with research in contrastive linguistics, comparative linguistics as well as translation studies.

The Oslo Multilingual Corpus (OMC) is an example of such a corpus (see Figure 2.4). It is the extension of the English-Norwegian Parallel Corpus (ENPC) (cf. [Johansson & Hofland, 1994], [Johansson et al., 1996], [Johansson, 1998b], [Johansson & Hofland, 2000], [Johansson, 2001a] and [Johansson, 2001d]). Built up and maintained at the Faculty of Arts at the University of Oslo, this corpus is used for analyses in contrastive linguistics and translation studies (cf. [Johansson & Ebeling, 1996], [Johansson, 1997a], [Johansson, 1997b], [Johansson & Løken, 1997], [Johansson, 1998a], [Johansson & Hasselgård, 1999], [Wikberg et al., 1999], [Johansson, 2001b], [Johansson, 2001c] and [Johansson, 2001e]) as well as for investigations regarding thematic structure in translation (cf. [Hasselgård, 1998]) and information packaging in translation (cf. [Fabricius-Hansen, 1998], [Fabricius-Hansen, 1999b] and [Fabricius-Hansen, 1999a]).

[Johansson, 1995, 23] analyzed the English and Norwegian sub-corpora of the OMC and found that the number of words in English translations exceeds the number of words in the respective Norwegian originals by 10%. Moreover, a slight increase could be recognized in terms of text length for

Norwegian translations compared to their English originals. Differences in text length (i.e. the number of words) are often due to morphological differences between languages, yet in Johansson's finding an increase in text length for both translation directions is particularly interesting, as it confirms the explicitation hypothesis.

Figure 2.4: The multilingually comparable and parallel corpus

[Fabricius-Hansen, 1999b] analyzed the German-English and German-Norwegian parallel sub-corpora of the OMC in view of information packaging in German-English translations (see Example 2.5 taken from [Fabricius-Hansen, 1999b, 177]).

(2.5) a. German: Die Jungen sind bei den plakatfarbigen Arten fast immer noch bunter, noch ortstreuer und noch bösartiger als die Erwachsenen, so daß man die zu untersuchenden Vorgänge an diesen Miniaturfischchen auf verhältnismäßig beschränktem Raum gut beobachten kann.

b. English: In the poster-coloured species, the young are nearly always not only more colourful and fiercer but also more firmly attached to their territories than the adults are. Since the young are small, we could observe their behavior in a comparatively limited space.

Example 2.5 shows clausal expansion (translating the German adjective "ortstreuer" into the English reduced relative clause "more firmly attached to their territories") as well as sententialization (dividing the German clause complex into two separate sentences).

As can be seen in Figures 2.1 to 2.4, the parallel and comparable corpora are sub-corpora of the bilingually comparable and parallel corpus and the bilingually comparable and parallel corpus is a sub-corpus of the multilingually comparable and parallel corpus. The multilingually comparable and parallel corpus could, again, form part of a larger corpus design consisting of a large number of sub-corpora in various languages and their translations into the other languages. Such a comprehensive corpus design could be the overall aim, integrating all corpus initiatives which build up parallel and comparable corpora worldwide.

2.2.2 Corpus-based register analysis

In general, register is described as functional variation (cf. [Quirk et al., 1985] and [Biber et al., 1999]), where language varies in field of discourse, medium and attitude. Variation according to the field of discourse depends on the activity the participants are engaged in. Medium refers to varieties such as written vs. spoken language. Attitude describes the speaker's intention, the relationship between speaker and addressee(s) and the purpose of the communication.
SFL looks at register from a different viewpoint, offering a theory of language variation according to situational contexts (cf. [Halliday, 1985], [Halliday & Hasan, 1989], [Matthiessen, 1993] and [Martin, 1992]). This kind of register theory focuses on the systematic relation between a context of situation and the language variety used in that context, i.e. the register. SFL defines language variation within the context of situation along the following parameters: field, tenor and mode of discourse. The field of discourse refers to social activity, tenor characterizes the relationship between speaker and listener, i.e. the participant and role relationships, and mode is concerned with the role played by language itself in a given context of situation. A more detailed concept of the context of situation within the SF register analysis is illustrated in Figure 2.5.

This or a similar kind of register analysis is widely used in text linguistics (cf. [Ghadessy, 1988], [Ghadessy, 1993], [Kittredge & Lehrberger, 1982] and [Ure & Ellis, 1977]). Applications of register analysis have recently been adopted in translation studies, too: register theory is used to assess the quality of translations (cf. [House, 1997] and [Hatim & Mason, 1990]), to evaluate translations (cf. [Steiner, 1998]) or to analyze SL texts as a preparatory step for translation (cf. [Steiner, 1997]).

Figure 2.5: Context of situation within SF register analysis

Another important application of register analysis can be found in the corpus-based description of registers, i.e. the lexico-grammatical features which occur significantly frequently or seldom in a given register are empirically analyzed and described (cf. [Biber et al., 1998] and [Biber & Finegan, 1993]). On the basis of such a corpus-based frequency analysis, the boundaries between registers can be identified in terms of the typical or untypical features of each register. In this context, factor analysis constitutes an empirical method which can be used to statistically analyze the co-occurrence of lexico-grammatical features in large corpora (cf. [Biber, 1990], [Biber, 1992] and [Biber, 1993]). This method can be applied in connection with the analysis of monolingual corpora (cf. [Biber, 1988]) as well as multilingual corpora (cf. [Biber, 1995] and [Biber & Hared, 1992]). [Biber, 1995], for example, employs cluster analysis to investigate different registers in four languages: English, Nukulaelae Tuvaluan, Korean and Somali. For English[4], his studies are based on the LOB Corpus and the London-Lund Corpus (cf. [Biber, 1995, 86 ff.]), including seventeen written registers (press

[4]This section focuses on Biber's synchronic investigations for English. The diachronic investigations and the analyses of the three remaining languages are neglected, since they are of no relevance to this study.

reportage, editorials, press reviews, religion, skills/hobbies, popular lore, biographies, official documents, academic prose, general fiction, mystery fiction, science fiction, adventure fiction, romantic fiction, humor, personal letters and professional letters) and six spoken registers (face-to-face conversation, telephone conversation, public conversations/debates/interviews, broadcast, spontaneous speeches and planned speeches). Altogether the corpus consists of 960,000 words. [Biber, 1995] automatically analyzed and provided the corpus with annotation encoding lexical, grammatical and syntactic information. In order to find register boundaries, he performed frequency counts on the following sixteen grammatical and functional categories:

- tense and aspect markers,
- place and time adverbials,
- pronouns and pro-verbs,
- questions,
- nominal forms,
- passives,
- stative forms,
- subordination features,
- PPs, adjectives and adverbs,
- lexical specificity,
- lexical classes,
- modals,
- specialized verb classes,
- reduced forms and discontinuous structures,
- co-ordination,
- negation.

The normed frequency counts of these categories served as a basis for the factor analysis [Biber, 1995] carried out. On the basis of the shared variance of the analyzed features, he then identified feature clusters and interpreted the results, classifying the co-occurring features into the following seven basic functional dimensions of variation:

- Dimension 1 – involved vs. informational production,
- Dimension 2 – narrative vs. non-narrative discourse,
- Dimension 3 – situation-dependent vs. elaborated reference,
- Dimension 4 – overt expression of argumentation,
- Dimension 5 – abstract vs. non-abstract style,
- Dimension 6 – on-line informational elaboration marking stance,
- Dimension 7 – academic hedging[5].

In the following, each functional dimension is briefly introduced by outlining its typical features and registers and contrasting these with the untypical features and registers of the respective functional dimension (cf. [Biber, 1995, 141 ff.]).

Dimension 1 – Involved vs. informational production. The positive features in Dimension 1 (see Table 2.1)[6] can be interpreted as features of involved production and interaction. For example, the co-occurrence of private verbs and the verb *be*, together with present tense, is a strong indicator of the speaker's involvement in the communicative situation. These features are characteristic of the expression of comments, attitudes and feelings, making the situation more interactive. On the basis of the co-occurrence of the positive features in Dimension 1, telephone conversations and face-to-face conversations can be considered as typical examples of involved production.
In contrast to that, the negative features are characteristic of informational

[5]This is only a tentative interpretation by [Biber, 1995] because the results of the factor analysis are not significant.

[6]All the tables in this section are taken from [Biber, 1995, 141 ff.]. The figures in the tables reflect the results of Biber's factor analysis. The features in brackets are not considered for Biber's interpretation of the results since their statistical values are too low.

production (see Table 2.1). A high frequency of nouns co-occurring with a high frequency of attribute adjectives results in a high information density. Additionally, PPs which occur as post-modifications and thus as specifications of NPs suggest that the information in the text is presented in a very dense fashion. Typical registers reflecting the features for informational production are official documents, press reports and academic prose.

Dimension 2 – Narrative vs. non-narrative discourse. The positive features of Dimension 2 reflect narrative discourse, whereas the negative features reflect non-narrative discourse (see Table 2.2). Past tense and perfect aspect, for example, are used to describe major processes in the past, which is crucial in narrative texts. Another positive feature is the use of third person pronouns. This feature is used to refer to participants, which is frequently the case in narrations. Furthermore, using public verbs, a narrator introduces direct or indirect speech acts of the participants involved in the narration. The classification of the positive features into narrative discourse is exemplified by the registers which incorporate these features: romance fiction, mystery fiction, science fiction, general fiction and adventure fiction, i.e. all sub-registers of fiction.
The negative features of Dimension 2, i.e. present tense and attributive adjectives (see Table 2.2), are used to describe contents. The occurrence of these features renders texts more expository and informational. Typical registers which reflect these characteristics are broadcasts, hobbies, official documents and academic prose.

Dimension 3 – Situation-dependent vs. elaborated reference. The positive features of Dimension 3 (mainly time adverbials and place adverbials) convey local and temporal information (see Table 2.3). Since this kind of information always refers to the actual physical context, these features can be grouped into situation-dependent features which can typically be found in broadcasts and telephone conversations.
In contrast to that, the negative features of Dimension 3, such as *wh*-relative clauses in subject and object position and pied-piping (see Table 2.3), specify the element they refer to. Thus, they make the text more explicit and context-independent. Official documents and professional letters are registers which are typical of this kind of elaborated reference.

Dimension 1	
positive features:	
private verbs	0.96
that-deletion	0.91
contradictions	0.90
present tense verbs	0.86
second person pronouns	0.86
do as pro-verb	0.82
analytic negation	0.78
demonstrative pronouns	0.76
general emphatics	0.74
first-person pronouns	0.74
pronoun *it*	0.71
be as main verb	0.71
causative subordination	0.66
discourse particles	0.66
indefinite pronouns	0.62
general hedges	0.58
amplifiers	0.56
sentence relatives	0.55
wh-questions	0.52
possibility modals	0.50
non-phrasal co-ordination	0.48
wh-clauses	0.47
final prepositions	0.43
(adverbs	0.42)
negative features:	
nouns	-0.80
word length	-0.58
prepositions	-0.54
type-token ratio	-0.54
attributive adjectives	-0.47
(place adverbials	-0.42)
(agentless passives	-0.39)
(past participial postnominal clauses	-0.38)

Table 2.1: Dimension 1 – involved vs. informational production

Dimension 2	
positive features:	
past tense verbs	0.90
third-person pronouns	0.73
perfect aspect verbs	0.48
public verbs	0.43
synthetic negation	0.40
present participial clauses	0.39
negative features:	
(present tense verbs	-0.47)
(attributive adjectives	-0.41)

Table 2.2: Dimension 2 – narrative vs. non-narrative discourse

Dimension 3	
positive features:	
time adverbials	0.60
place adverbials	0.49
adverbs	0.46
negative features:	
wh-relative clauses in object positions	-0.63
pied-piping constructions	-0.61
wh-relative clauses in subject positions	-0.45
phrasal co-ordination	-0.36
nominalizations	-0.36

Table 2.3: Dimension 3 – situation-dependent vs. elaborated reference

Dimension 4 – Overt expression of argumentation. In Dimension 4, only positive features can be found (see Table 2.4). These features, e.g., prediction modals, suasive verbs, necessity modals and possibility modals, express different positions the speaker and the receiver can be involved in, possible alternatives they can opt for and the arguments in favor of or against these alternatives. Professional letters and editorials exemplify this kind of persuasion or overt expression of argumentation.
The opposite of these overtly argumentative registers are, for example, broadcasts.

Dimension 5 – Abstract vs. non-abstract style. In Dimension 5, solely negative features, such as agentless passives, *by*-passives and past participials as adverbial and postnominal clauses, can be found (see Table 2.5). Since the past participial constructions are used to express passive

meaning, all of these features draw attention away from the agent and focus on the patient or the action itself. This focus results in a rather abstract representation of the contents, which is typical of, for example, technical and engineering prose, academic prose in general and official documents.

Dimension 4	
positive features:	
infinitives	0.76
prediction modals	0.54
suasive verbs	0.49
conditional subordination	0.47
necessity modals	0.46
split auxiliaries	0.44
(possibility modals	0.37)
[no negative features]	

Table 2.4: Dimension 4 – overt expression of argumentation

In contrast to the registers representing this kind of abstract style, non-abstract text production can be found in telephone and face-to-face conversations, in personal letters, spontaneous speeches and all sub-registers of fiction.

Dimension 5	
[no positive features]	
negative features:	
conjuncts	-0.48
agentless passives	-0.43
past participial adverbial clauses	-0.42
by passives	-0.41
past participial postnominal clauses	-0.40
other adverbial subordinators	-0.39

Table 2.5: Dimension 5 – abstract vs. non-abstract style

Dimension 6 – On-line informational elaboration marking stance. The positive features of Dimension 6 are, among others, *that*-clauses complementing verbs or adjectives and *that*-relative clauses in object positions (see Table 2.6). These features are strong indicators of informational spoken registers. On the one hand, they are informational since *that*-clauses

describe adjectives or nouns in greater detail, or specify what has been said. On the other hand, they are typical of spoken registers, given that they are produced under real-time constraints and not edited. Prepared and spontaneous speeches as well as public conversations are therefore typical registers. At the other end of the scale, the sub-registers of fiction represent edited-style text production.

Dimension 6	
positive features	
that-clauses as verb complements	0.56
demonstratives	0.55
that-relative clauses in object positions	0.46
that-clauses as adjective complements	0.36
(final prepositions	0.34)
(existential *there*	0.32)
(demonstrative pronouns	0.31)
(*wh*-relative clauses in object positions	0.30)
negative features:	
(phrasal co-ordination	-0.32)

Table 2.6: Dimension 6 – on-line informational elaboration marking stance

Dimension 7 – Academic hedging. In Dimension 7, the results of the factor analysis are not significant (see Table 2.7). Yet, since the features are coherent, [Biber, 1995] provides a tentative interpretation for this dimension. The positive features of Dimension 7 include, among others, the verbs *seem* and *appear* as well as downtoners. Expressing uncertainty, these features are typical of academic argumentation which involves hedging.

As can be seen in the descriptions of Dimension 1 to Dimension 7, by means of their distributional patterns, the registers can be assigned to the different ends on each dimension scale (e.g., public conversations and academic prose are frequently located on opposite directions). However, they also show that certain registers can be grouped together, and that they nearly always belong to the same extreme of each dimension scale (e.g., the sub-registers of fiction are frequently classified as one group). Thus, the register analysis as carried out by [Biber, 1995] clearly shows patterns of register classification which can be broken down into lexico-grammatical features.

Dimension 7	
positive features	
seam/appear	0.35
(downtoners	0.33)
(adverbs	0.31)
(concessive adverbial subordination	0.30)
(attributive adjectives	0.30)
[no negative features]	

Table 2.7: Dimension 7 – academic hedging

2.3 Cross-linguistic description

In empirical linguistics, it is important not only to work with large amounts of texts, but also to describe the results that have been gained against the background of a theoretical language model which offers a profound framework for interpretation. Hence, different methods of cross-linguistic description are introduced in this section. There are many theories regarding the description of language, but only few of them provide a comprehensive framework for the description of multilingual texts. The classical methods used to describe the phenomena occurring in comparative linguistics (cf. Section 2.3.1) and translation studies (cf. Section 2.3.2) are presented and discussed. Recently, different approaches have evolved in SFL which involve methods of describing contrastive data, on the one hand, and translated data, on the other (cf. Section 2.3.3). Thus, this section presents the state of the art of the cross-linguistic research which is of interest to this thesis.

2.3.1 Comparative typology

The analysis of multilingual texts requires knowledge of the major contrasts of the languages involved (cf., for example, [Haspelmath et al., 2001] or [Lang & Zifonun, 1996]). Since the present study focuses on the translation directions German-English and French-English, this section presents Hawkins' comparative typology of English and German (cf. [Hawkins, 1986, 11 ff.] and see also [Hawkins, 1981]) and Vinay and Darbelnet's comparative

stylistics of English and French (cf. [Vinay & Darbelnet, 1995, 99 ff.])[7].

The major contrasts between English and German can be found in terms of grammatical morphology, word order freedom, the basic grammatical relations and their semantic diversity, raising structures, extractions and deletions. In the following, each of these contrasts is introduced briefly and illustrated through examples (cf. [Hawkins, 1986, 11 ff.]).

Grammatical morphology. The grammatical morphology of German is richer than that of English. Basically, the set of morphological options in English constitutes a subset of the morphological options available in German. For example, both languages distinguish between singular and plural and between present and past tense. However, there are differences, for example, in gender distribution and case marking. The latter distinction is illustrated in Examples 2.6 to 2.8.

(2.6) a. German: Ich schlafe. Ich friere.
b. English: I am sleeping. I am freezing.

(2.7) a. German: Mich hungert. Mich friert.
b. English: I am hungry. I am freezing.
c. English: *Me hungers. Me freezes.

(2.8) a. German: Mir ist warm.
b. English: I am warm.
c. English: *To me is warm.

The English and German verbs used in Example 2.6 typically require nominative cases in both languages. Additionally, both accusative (see Example 2.7) and dative (see Example 2.8) can be found in German impersonal constructions. A direct translation into the corresponding English accusative (see Example 2.7(c)) and dative (see Example 2.8(c)) constructions is not grammatical. An appropriate English translation entails the use of the nominative case (see Examples 2.7(b) and 2.8(b)).

[7]The examples are taken either from [Hawkins, 1986, 11 ff.] or from [Vinay & Darbelnet, 1995, 99 ff.].

Word order freedom. Whereas, in English, the constituents of the sentence are determined by word order, which is more fixed than German word order, the rich grammatical morphology of German makes it possible to determine the constituents of a sentence on the basis of their morphology. Therefore, the options available in English form a subset of the options available in German, as shown in Examples 2.9 and 2.10.

(2.9) a. English: John saw the boy.

b. English: John has seen the boy.

c. English: I know that John saw the boy.

d. English: I know that John has seen the boy.

(2.10) a. German: Johann sah den Jungen.

b. German: Johann hat den Jungen gesehen.

c. German: Ich weiß, daß Johann den Jungen sah.

d. German: Ich weiß, daß Johann den Jungen gesehen hat.

Example 2.9 illustrates that the English word order is SVO independent of the tense which is used (see Examples 2.9(a) and 2.9(b)) or the type of clause it occurs in (see Examples 2.9(c) and 2.9(d)). In German, however, word order varies according to the tense which is used (see Examples 2.10(a) and 2.10(b)) and it is realized differently in main clauses and subordinate clauses (see Examples 2.10(c) and 2.10(d))[8].

Basic grammatical relations and their semantic diversity. The set of German realizations concerning subjects and direct objects is a subset of the English ones. Consequently, English is richer than German in terms of the semantic diversity of subjects and direct objects. This contrast is illustrated in Examples 2.11 and 2.12.

[8] A more detailed overview on contrastive word order as well as word order universals can be found in [Hawkins, 1983] and [Hawkins, 1984].

(2.11) a. English: This hotel forbids dogs.
b. German: In diesem Hotel sind Hunde verboten.
c. German: *Dieses Hotel verbietet Hunde.

(2.12) a. English: This tent sleeps four.
b. German: In diesem Zelt schlafen vier.
c. German: *Dieses Zelt schläft vier.

In Example 2.11, the English sentence uses a non-agentive NP subject, which cannot be directly translated into German. The use of such a non-agentive NP subject in the German sentence results in a personification which is not acceptable in German (see Example 2.11(c)). In Example 2.11-(b), a locational circumstantial in the form of PP is used to produce an appropriate German translation, thereby realizing the English active construction including a non-agentive NP subject by means of a German passive construction involving a locational PP circumstantial. In Example 2.12, the same transformation in terms of grammatical relations has taken place, the active voice, however, has been maintained.

Raising structures. In German, merely a handful of verbs allow subject-to-subject raising, which is a subset of the English verbs triggering these structures. One of these verbs is shown in Example 2.13.

(2.13) a. English: John seems to be ill.
b. German: Johann scheint krank zu sein.

(2.14) a. English: John happens to be ill.
b. German: *Johann geschieht krank zu sein.

(2.15) a. English: I believe John to be ill.
b. German: *Ich glaube Johann krank zu sein.

As to the use of raising structures, English is more productive than German. Many English verbs allow subject-to-subject raising, which cannot be translated into German by a raising structure (see Example 2.14). While

common in English, subject-to-object raising is not possible in German at all (see Example 2.15).

Extractions. The extraction of complements is possible both in English and in German. Again, similarly to raising constructions, the means of extraction available in German are fewer than the options available in English (see Examples 2.16 and 2.17).

(2.16) a. English: The man who I have tried to kill was your friend.

b. German: Der Mann, den ich zu töten versucht habe, war dein Freund.

(2.17) a. English: I don't know who the police thought that the guilty man was.

b. German: *Ich weiß nicht, wer die Polizei glaubte, daß der Schuldige sei.

As illustrated in Example 2.16, extraction out of an infinitival object complement is possible in both languages. Nonetheless, the extraction of finite object complements, while allowed in English, is not acceptable in German (see Example 2.17).

Deletions. Whenever German permits a deletion, it is possible in English as well, since German only offers a subset of the options available in English (see Examples 2.18 to 2.22).

(2.18) a. English: Fred saw the king and greeted the king.

b. English: Fred saw and greeted the king.

(2.19) a. German: Fritz sah den König und begrüßte den König.

b. German: Fritz sah und begrüßte den König.

(2.20) a. English: Fred saw the king and thanked the king.

b. English: Fred saw and thanked the king.

(2.21) a. German: Fritz sah den König und dankte dem König.

b. German: *Fritz sah und dankte dem König.

(2.22) a. English: The door opens.

b. German: Die Tür öffnet sich.

Example 2.18 shows the deletion of the English NP "the king". This NP deletion is possible in German, too (see Example 2.19), since both verbs ("sehen" and "begrüßen") take accusative case in German ("den König"). In Example 2.21 the German verb "sehen" takes accusative, whereas the German verb "danken" takes dative case. The combination of accusative and dative makes the NP deletion impossible in German, whereas it is still acceptable in English (see Example 2.20).
In Example 2.22, the reflexive noun, which can be found in the German sentence, is avoided through the use of a middle construction in the English sentence, showing another kind of deletion.

In general, it can be said that there are overlaps and interactions between the above-introduced typological contrasts. For example, due to the rather rich grammatical morphology in German, word order is relatively free, whereas with its rather restricted grammatical morphology, English offers more options for deletion. A summary of the contrastive English-German typology is displayed in Table 2.8.

English	German
less grammatical morphology	more grammatical morphology
less word order freedom	more word order freedom
more semantic diversity of GR	less semantic diversity of GR
more raising	less raising
more extraction	less extraction
more deletion	less deletion

Table 2.8: Contrastive English-German typology

With a view to analyzing the contrasts between English and French, [Vinay & Darbelnet, 1995] investigate the predominance of the French noun, the verb and the flow of action, different kinds of supplementation, gender, number, characterization, time/tense, voice, modality, verbal aspects, phraseology and ellipsis. In the following, each of these contrasts is introduced

briefly (cf. [Vinay & Darbelnet, 1995, 99 ff.]).

Predominance of the French noun. In French, the semantic focus usually is on a noun rather than on another word class. Thus, French nouns can be found in cases in which English uses verbs, adjectives or prepositions (see Examples 2.23 to 2.25).

(2.23) a. English: to total

b. French: atteindre le total de

(2.24) a. English: a hopeless undertaking

b. French: une entreprise sans espoir

(2.25) a. English: within two weeks

b. French: dans un délai de deux semaines

In Example 2.23, the English verb "total" is translated into the French expression "atteindre le total de", where "atteindre" is a verb with a rather general meaning and the focus is on the noun "total". Example 2.24 shows a similar phenomenon, in which the English adjective "hopeless" is translated into the French PP "sans espoir". In Example 2.25, the English preposition "within" is translated into the French PP "dans un délai de", in which the semantic head "délai" is added to the PP.

The verb and the flow of action. With regard to the verb and the flow of action, it can be said that English follows the straightforward order, such as the temporal sequence, whereas French usually puts the result before the means (see Examples 2.26 to 2.27).

(2.26) a. English: He gazed out of the open door into the garden.

b. French: Il a regardé dans le jardin par la porte ouverte.

(2.27) a. English: blown away

b. French: emporté par le vent

Example 2.26 shows how English follows the natural flow of action: first the gaze passes the open door, then it reaches the garden. In French, however, the gaze reaches the garden followed by an explanation how it reaches there. The same phenomenon is illustrated in Example 2.27, where in the English construction, the action is described before the result and in the French construction, it is the other way round.

Supplementation. In French, it is unusual for some function words to be used on their own. For this reason, they are strengthened by the supplementation of another word (see Examples 2.28 to 2.29).

(2.28) a. English: passengers to Paris

b. French: voyageurs à destination de Paris

(2.29) a. English: He stopped at the desk for his mail.

b. French: Il s'arrêta au bureau pour prendre son courier.

Example 2.28 shows how the English preposition "to" is translated into the French preposition "à" which is reinforced by the nominal construction "destination de". The supplementation through nouns is the most common form of supplementation and supports the above-mentioned predominance of the French noun. However, supplementation through other word classes is possible as well: Example 2.29 illustrates how the English PP "for his mail" is translated into the French PP "pour son courier" which is supplemented by the verb "prendre".

Gender. Both English and French distinguish between natural gender (i.e. male, female and asexual being) and grammatical gender (i.e. masculine, feminine and neuter). English has almost completely lost the grammatical gender and focuses on the natural gender, whereas French still distinguishes between the grammatical gender masculine and feminine and is thus dominated by grammatical gender. Nevertheless, French inhibits ambiguities such as the one illustrated in Example 2.30. The possessive pronoun in

French assumes the grammatical gender of the possessed noun instead of the possessor's natural gender, which is the case in English.

(2.30) a. English: his hat - her hat

b. French: son chapeau

(2.31) a. English: our readers

b. French: nos lecteurs - nos lectrices

(2.32) a. English: my cousin

b. French: mon cousin - ma cousine

Examples 2.31 and 2.32, however, illustrate that the natural gender is also taken into account in French. In these cases, the corresponding English expressions are more ambiguous.

Number. English frequently uses nouns which remain singular but express a collective meaning. In an English-French translation these nouns always have to be translated by a French plural noun (see Examples 2.33 and 2.34).

(2.33) a. French: des meubles - un meuble

b. English: furniture - a piece of furniture

(2.34) a. French: des éclairs - un éclair

b. English: lightning - a flash of lightning

In order to express the singular meaning of collective nouns, English uses a special form of expression, called the singulative, as shown in Example 2.33. Such collective nouns are unusual in French, yet there are exceptions which also use the singulative in French (e.g., "de la monnaie - une pièce de la monnaie"). In English, however, the singulative does not necessarily have to be expressed through "a piece of", as shown in Example 2.34.

Characterization. In English, characterization is usually expressed through adjectives or adverbs, whereas French tends to use nouns for characterization. This tendency again supports the predominance of the French noun (see Examples 2.35 and 2.36).

(2.35) a. English: a rural church

b. French: une église de campagne

(2.36) a. English: tolerantly

b. French: avec tolérance

In Example 2.35, the English adjective "rural" characterizes the English noun "church". In the French translation, the characterization of the noun "église" is realized through another noun within a PP ("de campagne"). A similar phenomenon is shown in Example 2.36: the English characterization adverb "tolerantly" is translated into the French PP "avec tolérance", in which the characterization is realized by the noun "tolérance".

Time and tense. Although the number of different tenses in English is comparable to that of French, their distributions and functions are divergent. This phenomenon is illustrated in Examples 2.37 to 2.43 on the basis of the use of present tense in both languages.

(2.37) a. English: I write.

b. French: J'écris.

(2.38) a. English: I am writing.

b. French: J'écris.

(2.39) a. English: I do write.

b. French: J'écris.

(2.40) a. English: I have been writing ...

b. French: J'écris depuis ...

(2.41) a. English: Nothing will last forever.
b. French: Rien ne dure.

(2.42) a. English: I have come to tell you that ...
b. French: Je viens vous dire ...

(2.43) a. English: I am coming.
b. French: Je viens.

Example 2.37 illustrates present for both languages, whereas Examples 2.38 to 2.40 contain English aspects which are often ignored in French or translated lexically. In Example 2.41 English uses future to express a prediction, whereas French uses present tense. The English Examples 2.42 and 2.43 convey semantic nuances which are ignored in French: in Example 2.42, the speaker has arrived shortly before the speech act, whereas in Example 2.43, the speaker is not yet there and announces his arrival.

Voice. Both languages, English and French, distinguish between three voices: active, passive and middle. They differ, however, in terms of how the middle is formed and when it is used (see Examples 2.44 to 2.46).

(2.44) a. English: The door is closed.
b. French: La porte est fermée.

(2.45) a. English: The door is closed (automatically).
b. French: La porte se ferme (automatiquement).

(2.46) a. English: The door closes by itself.
b. French: La porte se ferme d'elle-même.

Instances of how English and French form and use passive voice are given in Example 2.44. In Example 2.45, passive is again contained in the English sentence, whereas middle is used in the French sentence. Owing to the middle construction, the patient is in subject position and the agent is omitted. In Example 2.46, both languages use middle constructions. The French

middle construction involves reflexivation, whereas the English middle uses the finite verb form.

Modality. Modality expresses the speaker's attitude in English as well as in French. Besides auxiliary verbs, French uses the subjunctive to express modality, whereas the subjunctive is rare in English (refer to Examples 2.47 and 2.48).

(2.47) a. French: Je ne dirai pas qu'il l'a/l'ait fait exprès.

b. English: I won't tell anyone he did it on purpose.

c. English: I won't go so far as to say he did it on purpose.

(2.48) a. French: Je cherche un livre qui contient/contienne ce renseignement.

b. English: I am looking for such a book as might contain this information.

In Example 2.47, the nuances contained in the French sentence can be expressed by lexical means in English, whereas this distinction cannot be made in the English version of Example 2.48.

Verbal aspects. Aspects are distributed and realized differently in English and French. The progressive aspect, for instance, is characteristic of English, while the durative or imperfective aspect can easily be expressed in French, as can be seen from Examples 2.49 to 2.51.

(2.49) a. English: He is working.

b. French: Il est en train de travailler.

(2.50) a. French: Il se taisait.

b. English: He remained silent.

(2.51) a. French: Il s'est tu./Il se tut.

b. English: He fell silent.

Example 2.49 shows the typical use of the progressive aspect in English. In order to compensate the lack of this kind of aspect, French uses lexical means, such as "est en train de". To express the durative aspect, French uses *imparfait* for the description of states or continuing actions (see Example 2.50) as opposed to *passé composé* for the occurrence of an event (see Example 2.51). This distinction has to be realized lexically in English.

Phraseology. Both English and French distinguish between fixed compounds and NPs which are freely combined (see Examples 2.52 and 2.53).

(2.52) a. English: the trunk of a tree
b. French: le tronc d'un arbre

(2.53) a. English: a tree trunk
b. French: le tronc d'arbre

Example 2.52 illustrates the use of freely combined NPs in English and French. Example 2.53 shows how English forms compounds by simply combining two or more nouns, whereas French uses prepositions to form a compound.

Ellipsis. English has freer rules concerning ellipsis than French, which uses pronouns in order to avoid ellipses (see Examples 2.54 and 2.55).

(2.54) a. English: You did not tell me.
b. French: Vous ne m'en aviez pas parlé.

(2.55) a. English: He is satisfied, but I am not.
b. French: Il est satisfait, mais je ne le suis pas.

In Example 2.54, the French pronoun "en" is used to translate the English elliptical expression. Example 2.55 shows that, in both languages, the same solutions are used respectively (i.e. ellipsis in English and pronoun in French) to avoid the repetition of the attributive adjective within the comparison.

As in the summary of the English-German typology, overlaps and interactions also exist in connection with the English-French typological contrasts introduced above. For example, the supplementation with French nouns or the use of French nouns in characterization encourage the predominance of the French noun. An overview of the contrastive English-French typology is given in Table 2.9.

English	French
no dominant word class	predominance of the noun
natural order of flow of action	focus on result
no supplementation	frequent use of supplementation
focus on natural gender	focus on grammatical gender
frequent use of collective nouns	rare use of collective nouns
characterization through adjectives and adverbs	characterization through nouns
more expressive tense system	less expressive tense system
middle through finite verb	middle through reflexivation
auxiliary verbs and lexical means	auxiliary verbs and subjunctive
focus on progressive aspect	focus on duration aspect
compounds through noun combinations	compounds through prepositions
free ellipsis rules	restricted ellipsis rules

Table 2.9: Contrastive English-French typology

In summary, the English-German comparative typology and the English-French comparative stylistics introduced in this section show contrasts and commonalities between English and German as well as English and French, but they also illustrate contrasts and commonalities between German and French. For example, the English middle has to be translated into German and French reflexive constructions. However, it is rather problematic to compare the two typologies since they are not systemically motivated and the English-French typology focuses on stylistics. In order to compare the English-German to the English-French typology, a description of the different language systems is needed (such as provided in [Teich, 2001a] and [Teich, 2001b] for English-German).

2.3.2 Translation techniques

To give an overview of the wide range of techniques to transfer an SL text into a TL text, the notions of translation method and translation procedure as described in [Vinay & Darbelnet, 1995, 30 ff.] and [Newmark, 1988, 68 ff.] are introduced and compared in this section[9]. First of all, both make a distinction between the translation method (relating to the whole text) and the translation procedure (relating to sentences or smaller units). [Vinay & Darbelnet, 1995] differentiate between two translation methods and seven translation procedures, whereas [Newmark, 1988] discriminates between eight translation methods and sixteen translation procedures. A description thereof is given in the following, taking into account the differences and commonalities.

Translation methods. Both [Vinay & Darbelnet, 1995] and [Newmark, 1988] speak of SL-oriented translation as opposed to TL-oriented translation (cf. [Vinay & Darbelnet, 1995, 30 ff.] and [Newmark, 1988, 68 ff.]). According to this very rough classification, [Vinay & Darbelnet, 1995] call these methods direct/literal translation and oblique translation. The method of direct or literal translation can be applied when structural and/or conceptual parallelisms exist between the SL and the TL, whereas oblique translation is intended to bridge gaps between the SL and the TL. Newmark's classification is more delicate as SL-oriented translations are split up into word-for-word translation, literal translation, faithful translation and semantic translation. As regards the TL-oriented ones, a distinction is made between communicative translation, idiomatic translation, free translation and adaptation, depending on how literally or freely they are translated (cf. [Newmark, 1988]). The word-for-word translation is frequently applied to produce an interlinear translation of a text, reflecting the word order of the SL and its grammatical constructions. Literal translation goes a step further than word-for-word translation as it transfers the word order and the grammatical constructions of the SL into their most obvious counterparts in the TL. In a faithful translation, the contextual meanings of the

[9]There are of course other authors dealing with this topic (cf., for example, [Königs, 2000] or [Fawcett, 1997]), but the combination of [Vinay & Darbelnet, 1995] and [Newmark, 1988] gives a very comprehensive view on translation methods and procedures. Furthermore, the notion of equivalence is not discussed at all, since it is of little importance for this study.

words are taken into account and transferred, yet it remains strictly in line with the intentions of the writer of the SL text. Semantic translation, in which the stylistic and cultural specificities of the SL text are taken into account and neutrally rendered into the TL, represent the last SL-oriented method. Communicative translation belongs to the TL-oriented methods and transfers SL content and language into the TL so that it can easily be understood by the readership. To obtain an idiomatic translation, the translator tries to transfer the message of the SL text according to the norms of the TL, accepting that semantic details may get lost in the process. The purpose of a free translation is to give an intralingual translation, i.e. the content is rendered into the TL without taking into account the form of the SL text. Adaptation is frequently used to translate plays and poetry. With this method, the relevant cultural aspects of the SL text are transferred into the TL culture and the text is then rewritten in the TL.
In essence, it can be said that Newmark's word-for-word translation, literal translation, faithful translation and semantic translation are comparable to Vinay and Darbelnet's direct translation, whereas Newmark's communicative translation, idiomatic translation, free translation and adaptation can be compared to Vinay and Darbelnet's oblique translation.

Translation procedures. As regards the classification of the translation procedures, there are five procedures which [Vinay & Darbelnet, 1995] and [Newmark, 1988] have in common, two procedures which are only described in [Vinay & Darbelnet, 1995, 30 ff.] and eleven procedures which are only dealt with in [Newmark, 1988, 68 ff.]. First, a description will be given of the five procedures which [Vinay & Darbelnet, 1995] and [Newmark, 1988] have in common, with Vinay and Darbelnet's terminology followed by Newmark's terminology.[10]
Borrowing/transference refers to the process of introducing an SL word (usually an unknown technical or cultural concept) to the TL (see Example 2.56).

(2.56) a. English: à la mode

b. French: science-fiction

[10]Most of the examples are taken from [Vinay & Darbelnet, 1995, 30 ff.]. References to other sources can be found in the text.

Example 2.56 shows that the French expression “à la mode” is introduced to the English language and the English expression “science-fiction” to the French language. This translation procedure is also termed emprunt, loan word or transcription.
Calque/naturalization represents a special kind of borrowing/transference, where an SL expression is introduced to the TL but adapted to the TL pronunciation and/or morphology, or each of its elements is translated literally into the TL (see Example 2.57).

(2.57) a. English: Governor General

b. French: Compliments de la Saison

In Example 2.57, the English expression “Governor General” is a calque/naturalization of the French source “Gouverneur Général”, whereas the French calque/naturalization “Compliments de la Saison” originates from the English expression “Compliments of the Season”. Both expressions are borrowed from the SL, but follow the norms of the TL morphology and pronunciation.
With literal translations[11], the SL text is translated word for word into the TL, taking into consideration the grammar and the idiomatic expressions of the TL (see Example 2.58).

(2.58) a. English: The book is on the table.

b. French: Le livre est sur la table.

Example 2.58 shows that, in this case, a word-for-word translation from English into French is possible which obeys the grammatical rules of the TL. In some cases, it is not possible to translate an SL sentence literally without losing its meaning or neglecting the grammar of the TL (see Example 2.59(a) and (c)); in such cases, one of the other translation procedures has to be used. If a translator tries to render such an English sentence into a meaningful translation, an idiomatic expression has to be used as in Example 2.59(b).

[11] [Vinay & Darbelnet, 1995] and [Newmark, 1988] use the same terminology for this procedure.

(2.59) a. English: He looked the picture of health.

b. French: Il avait l'air en pleine forme.

c. French: *Il paraissait l'image de la santé.

Literal translation occurs very frequently where translating involves language pairs from the same language family (such as French and Italian). Transposition[12] or shift means that rendering the SL text into the TL entails grammatical changes such as a change from singular to plural (see Example 2.60 from [Newmark, 1988, 85]) or a change to compensate a lexical gap (see Example 2.61 from [Newmark, 1988, 85]) or a lacking grammatical construction (see Example 2.62 from [Newmark, 1988, 85]).

(2.60) a. English: advice

b. French: des conseils

(2.61) a. English: it totals

b. French: il atteint le total

(2.62) a. German: Das Interessante ist, daß ...

b. English: What is interesting is that ...

c. or: The interesting thing is that ...

d. or: It's interesting that ...

e. or: The interest of the matter is that ...

The German construction of a nominalized adjective as subject does not exist in English (see Example 2.62). For this reason, another grammatical option has to be chosen for the translation of this construction into English. The replacement of one word class with another is another type of transposition (see Example 2.63).

(2.63) a. English: After he comes back ...

b. French: Après son retour ...

[12] Again, [Vinay & Darbelnet, 1995] and [Newmark, 1988] use the same terminology for this procedure.

In Example 2.63, a temporal subordinate clause is replaced by a PP serving as a temporal adjunct. This is an example of an optional transposition (as opposed to the obligatory ones discussed in Examples 2.60, 2.61 and 2.62), since the transposed expression could have been translated using the same word classes in the TL as well. This indicates that transposition can also occur within one language. Example 2.64 displays the different options of transposition available in English and French with regard to the sentence discussed in Example 2.63.

(2.64) a. English: After he comes back ...

b. English: After his return ...

c. French: Après qu'il sera revenu ...

d. French: Après son retour ...

Not only do transpositions include word class changes from verb to noun and vice versa, but also any other possible word class change.
Modulation[13] describes the translation of an SL expression into a TL expression which involves a change in terms of point of view or perspective. There are ten different types of modulation a translator can choose from: abstract for concrete or metonymy (see Example 2.65), explicative modulation or cause for effect (see Example 2.66), part for the whole or synecdoche (see Example 2.67), one part for another or metonymy (see Example 2.68), reversal of terms (see Example 2.69), negation of the opposite or litotes (see Example 2.70), active to passive and vice versa (see Example 2.71), space for time or metalepsis (see Example 2.72), exchange of intervals for limits in space and time (see Example 2.73) and change of symbol (see Example 2.74).

(2.65) a. English: Buy Coca-Cola by the carton.

b. French: Achetez Coca-Cola en gros.

(2.66) a. English: You're quite a stranger.

b. French: On ne vous voit plus.

[13]Once again, [Vinay & Darbelnet, 1995] and [Newmark, 1988] use the same terminology for this procedure.

(2.67) a. English: He shut the door in my face.
b. French: Il me claqua la porte au nez.

(2.68) a. English: He read the book from cover to cover.
b. French: Il lut le livre de la première à la dernière page.

(2.69) a. English: You can have it.
b. French: Je vous le laisse.

(2.70) a. English: It does not seem unlikely.
b. French: Il est fort probable.

(2.71) a. English: Only a miracle saved the world.
b. French: Le monde n'a été sauvé que par miracle.

(2.72) a. English: This in itself presented a difficulty.
b. French: Cette opération présentait déjà une difficulté.

(2.73) a. English (in time): For the period under review ...
b. French (in time): Depuis notre dernier numéro ...
c. English (in space): No parking between signs.
d. French (in space): Limite de stationnement.

(2.74) a. English: as like as two peas
b. French: comme deux gouttes d'eau

Again, like transpositions, modulations may be optional or obligatory.
In the following, two translation procedures are presented which are only described in [Vinay & Darbelnet, 1995, 30 ff.] (i.e. equivalence and adaptation) and eleven procedures are introduced which are only dealt with in [Newmark, 1988, 68 ff.] (cultural equivalent, functional equivalent, descriptive equivalent, synonymy, through-translation, recognized translation, translation label, componential analysis, reduction and expansion, compensation

as well as paraphrase).[14]
Equivalence refers to a translation procedure which uses the stylistic and grammatical means offered by the TL to produce the same situation in the TL as described in the SL text (see Example 2.75).

(2.75) a. English: Like a bull in a china shop.

b. French: Comme un chien dans un jeu de quilles.

Example 2.75 illustrates this translation procedure which is frequently used to translate proverbs. It is also used for the translation of idioms, clichés and fixed nominal or adjectival phrases.
With adaptation, an equivalent situation is created in order to translate an SL message unknown to the TL recipient (see Example 2.76).

(2.76) a. English: before you could say Jack Robinson

b. French: en un clin d'oeil

This translation procedure is similar to Newmark's cultural equivalent (cf. [Newmark, 1988, 68 ff.]). See Example 2.77 for a comparison.
A cultural equivalent is used to translate an expression typical of the SL into an expression typical of the TL, as shown in Example 2.77.

(2.77) a. English: tea break

b. French: café-pause

The difference between adaptation and cultural equivalent is that translators choose an adaptation to translate a cultural concept unknown in the TL by creating a similar TL situation as opposed to a cultural equivalent, which is intended to transfer an SL cultural concept to a similar TL concept.
A functional equivalent is chosen to transfer an SL cultural concept by using a neutral or generalized expression, as shown in Example 2.78.

(2.78) a. English: Roget

b. French: dictionnaire idéologique anglais

[14]Examples are taken from [Vinay & Darbelnet, 1995, 30 ff.] and [Newmark, 1988, 68 ff.] respectively.

This translation procedure is used in cases where the cultural concept of the SL cannot be transferred into a similar TL context. It implies the analysis of the SL expression and its deculturalization in the TL.
A descriptive equivalent provides a detailed description of the SL expression (see Example 2.79).

(2.79) a. Spanish: machete

b. English: a Latin American broad, heavy instrument whose function is cutting or aggression

As can be seen in Example 2.79, the TL text gives a description for the SL expression. Again, like the functional equivalent, this procedure is used if the cultural concept of the SL is unknown to the TL recipient and cannot be transferred into a similar TL context. The additional information of a functional or a descriptive equivalent can occur within the text or in footnotes.
Synonymy means that the TL offers an equivalent close in meaning for an expression which cannot be translated literally. This translation procedure implies a slight shift in meaning, as shown in Example 2.80.

(2.80) a. English: kind person

b. French: personne gentille

Synonymy is frequently used in connection with adjectives and adverbs of quality for which it is difficult to find a TL equivalent covering all the semantic nuances of the SL word.
Through-translation is used to translate names of organizations and acronyms, as can be seen in Example 2.81.

(2.81) a. English: NATO

b. French: OTAN

It also serves the literal translation of compounds, collocations and phrases and is thus similar to Vinay and Darbelnet's calque (cf. [Vinay & Darbelnet, 1995, 30 ff.]).
Recognized translation means that the translator chooses an officially accepted TL equivalent of an institutional term of the SL (see Example 2.82).

(2.82) a. German: Rechtsstaat

b. English: constitutional state

If translators do not agree with the officially recognized equivalent, they usually gloss it.
Translation label refers to the literal translation of an institutional term, usually put between inverted commas, as shown in Example 2.83.

(2.83) a. German: Erbschaftssprache

b. English: 'heritage language'

The translation label often serves as a provisional translation of a new institutional term.
Componential analysis refers to the analysis of the different components included in one lexical unit, as can be seen in Example 2.84.

(2.84) a. German: stürzen

b. English: to fall + suddenly + heavily + may refer to an important person or entity

This translation procedure often results in one-to-two or one-to-many translations.
Reduction and expansion are translation procedures which are used rather intuitively (see Examples 2.85 and 2.86).

(2.85) a. French: science linguistique

b. English: linguistics

(2.86) a. French: cheveux égaux

b. English: evenly cut hair

As shown in Example 2.85, reduction is frequently applied to NPs in which the SL adjective can be dropped or rendered into a TL noun. Expansion frequently occurs where an SL adjective can be transferred into a TL adverb and a participle is added to the TL expression (see Example 2.86).
The last two translation procedures, compensation and paraphrase, are

rather rare and are applied to translate larger text segments. For this reason, they are explained in the following, without using examples for illustration. Compensation means that a loss in meaning in one part of a translated text segment is compensated in another part of the translated text segment. Paraphrase is used to explain or amplify a meaning lacking in a text segment.

It is important to mention that a translation procedure can be used more than once or even in combination with other translation procedures in the same translation unit, as shown in Examples 2.87 and 2.88 (taken from [Vinay & Darbelnet, 1995, 252 ff.])[15].

(2.87) a. English: Let sleeping dogs lie.

b. French: Il ne faut pas réveiller le chat qui dort.

(2.88) a. English: Men will not always die quietly.

b. French: Les hommes ne mourront pas toujours sans se plaindre.

In Example 2.87, three different types of modulation can be found in the translated expression (modulation by negation of the opposite: "il ne faut pas" for "let", modulation by change of symbol: "chat" for "dog" and modulation by reversal of terms: "réveiller" for "lie"). Example 2.88 combines a transposition (word class change from adverb to verb: "sans se plaindre" for "quietly") and a modulation (modulation by negation of the opposite: "pas toujours sans se plaindre" for "not always quietly").

Another important point made by [Vinay & Darbelnet, 1995] is that the different translation procedures can be employed "at the three planes of expression, i.e. lexis, syntactic structure and message" (see Examples 2.89, 2.90 and 2.91).

(2.89) a. English: from

b. French: expéditeur

[15] [Newmark, 1988] calls this phenomenon couplet, triplet or quadruplet depending on the number of translation procedures which add up or combine.

(2.90) a. English: As timber becomes more valuable ...

b. French: Depuis la revalorisation du bois ...

(2.91) a. English: No smoking!

b. French: Défense de fumer!

Example 2.89 illustrates a transposition at the lexical level, Example 2.90 shows a transposition at the syntactical level, and in Example 2.91, a transposition at the semantic level can be found.

Comparing direct translation methods (i.e. methods with SL emphasis) to oblique translation methods (i.e. methods with TL emphasis) on the basis of the translation procedures used for each method, the following classification can be made: Borrowing, calque, literal translation, transference, naturalization, synonymy, through-translation, recognized translation and translation label are preferably used for direct translations, whereas transposition, modulation, equivalence, adaptation, cultural equivalent, functional equivalent, descriptive equivalent, componentional analysis, reduction and expansion, compensation as well as paraphrase are preferably used for oblique translations.

2.3.3 SFL and grammatical metaphor

Within the Systemic Functional (SF) model (cf. [Halliday, 1961], [Halliday, 1963], [Halliday, 1967], [Halliday, 1968], [Halliday, 1973], [Halliday, 1976], [Halliday, 1978], [Halliday, 1985] and [Matthiessen, 1995]), language is looked upon as a social semiotic. SFL takes the cultural and situational context into account. This view of language is based on Firth's model of language (cf. [Firth, 1957]) and is motivated by anthropological principles (cf. [Malinowski, 1935]). It supports Whorf's theory (cf. [Whorf, 1956]) according to which thought is determined by language (cf. [Halliday & Matthiessen, 1999]). In the following, the main characteristics of the SF model are briefly introduced.

- **Stratification:** Within the SF model, language is tristratal: the three strata are semantics, lexico-grammar and phonology. In this context,

semantics constitutes a resource for meaning, whereas lexico-grammar represents a resource for wording, i.e. a resource of grammatical and lexical items for realizing semantics. The relation between semantics and grammar is influenced by the creation of new modes of meaning termed grammatical metaphor. Phonology is a resource for sounding, which includes intonation, rhythm, syllabic and phonemic articulation, however, merely the intonational resources serve the realization of the grammatical choices. Phonology belongs to the expression plane, whereas lexico-grammar and semantics constitute the content plane. Moreover, these three strata of the linguistic system are all embedded in a higher-level system of context.

- **Metafunctions:** The content plane of language (i.e. lexico-grammar and semantics) comprises the ideational, interpersonal and textual metafunction. Firstly, the ideational metafunction refers to the propositional content. It is further divided into sub-types: the experiential and the logical metafunction. The experiential metafunction represents classes of phenomena as a whole, whereas the logical metafunction represents combinations of phenomena through highly generalized relations. Secondly, the interpersonal metafunction is concerned with discourse participants and their role relations. Thirdly, the textual metafunction represents the linguistic patterns which create a cohesive and coherent text. While the metafunctions are principles of organizing the content plane, the context of situation contains three categories of higher-level semiotic organization: field, tenor and mode. The field of discourse deals with the nature of the social action that is taking place, whereas tenor describes the nature of the participants, their statuses and roles, and mode refers to the role language plays in that situation. Field, tenor and mode together make up the context of situation within the register of a given text.

- **Axiality:** In the context of the SF model, the linguistic description of each stratum takes into account two aspects, namely the paradigmatic and the syntagmatic one. The paradigmatic aspect represents linguistic systems, whereas the syntagmatic aspect deals with the realizations of these systems. This concept is referred to as axiality. The paradigmatic axis is represented by means of system networks, whereas the

specifications of the syntagmatic axis are derived by means of realization statements which are associated with paradigmatic choices.

- **Delicacy:** The systems of a system network are ordered according to their delicacy (from least delicate to most delicate). In the SF model, the least delicate system is the most general one, while the most delicate system is the most specific one. All strata and all metafunctions are based on system networks.

- **Ranks:** The organization of structure within lexico-grammar and phonology is rank-based. The hierarchic rank scale is sorted from the largest unit to the smallest unit. Units of a higher rank (e.g., group/phrase in grammar or syllable in phonology) consist of units of the rank below (e.g., words and phonemes, respectively). In this case, rankshift entails downranking.

These five SF dimensions provide a profound fundament for the linguistic description of language. Figure 2.6 (taken from [Teich, 2001a, 34]) illustrates the relationship between these dimensions within the SF model.

Figure 2.6: Dimensions of SF description

For a SF analysis of texts, it is important to emphasize that a text is the instantiation of a linguistic system in a particular cultural or situational context (cf. [Teich, 2001a]). This makes it possible to look at a register from two different perspectives: the system-based and the text-based one. Furthermore, the creation of texture is another important aspect of the SF model (cf. [Halliday, 1985] and [Halliday & Hasan, 1976]). Texture is

split up into a structural level (focusing on thematic structure, information structure and focus) and into a cohesive level (which comprises reference, ellipsis, substitution, conjunction and lexical cohesion).

In the recent past, several attempts have been made to adapt the SF model to other languages than English or to multilingual contexts (cf., for example, [Caffarel, 1996], [Teich, 1995a], [Matthiessen, 2001], [Halliday, 2001], [Steiner, 2001a] and [Steiner, 2001b]). The analysis and comparison of any two languages will always reveal similarities and differences. Thus, a model of language comparison has to encompass both aspects – cross-linguistic commonality and divergence. Since SFL covers both aspects, in the following, the above-introduced SF dimensions are applied to contrastive language description (cf. [Teich, 1999], [Teich, 2001a] and [Teich, 2001b]).

Stratification. Languages tend to bear more similarity on the semantic plane, whereas they show more differences on the grammatical plane. This means that the same proposition can be expressed by different lexico-grammatical means in different languages (see Example 2.92)[16].

(2.92) a. English: The guitar broke a string.

b. German: An der Gitarre riß eine Saite.

c. German: *Die Gitarre zerriß eine Saite.

In Example 2.92[17], the same ideational meaning is expressed in English and German. In English, a transitive construction is used (with "The guitar" as non-agentive NP subject), whereas in German, a middle construction is chosen (with "An der Gitarre" as a locational circumstantial). This choice is due to typological differences between English and German (cf. [Hawkins, 1986, 58 ff.]). The literal translation of the English sentence (see Example 2.92(c)) with the non-agentive NP subject "Die Gitarre" would lead to a personification which is not acceptable in German. Thus, it is not possible to realize the same ideational meaning with the same lexico-grammatical means in English and German for expressions like the one in Example 2.92.

[16]The examples illustrating the SF dimensions of contrastive linguistic description are taken from [Teich, 2001a, 45 ff.].

[17]This example is taken from [Hawkins, 1986, 58].

Metafunctions. In different languages, a particular grammatical realization may serve different metafunctions, e.g., word order serves the interpersonal or experiential metafunction in English, English being a grammatical word order language. However, in Russian, which is a pragmatic word order language, it serves the textual metafunction (see Example 2.93).

(2.93) a. English: A girl went to Moscow.

b. English: The girl went to Moscow.

c. Russian: Poekhala devushka v Moskvu.

d. Russian: Devushka poekhala v Moskvu.

In Example 2.93, the English word order marks syntactic functions, i.e. the subject of the sentence ("A girl" and "The girl"). Thus, it serves the interpersonal metafunction. In the Russian sentences, word order serves the textual metafunction, marking theme/rheme and given/new distribution (with the verb "Poekhala" in initial position as a marker for identifiability and the noun "Devushka" in initial position as a marker for non-identifiability). In contrast to Russian, this feature is realized through the use of indefinite ("A") and definite ("The") articles in the English sentences.

Axiality. Languages tend to be similar concerning their paradigmatic systems, while differing in terms of syntagmatic realizations. The English mood system is a good example of this: English, French and German distinguish between declarative, interrogative and imperative, yet their respective realizations of the imperative diverge (see Example 2.94).

(2.94) a. English: Leave me alone!

b. French: Laisse-moi tranquille!

c. German: Laß mich in Ruhe!

The English imperative is realized through a non-finite verb form (here: "Leave"), whereas the French and German imperatives are realized through finite verb forms (here: "Laisse" and "Laß").

Delicacy. Grammatical systems of low delicacy (i.e. less specific grammatical types) show the tendency to be more similar across languages than

grammatical systems of high delicacy (i.e. more specific grammatical types). Again, the system of mood is a good example (see Example 2.95).

(2.95) a. English: Leave me alone!

b. French: Laisse-moi tranquille!

c. French: Laissez-moi tranquille!

d. German: Laß mich in Ruhe!

e. German: Laßt mich in Ruhe!

f. German: Lassen Sie mich in Ruhe!

Example 2.95 shows that, in English, there is only one way of realizing imperatives ("Leave"), whereas several forms are available in French and German according to the number of addressees and to the degree of social distance between the participants (the French "Laisse" indicating a low degree of social distance and singular, the French "Laissez" indicating a low or high degree of social distance for plural or a low degree of social distance for singular, the German "Laß" indicating a low degree of social distance for singular, the German "Laßt" indicating a low degree of social distance for plural, and, finally, the German "Lassen" indicating a high degree of social distance for singular or plural).

Ranks. Different languages may express the same contents through different ranks, as can be seen from Example 2.96, which gives an instance of the marking of referents in terms of identifiability.

(2.96) a. English: Specify the first point of the multiline.

b. Bulgarian: Zadaite purvata tochka na multilinijata.

Example 2.96 shows that English realizes the marking of referents at nominal group rank through the definite article "the", whereas in Bulgarian it is expressed at morpheme rank through the affix "ta".

Instantiation. Different languages may diverge in terms of instantiation despite commonalities in the language systems (as can be seen from Example 2.97).

(2.97) a. English: Set the switch OFF when carrying the radio to prevent turning the power on accidentally.

b. French: Réglez-la sur OFF quand vous transportez la radio pour éviter de la mettre sous tension accidentellement.

c. German: Beim Transportieren sollte der Schalter auf OFF gestellt werden, damit das Gerät nicht versehentlich eingeschaltet wird.

In Example 2.97, the speech act *command* is expressed through different lexico-grammatical constructions: through an imperative in the English example ("Set"), through a polite imperative in the French example ("Réglez") and by means of a declarative in the German example ("sollte ... gestellt werden"). Given the systemic differences described above, the divergence between English and French is predictable. Nevertheless, the choice of the German declarative is not a consequence of systemic differences (since the same options as in English and French are available in German, too). It is rather motivated through register-specific language use.

Figure 2.7: Dimensions of SF cross-linguistic variation in system and text

All these SF categories make up a framework for cross-linguistic analysis and comparison as well as for the description of possible differences and similari-

ties in system and text. Figure 2.7 (taken from [Teich, 2001a, 54]) illustrates all the SF dimensions of contrastive linguistic description explained above.

Figure 2.8: Grammatical metaphor

The notion of grammatical metaphor is another important aspect of the SF model (cf. [Halliday, 1985] and [Halliday & Matthiessen, 1999]). Grammatical metaphor can be described as the encoding of the same ideational meaning by means of different phrasal categories, such as clause complex, clause, phrase, group, word or morpheme (cf. [Steiner, 2001b, 7 ff.]), the wording that is higher in rank being more metaphorical than the congruent variants lower in rank. Similar concepts can be found in Tesnière's translation (cf. [Tesnière, 1959]), in Melchuk's part-of-speech lexical functions (cf. [Melchuk, 1988]), in the notion of paraphrase (used in many semantic theories), in the distinction Polenz makes between explicit and compressed expression (cf. [Polenz, 1988]), in the mapping of entities onto predicates (cf. [Ehrich & Rapp, 2000]) and in notions of derivation and category-changing transformation (such as nominalizations; cf. [Chomsky, 1970]). Since [Halliday & Matthiessen, 1999] explain that "the wording that is lower in rank will contain less information" (cf. [Halliday & Matthiessen, 1999, 231]), it can be argued that grammatical metaphor is related to the concept of explicitation (cf. [Blum-Kulka et al., 1989] and [Baker, 1996]). However, the relationship between grammatical metaphor and explicitation or information density is rather a complex one and is not discussed in this thesis. Figure 2.8 illustrates the notion of grammatical metaphor in greater detail (taken from [Halliday & Matthiessen, 1999, 231]). It shows that the sentence "Lung cancer death rates are clearly associated with increased smoking" is less explicit and thus more metaphorical than possible

corresponding combinations of [1] to [3] (see Example 2.98)[18].

(2.98) a. Lung cancer death rates are clearly associated with increased smoking.

b. (It is clear that) if more people smoke, then more people die of lung cancer.

c. (It is clear that) some people smoke more, so they die faster of lung cancer.

In Example 2.98, (a) consists of a relational process and rather complex phrases, whereas in (b) and (c) the same ideational meaning is expressed in a clause complex. Concerning grammatical shift, there are verbs and conjunctions in (b) and (c) which become adjectives and nouns in (a). Thus, (b) and (c) are congruent (and more explicit) versions of the metaphorical (and more ambiguous) version in (a) (cf. [Steiner, 2001b, 8]).

Other types of grammatical shift which can take place in grammatical metaphors are listed in Table 2.10, which is an abridged version of Halliday and Matthiessen's classification (cf. [Halliday & Matthiessen, 1999, 246 ff.]). At this point it is important to mention that grammatical metaphor cannot be reduced to a mere change in word classes (transcategorization), but it implies a semantic tension between the congruent and the metaphorical realization of the same ideational meaning (cf. [Steiner, 2001b] and see Example 2.98 and the examples given in Figure 2.8 and in Table 2.10).

[18]This example (cf. [Steiner, 2001b, 8]) is an adapted version of Halliday and Matthiessen's example (cf. [Halliday & Matthiessen, 1999, 231]).

Grammatical shift	**Example**
(congruent ⇒ metaphorical; explicit ⇒ implicit)	
adjective ⇒ noun	unstable ⇒ instability
verb ⇒ noun	transform ⇒ transformation
auxiliary ⇒ noun	could ⇒ possibility
semi-auxiliary ⇒ noun	want to ⇒ desire
preposition ⇒ noun	with ⇒ accompaniment
PP ⇒ noun	dust is on the surface ⇒ surface dust
conjunction ⇒ noun	if ⇒ condition
verb ⇒ adjective	poverty increases ⇒ increasing poverty
auxiliary ⇒ adjective	was ⇒ previous
semi-auxiliary ⇒ adjective	begin to ⇒ initial
preposition ⇒ adjective	with ⇒ accompanying
PP ⇒ adjective	marks are on the surface ⇒ superficial marks
conjunction ⇒ adjective	before ⇒ previous
preposition ⇒ verb	instead of ⇒ replace
PP ⇒ verb	put in a box ⇒ box
conjunction ⇒ verb	then ⇒ follow
conjunction ⇒ preposition	because ⇒ because of
conjunction ⇒ PP	so ⇒ as a result
0 ⇒ noun	[x] ⇒ the fact of [x]
0 ⇒ verb	[x] ⇒ have [x]
0 ⇒ causative	make [x:y] ⇒ impose [y on x]
0 ⇒ phase auxiliary	started to survey ⇒ started a survey
NP head ⇒ post-mod.	the government decided ⇒ decision by the government
NP head ⇒ deictic	the government decided ⇒ the government's decision
NP head ⇒ classifier	the government decided ⇒ governmental decision
adverb ⇒ adjective	decided hastily ⇒ hasty decision
PP ⇒ adjective	argued for a long time ⇒ a lengthy argument
adverb ⇒ various	announced yesterday ⇒ yesterday's announcement
PP ⇒ various	departed for the airport ⇒ departure for the airport

Table 2.10: Types of grammatical metaphor (monolingual)

Recently, the notion of grammatical metaphor has been used to describe the specific properties of translations (cf. [Steiner, 2001b]), in the sense that the concept of grammatical metaphor can be adapted to multilingual contexts (see Example 2.99 taken from [Steiner, 2001b, 8]).

(2.99) a. Lungenkrebssterblichkeitsraten sind klar assoziiert mit verstärktem Rauchen.

b. (Es ist klar, daß), wenn mehr Menschen rauchen, eine größere Zahl von ihnen an Lungenkrebs stirbt.

c. (Es ist klar, daß) einige Menschen mehr rauchen und sie deshalb schneller an Lungenkrebs sterben.

d. Die Sterblichkeitsrate bei/an Lungenkrebs steht in deutlicher Beziehung zum Anstieg des Rauchens.

e. Die Sterblichkeitsrate bei/an Lungenkrebs ist klar korreliert mit erhöhtem Tabakkonsum.

f. Bei verstärktem Tabakkonsum erhöht sich nachgewiesenermaßen die Sterblichkeitsrate bei/an Lungenkrebs.

g. Es gibt eine deutliche Beziehung zwischen erhöhtem Tabakkonsum und steigenden Sterblichkeitsraten bei/an Lungenkrebs.

Example 2.99 lists German versions of the English examples given in Example 2.98 and in Figure 2.8, (a) being the literal translation and (b) to (g) are more congruent (or de-metaphorized) versions of (a) as well as of the English Example 2.98. Table 2.11 contains a list of grammatical shifts which can take place in grammatical metaphor across languages (taken from [Steiner, 2001b, 11][19]). At this point it is important to mention that grammatical metaphor across languages is not the same as the translation procedure termed *transposition* (cf. [Vinay & Darbelnet, 1995, 30 ff.] and [Newmark, 1988, 68 ff.]), since it implies a semantic tension between congruent expression in a language A and its metaphorical realization in a language B (cf. [Steiner, 2001b] and see Example 2.99 and the examples given in Table 2.11).

[19]This multilingual classification is based on Halliday and Matthiessen's monolingual classification (cf. [Halliday & Matthiessen, 1999, 246 ff.]) introduced in Table 2.10

Grammatical shift	**Example**
(metaphorical ⇒ congruent; implicit ⇒ explicit; English ⇒ German)	
noun ⇒ adjective	instability ⇒ instabil
noun ⇒ verb	transformation ⇒ transformieren
noun ⇒ auxiliary	the possibility of ⇒ können
noun ⇒ semi-auxiliary	the desire to ⇒ wollen
noun ⇒ preposition	accompaniment ⇒ mit
noun ⇒ PP	floor dust ⇒ Staub auf dem Boden
noun ⇒ conjunction	condition ⇒ wenn
adjective ⇒ verb	rising poverty ⇒ Armut steigt
adjective ⇒ auxiliary	the previous, past ... ⇒ war
adjective ⇒ semi-auxiliary	the initial ... ⇒ beginnen
adjective ⇒ preposition	the accompanying ⇒ mit
adjective ⇒ PP	superficial ... ⇒ auf der Oberfläche
adjective ⇒ conjunction	previous ⇒ vor
verb ⇒ preposition	replace ⇒ anstatt
verb ⇒ PP	to box ⇒ in eine Schachtel
verb ⇒ conjunction	follow ⇒ dann
preposition ⇒ conjunction	because of ⇒ weil
PP ⇒ conjunction	as a result ⇒ deshalb
noun ⇒ 0	the fact that ... ⇒ 0
verb ⇒ 0	have an influence ⇒ beeinflussen
causative ⇒ 0	X imposes work on Y ⇒ X läßt Y arbeiten
phase auxiliary ⇒ 0	to begin an examination ⇒ beginnen zu untersuchen
post-mod. ⇒ NP head	a decision by the government ⇒ Die Regierung entscheidet
deictic ⇒ NP head	the government's decision ⇒ Die Regierung entscheidet
classifier ⇒ NP head	governmental decision ⇒ Die Regierung entscheidet
adjective ⇒ adverb	a hasty decision ⇒ entscheidet hastig
adjective ⇒ PP	a lengthy argument ⇒ stritten für eine lange Zeit
various ⇒ adverb	yesterday's quarrel ⇒ stritten gestern
various ⇒ PP	departure for the station ⇒ fahren zum Bahnhof

Table 2.11: Types of grammatical metaphor (multilingual)

2.4 Experimental psycholinguistics

This section presents the state of the art of the psycholinguistic research relevant to this thesis. In experimental psycholinguistics, distinctions are made with regard to the research methods, as can be seen from Table 2.12[20].

spoken	written
off-line methods (non-reaction time methods)	on-line methods (reaction time methods)
comprehension	production

Table 2.12: Types of methods in experimental psycholinguistics

Table 2.12 shows that the methods in experimental psycholinguistics are distinguished according to the mode of language, i.e. spoken vs. written. Furthermore, a distinction can be made between on-line and off-line methods. For the first, the reaction time of each subject is recorded, for the second, time does not play a role. Finally, the methods can be classified according to the research type they are applied in, i.e. they can either be applied in research on language comprehension or language production. In the following, firstly, the methods in experimental psycholinguistics and their applications are introduced and, secondly, they are classified according to the types displayed in Table 2.12 (see Table 2.13).

Eye-tracking method. In eye-tracking, a text is presented on a computer screen. The eye movements of the subject looking at it are recorded with the help of light reflections of the cornea. Vertical and horizontal eye movements are measured up to thousand times per second. Thus, it can be evaluated where the subjects direct their eyes and for how long the subjects look at a particular point. Thus, saccades (usually from left to right), fixiations and regressions (to previous locations) are revealed. There are several variants of the eye-tracking method, e.g., the text can be changed according to the position of the eye (cf. [Haberlandt, 1994]).
There are various applications for eye-tracking methods. [Inhoff, 1987], for instance, analyzed the relationship between word structure and eye movement and found that the processing of the text is easier for the subject if the

[20]In this section, experimental methods in psycholinguistic research are presented and some major applications are introduced. Hypotheses concerning the respective method are not introduced since only the methodological issues are relevant to this thesis.

first three letters of the next word are displayed. [Just & Carpenter, 1980] investigated the effect of word length and word familiarity on gaze durations. They found, as independent effects, that short and familiar words are easier to process. Additionally, they analyzed syntactic processing and recorded longer gaze durations at final words in clauses, which reveals clause and sentence boundaries and ambiguous structures. [Schustack et al., 1987] and [Zola, 1984] found that words are easier to process if their concepts can be predicted on the basis of the context. [O'Brien et al., 1988], [Rayner et al., 1989] and [Garrod et al., 1990] concluded from gaze durations that the identification of a target concept is easier if it can be inferred from the context.

Window method. With this method, the text is presented on a monitor and the subject presses a key to look at successive windows (i.e. segments) of the text. The reading times for each window are measured when the key is pressed. The windows can be paragraphs, sentences, phrases, words or the whole text. There are several window methods which can be applied. The most popular one is the moving window method. This method uses dashes to mask each word and only one word at a time is revealed when the key is pressed. The words seem to run through the screen since at each key press the previous word is masked and the next in line revealed. In contrast to this method, the stationary window method uses the same location for each word which is revealed to the subject. In the cumulative window method, the revealed words remain unmasked on the screen, building up a whole sentence. With the pointing method, the words are partially masked, but the subject can recognize the length and shape of the words. The subject's pointing to a word with a special device like a mouse causes the words to be revealed (cf. [Haberlandt, 1994]).
Window methods can be used to investigate frequency and length effects (cf. [Haberlandt & Graesser, 1985], [Just et al., 1982] and [Mitchell, 1984]) and spillover effects (cf. [Haberlandt & Graesser, 1990] and [Mitchell, 1984]) at word level. Furthermore, [Aaronson & Ferres, 1984] and [Just et al., 1982] used window methods to investigate sentence wrap-up processes and found that the processing time for a word at the end of a clause and especially at the end of a sentence is longer than for clause-initial words. Another interesting observation relating to the sentence level was made by [Clifton

et al., 1991], who confirmed the minimal attachment principle: reading times for verb attachments are shorter than for noun attachments. At text level, [Bloom et al., 1990] found that word reading times are longer if the subject recognizes causal links (through antecedents) between clauses and sentences. Additionally, [Haberlandt, 1984] investigated expository passages, which are harder to process even at word level. The processing of new information takes longer than the processing of repeated or old information. [Haberlandt & Graesser, 1989] made this observation in connection with the introduction of new information in passages, whereas [Haberlandt & Graesser, 1990] and [Just et al., 1982] gained these results for clausal boundaries. [Sturt et al., 1999] used the self-paced reading method to analyze structural change and reanalysis difficulties in language comprehension.

Recognition Method. In connection with the item-recognition method, a subject reads a passage and, during or shortly after reading, one or more words are shown on a monitor. The subject then has to indicate whether or not the word appeared in the passage by pressing a key. Negative target words, which did not occur in the passage, are used to analyze inferences. The subject receives feedback on the accuracy and the speed of the results (cf. [Haberlandt, 1994]).
In [Ratcliff & McKoon, 1978] and [McKoon & Ratcliff, 1980], this method was used to inquire into the memory structure concerning propositional representations. Furthermore, [Glenberg et al., 1987] and [Morrow et al., 1990] analyzed the involvement of mental models in reading and recognizing propositionally similar text fragments. [Dell et al., 1983] used the recognition method in order to show that antecedent information is still active and easier to process than new information (see also [Gernsbacher, 1989], [O'Brien et al., 1986] and [Speelman & Kirsner, 1990]). Using item-recognition techniques, [McKoon & Ratcliff, 1986] and [McKoon & Ratcliff, 1990] analyzed inferences and the predictability of events.

Lexical decision task. The lexical decision task presents a word and a non-word to the subject who then decides whether or not it is a correct English word by pressing one of two keys (cf. [Haberlandt, 1994]). In the cross-modal lexical decision task, a subject listens to sentences which are semantically related to one of the meanings of visually presented polyse-

mous words (cf. [Ferreira & Anes, 1994]). With both variants, accuracy and reaction times are recorded.
[Gordon, 1983] found that decision times for familiar words are shorter. Moreover, it takes less time to process words that are embedded in a related context (cf. [Schuberth & Eimas, 1977], [West & Stanovich, 1982], [Sharkey & Mitchell, 1985] and [Potts et al., 1985]). [Swinney, 1979] used the cross-modal lexical decision task to show that the lexical decision between polysemous words is biased by the semantically associated context (see also [Gernsbacher & Faust, 1990], [Lucas, 1987], [Onifer & Swinney, 1981], [Tanenhaus et al., 1979] and [Till et al., 1988]).

Naming method. In connection with the word-naming technique, a text following a target word is presented visually to the subject. The subject then has to vocally name the target word. There are variants of this method where the subject has to make a lexical decision or to choose a color which encodes the target word (cf. [Haberlandt, 1994]). In the cross-modal naming task, a subject listens to sentences while the target word is presented visually. The subject then has to name the target word, to make a lexical decision or to choose a color (cf. [Ferreira & Anes, 1994]). All variants involve the recording of accuracy and reaction times.
[Baluch & Besner, 1991] and [West & Stanovich, 1982] used naming methods and found that word processing is faster for pairs of words (like "uncle" and "aunt"). The processing could be optimized when the duration of the prime was increased (cf. [Warren, 1977]). [Simpson & Krueger, 1991] as well as [Forster, 1981] studied the effect of the context of the sentence or passage on word processing. [Ferreira, 1991] found that the complexity of sentences has a negative influence on the time required for processing. At text level, reinstatement searches (cf. [O'Brien, 1987] and [O'Brien & Albrecht, 1991]), anaphoric references (cf. [O'Brien et al., 1990]), the centrality of concepts (cf. [Albrecht & O'Brien, 1991]) and predictable inferences (cf. [Potts et al., 1985]) were investigated with the help of naming methods. The cross-modal naming task was used to investigate the processing of lexical ambiguities in sentences (cf. [Blutner & Sommer, 1988]).

Perceptual identification. This method is the most straightforward method in the field of research on spoken language comprehension. A word is presented auditorily and the subject has to identify it. In an open-set identification task, the subject can name any word, whereas, in a closed-set identification task, the subject has to select the correct word from a choice of different given alternatives (cf. [Lively et al., 1994]).
This method is used to obtain information about word storage in the mental lexicon and the influence word frequency and lexical similarity have on it (cf. [Luce, 1986]). Using the open-set identification task, [Savin, 1963] investigated typical error patterns a subject makes under difficult listening conditions.

Shadowing. This term refers to a method where the subject tries to repeat the spoken language of a speaker simultaneously (cf. [Lively et al., 1994]). The shadowing task is used to look into the relationship of different levels of language processing. [Marslen-Wilson, 1985], for example, found that subjects start syntactic and semantic analysis of the input during the shadowing task, i.e. the accuracy of the shadowing decreases when the input is syntactically or semantically anomalous. Another application of this method consists in identifying mispronunciation.

Phoneme restoration. In the phoneme restoration task, a phoneme in a word is replaced with noise or noise is added to the phoneme. The subject then has to identify the target word by restoring the masked phoneme and has to decide whether the noise is an addition to or a replacement of the phoneme (cf. [Lively et al., 1994]).
Using phoneme restoration, [Warren, 1970] found that the subjects hear the intact representation of the target word, when the noise replaces the phoneme. This task is also used to investigate the reception of word-initial vs. word-final information, word vs. non-word manipulations, priming effects as well as the relationship between speech perception and word recognition (cf. [Warren, 1970]).

Click detection task. In the click detection task, a subject listens to a sentence and at the same time to a non-linguistic sound (such as a click). Afterwards, the subject has to identify when the non-linguistic sound occurred (cf. [Ferreira & Anes, 1994]).
This method is used to test whether the click which occurs in the middle of a constituent is perceived in the middle or at the end of the constituent and thus to show whether or not constituents are perceptual units (cf. [Fodor et al., 1974]). [Fodor & Bever, 1965] and [Garrett et al., 1965] found that the click even tends to migrate to clause boundaries. In addition to these studies, [Geers, 1978] used the click detection task to investigate the relationship of prosody and sentence segmentation and found that prosody facilitates the perceptual click detection at clause and sentence boundaries.

Sentence memory task. In the sentence memory task, a text consisting of three sentences is presented auditorily or visually to the subject, with the last line always including an anaphor. When the anaphor occurs, the subject has to recall its antecedent. Having heard or read the text, the subject has to answer a comprehension question (cf. [Ferreira & Anes, 1994]).
[Jakimik & Glenberg, 1990] used the sentence memory task to obtain information about auditory language comprehension on the basis of temporal anaphors (such as "former" and "latter"). They concluded that it was easier for the subjects to remember the antecedents if the text is presented auditorily rather than visually. [Gernsbacher & Shroyer, 1989] used a variant of the sentence memory task to analyze cataphoric processing.

Picture-sentence verification task. In the picture-sentence verification task, first, a sentence is presented auditorily to the subject, then, a picture is shown. The subject decides whether or not the sentence was truthful according to the picture. Accuracy and reaction times are recorded (cf. [Ferreira & Anes, 1994]).
Studies based on the picture-sentence verification task show that the decision takes longer if the sentence is not truthful (cf. [Clark & Chase, 1972] and [Clark & Chase, 1974]), if the sentence contains a negation (cf. [Clark & Chase, 1972]) or if the sentence contains a passive construction (cf. [Gough, 1965], [Gough, 1966] and [Slobin, 1966]).

Monitoring task. Two kinds of monitoring tasks can be distinguished: the phoneme monitoring task and the word monitoring task. With both methods, a sentence or a part of a sentence is presented auditorily and the subject has to identify a particular morpheme or word by pressing a key. Accuracy and reaction times are then recorded (cf. [Ferreira & Anes, 1994]).
This method is used to study auditory phoneme processing (cf. [Cutler et al., 1986], [Cutler et al., 1987a], [Cutler et al., 1987b], [Dupoux & Mehler, 1990], [Frauenfelder et al., 1990] and [Schriefers et al., 1991]) as well as auditory word processing (cf. [Marslen-Wilson & Tyler, 1980] and [Tyler & Warren, 1987]).

End-of-sentence comprehension time. In the course of the end-of-sentence comprehension time task, a sentence is presented auditorily to the subjects who are then to press a key to indicate whether they have completely understood the sentence (cf. [Ferreira & Anes, 1994]).
[Slowiaczek, 1981] employed this method to investigate garden-path sentences in combination with prosody.

Gating task. Here, a fragment of a word (in the word-gating task) or a fragment of a structurally ambiguous sentence (in the sentence-gating task) is presented auditorily to the subject and the subject has to predict how the word or sentence continues. In a variant of this method, parts of the word or sentence are masked by noise (cf. [Ferreira & Anes, 1994] and [Lively et al., 1994]).
[Beach, 1991] used this method to investigate the relationship between syntax and prosody and found that the duration of the presentation of verbs has an effect on the continuation of the sentence. [Grosjean, 1980] investigated word recognition on the basis of this method masking either the beginning or the end of the word.

Sentence completion task. Similarly to the sentence-gating task, a fragment of an ambiguous sentence is presented to the subject, but on a written basis. The subjects then have to complete the sentence by writing down how they expect the sentence to continue.

[Scheepers & Corley, 2000], [Branigan et al., 1995] and [Pickering & Branigan, 1998] used the sentence completion task to investigate syntactic priming (i.e. the tendency to re-use syntactic structures) for written language.

Corpus of written language. Using a corpus of written language means that a large number of texts representative of the written language under investigation can be exploited to analyze language performance. The corpus can be investigated as raw text, but also enriched with information on part-of-speech, syntax, semantics, discourse etc. Additionally, the corpus is usually marked up with meta-information concerning the individual texts or sub-corpora. The most common form of corpus analysis consists in using a treebank, i.e. a corpus linguistically interpreted with part-of-speech and syntactic annotation (cf. [Marcus et al., 1993]).
[Doherty, 1999a] uses an English-German parallel corpus of raw texts to investigate English cleft sentences in connection with their corresponding German realizations and to describe the strategies the translators used (see also [Doherty, 1998b]). [Uszkoreit et al., 1998] analyzed a German treebank in order to test Hawkins's hypotheses on German relative clause extraposition (cf. [Hawkins, 1994]). Similar studies of word order variation in German have been undertaken by [Kurz, 2000] and [Kurz et al., 2000]. Corpora are furthermore used for pretests: [Sturt et al., 1999] investigated a random sample of the British National Corpus to gain results regarding verb bias, which they used as a pretest for a window experiment dealing with structural change and reanalysis difficulties in language comprehension.

Corpus of spoken language. Similarly to a corpus of written language, a corpus of spoken language comprises a large number of texts representative of the spoken language under investigation. Usually, such a corpus is transcribed and meta-information is added. The corpora can, again, be investigated as raw text, but also be enriched with information on part-of-speech, syntax, semantics, discourse etc. (cf. [Bird & Harrington, 2001]).
[Pöchhacker, 1994] uses a corpus of interpreted texts to study simultaneously interpreted conference speeches and to describe the interpreter's strategies.

Rating. This term refers to methods where subjects make a statement about the acceptability of sentences with the help of a scale in which 1, for instance, stands for "not acceptable" and 7 means "good". The subjects are to take into account the plausibility and grammaticality of the sentences (cf. [Haberlandt, 1994]).
This kind of experiment is frequently used as a pretest for or in combination with other experiments: [Sturt et al., 1999] analyzed the relation between reanalysis difficulties and the plausibility of misanalyses. The results were gained by performing a rating task and served as a pretest for a window experiment carried out to gain insight into structural change and reanalysis difficulties in language comprehension.

Magnitude estimation. Magnitude estimation means that subjects decide whether a sentence is acceptable in terms of plausibility and grammaticality. This decision is taken with respect to a given reference sentence, meaning that a sentence can be, for example, twice as good or half as good as the reference sentence with the marks being freely chosen by the subjects (cf. [Haberlandt, 1994]).
Like ratings, magnitude estimations often serve as pretests for or in combination with other experiments: on the basis of magnitude estimations, [Uszkoreit et al., 1998] analyzed the acceptability of relative clauses which are extraponed and of relative clauses which are not. These results were used in combination with the results gained from a corpus study to test hypotheses relating to German relative clause extraposition.

Electroencephalogram (EEG). With the help of scalp electrodes, this method allows the measurement of the brain's electrical activity in terms of ion flow. In order to gain insight into language comprehension and production, the subject is involved in different language-related tasks during the EEG (cf. [Kutas & van Petten, 1994]).
This method was applied to show that the right hemisphere is more active during a verbal task than that the left hemisphere. Furthermore, the use of EEG also shows that the converse is true for the non-verbal task (cf. [J. Doyle & Galin, 1974], [Ehrlichman & Wiener, 1980] and [Galin & Ornstein, 1972]).

Event-related potential (ERP). The ERP is similar to the EEG since scalp-recorded electrical ion activity is measured. The main difference, however, is that the activity can be time-locked (i.e. synchronized) to some external event which can be a specific, language-related stimulus. The measurement of magnetic event-related fields, which is called magnetaencephalogram (MEG) is a variant of the ERP (cf. [Kutas & van Petten, 1994]).
This method can, for instance, be employed to investigate word recognition and the relationship between word frequency and sentence context (cf. [van Petten & Kutas, 1990] and [van Petten & Kutas, 1991]). [Schmidt et al., 1989] investigated language-processing paradigms using the MEG.

Magnetic resonance imaging (MRI). With this method, the body is covered by electromagnets which reveal densities of hydrogen atoms and their interactions with other tissues. Since, however, the forming of images on the basis of MRI is very slow, a faster variant, called echo-planar imaging (EPI), was developed. This method visualizes images of the functional activity of the brain within 30 milliseconds and displays the images with a high resolution (cf. [Solso, 1995] and [Eysenck & Keane, 1990]).
These methods are used for visualizing and localizing brain structures as well as processes involved in language-related tasks (cf. [Solso, 1995]).

Tomography scan. In connection with tomography scanning methods, three different kinds of scans can be distinguished: the computerized axial tomography scan (CAT scan), the dynamic spatial reconstructor scan (DSR scan) and the positron emission transaxial tomography scan (PETT scan) (cf. [Solso, 1995] and [Eysenck & Keane, 1990]). During the CAT scan, an X-ray machine rotates around the head, sending X-ray beams through the brain. Sensitive detectors are fixed on the opposite side of the X-ray's source to receive the beam and record the corresponding data. The pictures produced by every single X-ray beam together provide a cross section of the brain which shows the flow of blood in the brain. Since more blood flows through active parts of the brain, cognitive structures can be detected on the basis of tomography scans. Revealing internal brain structures in a three-dimensional way, the DSR method is a more elaborated variant of the CAT scan. PETT scans are similar to CAT scans except for the fact

that they use detectors for the measurement of radioactive particles in the blood. As mentioned above, active parts of the brain require more blood and thus attract more radioactive tracers. These tracers emit rays which are used to create visual maps of the cognitive structures (cf. [Solso, 1995] and [Eysenck & Keane, 1990]).
Tomography scans enable researchers to determine which part of the brain is active during spoken language comprehension in contrast to written language comprehension as well as spoken language production in contrast to written language production. Furthermore, this method reveals which part of the brain is involved in which kind of language-related task (cf. [Solso, 1995]).

Think-aloud protocol (TAP). TAPs are used to gain verbal data with the help of introspection. In order to record a TAP, the subject has to "think aloud", i.e. to verbalize whatever goes on in the head. The subject's verbalization is recorded on tape or videotape, and afterwards the protocols are transcribed. TAPs have a long tradition as an experimental method in general psychological studies (cf. [Claparède, 1932] and [Duncker, 1935]).
A more recent field of application of TAPs is concerned with translation studies, i.e. the investigation of translation performance, translation strategies and the translation process (cf. [Kussmaul & Tirkkonen-Condit, 1995], [Tirkkonen-Condit, 1997], [Lörscher, 1991c], [Lörscher, 1991b], [Lörscher, 1991a], [Krings, 1986b], [Krings, 1987] and [Königs, 1987]). Furthermore, TAPs are used to gain insight into the behavior of translation students or language learners (cf. [Krings, 1986a], [Jääskeläinen, 1993] and [Séguinot, 1991]) and make it possible to carry out comparisons between professional and non-professional translation (cf. [Jääskeläinen, 1989], [Tirkkonen-Condit, 1989] and [Jääskeläinen & Tirkkonen-Condit, 1991]).

From this introduction of the different methods available in the field of experimental psycholinguistics, it can be seen that overlaps exist between the different methods (e.g., the sentence completion task can be interpreted as the written version of the gating task) and that the different methods can be combined (e.g., rating can be used as a pretest to the window method). These overlaps or combinations are applied in many analysis scenarios. Table 2.13 contains a classification of the methods according to the different

types of methods introduced in Table 2.12.

eye-tracking method	written	on-line	comprehension
window method	written	on-line	comprehension
recognition method	written	on-line	comprehension
lexical decision task	written/spoken	on-line	comprehension
naming method	written/spoken	on-line	comprehension
perceptual identification	spoken	off-line	comprehension
shadowing	spoken	on-line	comprehension
phoneme restoration	spoken	on-line	comprehension
click detection task	spoken	off-line	comprehension
sentence memory task	written/spoken	off-line	comprehension
picture-sentence verification task	spoken	on-line	comprehension
monitoring task	spoken	on-line	comprehension
end-of-sentence comprehension time	spoken	on-line	comprehension
gating task	spoken	off-line	production
sentence completion task	written	off-line	production
corpus of written language	written	off-line	production
corpus of spoken language	spoken	off-line	production
rating	written	off-line	comprehension
magnitude estimation	written	off-line	comprehension
EEG	written/spoken	on-line	production/ comprehension
ERP	written/spoken	on-line	production/ comprehension
MRI	written/spoken	on-line	production/ comprehension
tomography scan	written/spoken	on-line	production/ comprehension
TAP	written/spoken	on-line	production/ comprehension

Table 2.13: Classification of methods in experimental psycholinguistics

2.5 Summary and conclusions

The purpose of this chapter was to give insight into theoretical issues pertaining to product-based as well as process-based linguistic research. First, the empirical methods as used in linguistics were introduced in Section 2.2. Within this section, corpus-based translation studies (cf. Section 2.2.1) and corpus-based register analysis (cf. Section 2.2.2) were presented. On

the basis of the exploitation of these empirical methods (cf. Section 3.2), it is possible to generate hypotheses (cf. Section 5.2) and to develop an appropriate methodology (cf. Section 4.2) for the cross-linguistic analysis. Secondly, different methods of cross-linguistic description were introduced (cf. Section 2.3). In Section 2.3.3, the multilingual SFL approach, which provides the theoretical framework (cf. Section 3.3) for the methodology developed in Section 4.3, and the analysis carried out in Section 5.3, was discussed. Additionally, a summary of classical approaches in contrastive linguistics (cf. Section 2.3.1) and translation studies (cf. Section 2.3.2) was presented. Finally, a comprehensive overview of the potential experimental methods available for the psycholinguistic experiment designed and carried out in Sections 3.4, 4.4 and 5.4 was given in Section 2.4.

This chapter showed that a detailed description of the empirical, cross-linguistic and psycholinguistic methods is the basis for deciding which methods are eligible in connection with the interdisciplinary methodology for investigating the specific properties of translations. However, since a combination of the different methods, i.e. of product- and process-oriented research, is requisite, Chapter 3 discusses which empirical, cross-linguistic and psycholinguistic methods are suitable to be chosen for this thesis. In this context, it is also spelled out which existing theoretical concepts can be employed and which concepts have to be newly introduced.

Chapter 3

Theoretical foundation

3.1 Introduction

The motivation for choosing the methods introduced in Sections 2.2 to 2.4 has been the development of a methodology for the investigation of the nature of translated text (cf. Chapter 4) and its realization (cf. Chapter 5). To gain a better idea of the usability of these methods especially for this thesis, in the following the advantages and disadvantages of each method are briefly discussed. For each approach (empirical, cross-linguistic and psycholinguistic), it is explained which theoretical concepts can be used in this thesis, which concepts have to be extended or narrowed and which have to be newly introduced.

Section 3.2 deals with the theoretical foundations of empirical linguistics. Section 3.3 discusses the advantages and disadvantages of the methods of cross-linguistic description. In Section 3.4, the usability of the different psycholinguistic experiments is assessed. The chapter concludes with a summary (cf. Section 3.5), which gives an overview of the theoretical methods employed in order to develop an interdisciplinary methodology for the investigation of the nature of translated text (cf. Chapter 4).

3.2 Empirical linguistics

In translation studies, Baker's approach (cf. [Baker, 1996]) is revolutionary in the sense that she was the first to investigate translated text in its own

right and not only for translation evaluation and criticism. This crucial development is highly relevant in connection with this thesis. It allows the empirical analysis of rather intuitively developed hypotheses relating to the nature of translated text and the translation process.

The problem with Baker's approach to corpus-based translation studies, which is the prototypical approach in this area, is that the features used for analysis (such as word counts) are rather shallow, which makes it difficult to explain or interpret the rather abstract hypotheses on the universal features of translations on the basis of this kind of features (e.g., optional *that* for explicitation). Without a comprehensive theory to substantiate the interpretation of the results gained empirically, a gap remains between hypothesis generation and hypothesis testing. Another problem with this kind of features is that they cannot be used cross-linguistically since the typological differences have to be taken into account. The analysis of, for example, simplification on the basis of type-token ratio, lexical density and average sentence length, which entails word counts, is rather problematic for the comparison of English and German, since the different morphological typologies bias the results (cf. [Hansen & Teich, 1999]). Thus, other ways have to be found of operationalizing the testing of the rather abstract hypotheses on a text basis, on the one hand, and for multilingual environments, on the other. Another shortcoming of Baker's approach is the corpus design that was chosen, as the corpus consists of translations only, while SL texts are neglected. This rules out the investigation of SL texts and their influence on the translations. Whether the SL influence can be leveled out through the collection of texts translated from various SLs is questionable. However, even if this is the case, the investigation of the influence of an individual SL on its translation is not possible. Another problematic issue concerning Baker's corpus design is that the corpus comprises rather broad registers. This means that register-related properties which can be found in translations cannot be detected. The notion of register is also crucial for the investigation of normalization, because it offers a profound framework for defining what the norm could be (cf. [Teich, 2001a]). The broader the register is, the harder it is to define this notion of norm. However, as can be seen in Section 2.2.1, many projects have started from Baker's approach and have then geared their own corpus designs to the requirements of their hypothesis testing, thus overcoming the disadvantages of Baker's approach.

The advantage of the different functional dimensions [Biber, 1995] found in his corpus-based register analysis is that they are defined on the basis of the statistically significant co-occurrence of lexico-grammatical features. Since [Biber, 1995] used the functional dimensions for cross-linguistic comparison, the lexico-grammatical features chosen for the register analysis can be applied to other languages as well.
A disadvantage regarding Biber's approach is that it is unclear why some features are chosen while others are left aside, and whether this omission has an influence on the definition of the functional dimensions (cf. [Lee, to appear]). Consequently, it is completely unclear which role the chosen features play within the language system. Another problem with Biber's approach is that the grammatical features which are analyzed are not classified according to their form (e.g., *wh*-clauses, nouns or preposition) and function (e.g., *that*-clauses in object position). This distinction is of particular interest when different language systems, i.e. formal and structural realizations of different languages, are involved. It is also unclear where the functional dimensions are located in the language system. Since they are abstractions from the co-occurring lexico-grammatical features, their interpretations can be derived from either semantics or discourse or pragmatics. Thus, the degree of abstraction in terms of linguistic strata cannot be ascertained. Finally, the different functional dimensions being strongly correlated, it may be doubted whether they can be grouped into only one dimension which distinguishes between spoken vs. written mode. In some cases, it is even unclear whether the dimension is named after the lexico-grammatical features or after the register representing it (e.g., the narrative vs. non-narrative dimension). In these cases, the classification of the functional dimensions would be predictable and does not offer novel insight.

With a view to the empirical investigation of the specific properties of translated text, the present thesis makes use of both Baker's corpus design and Biber's empirical register analysis. The notion of register in general and Biber's register features in particular are used in order to define what is normal (cf. [Teich, 2001a]) and, consequently, what can be interpreted as normalization. Thus, Biber's register features are used for the comparison of translations and originals in the TL in order to test Baker's normalization

hypothesis. A problem would arise, however, if the translations showed a contrary tendency of normalization. Therefore, the term *anti-normalization* is introduced in view of translations which show a tendency contrary to normalization in the sense that they do not use patterns typical of the TL. Furthermore, Biber's functional dimensions are used to explain the register shifts in translations which can be found in the cross-linguistic analysis.

3.3 Cross-linguistic description

Hawkins' typology of English and German (cf. [Hawkins, 1986]) as well as Vinay and Darbelnet's comparative stylistics of French and English (cf. [Vinay & Darbelnet, 1995]) offer a rather example-based description of the typological differences between the languages involved. The problem, however, is that it is not obvious why the differences they describe were chosen and whether any other differences may be neglected in their typologies. Furthermore, a comparative overview of the commonalities would be useful as well. For this reason, a description of the different language systems or one that is motivated by the different language systems (provided, for instance, in [Teich, 2001a] and [Teich, 2001b] for English-German) would be more comprehensive.

The same phenomenon as can be observed for the comparative typologies applies to the translation techniques. [Vinay & Darbelnet, 1995] and [Newmark, 1988] both provide a rather example-based description of the translation techniques. The problem, again, is that it is unclear why the techniques they describe are chosen and whether there are other techniques which are left aside in their observations. A description of all phenomena occurring in TL texts, as compared to SL texts, motivated by the different language systems involved (as suggested in [Steiner, 2001b] for English-German) would be more comprehensive.

There are many points of favor of using SFL as a method for cross-linguistic description (see also [Matthiessen & Bateman, 1991]):

- Historical perspective: the SF model has been used in many fields of application, for instance in the area of text generation (cf. [Bateman et al., 1991], [Teich, 1995b] and [Teich et al., 1996]) and translation studies (cf. [Steiner, 1997], [Steiner, 1998] and [House, 1997]).

- Theoretical comprehensiveness: its comprehensive character represents another advantage of the SF model. Many other theories deal with one (or only a few) of the different strata, but do not encompass all strata of language. However, it is extremely important to take on a comprehensive approach because the different strata play an important role for the contrastive analysis of translations and comparable texts.

- Paradigmatic orientation: with a view to the description of translation methods, it is extremely important to understand the paradigmatic choices. Thus, a model is required which emphasizes the importance of paradigmatic choices and reflects the conception of language as a resource. The system networks of the SF theory serve as a paradigmatic basis, whereas grammatical realizations are considered separately.

- Functional orientation: working with multilingual texts requires a theory with a functional orientation. The lack of such an orientation would make it extremely difficult to determine whether the linguistic patterns of the SL texts, the TL texts and the comparable texts fulfill the same or different functions.

- Theoretical research questions: the SF model is comprehensive enough to develop parameters for cross-linguistic description. Thus, it provides an ideal basis for generating new hypotheses concerning the specific features of translations.

- Exchange between theory and application: within the SF model, not only is it possible to generate hypotheses, but also to test them and to describe them with the help of the SF parameters for cross-linguistic description. Thus, while hypotheses remain rather abstract and are not sufficiently operationalized with many other theories, the SF model provides concrete linguistic parameters for such an analysis.

The only disadvantage of SFL is that the actual analyses involve a high degree of interpretation. Although it offers a set of operationalization rules, the analyses are conditioned by subjective factors. The fact that the operationalization rules are well documented for English, but not for German, French or other languages, represents a further drawback of SFL.

For the present thesis, Hawkins' typology of English and German as well as Vinay and Darbelnet's comparative stylistics of French and English are used to explain the typologically motivated translations to be found in the cross-linguistic analysis. Since the translation techniques described in [Vinay & Darbelnet, 1995] and [Newmark, 1988] are not sufficiently systemized, SFL is used to describe the relationship between the SL text and its translation. For this purpose, the notion of grammatical metaphor provides the basis for describing the transposition-like translation techniques: the term *metaphorization* is used in cases where the TL expression is more metaphorical than the SL expression, while the term *de-metaphorization* describes cases in which the SL is more metaphorical than the TL expression. The term *re-metaphorization* labels translations showing the same degree of grammatical metaphoricity as the SL expression. For the description of other translation techniques (cf. [Vinay & Darbelnet, 1995] and [Newmark, 1988]), the notion of texture provided by SFL is introduced: changes in texture can take place at a structural level (i.e. thematic structure and information structure and focus) and at a cohesive level (i.e. reference, ellipsis, substitution, conjunction and lexical cohesion). In view of comparing an SL text and its translation, this means that cross-lingual changes in texture often trigger an ambiguation or disambiguation of the translated text. If, for example, an ellipsis used in the SL text is spelled out in the TL text, the translation will be less ambiguous than the corresponding passage in the SL text. Thus, the terms *disambiguation-cohesion*, *ambiguation-cohesion*, *disambiguation-structure* and *ambiguation-structure* indicate whether the TL texts are ambiguated or disambiguated in terms of structure or cohesion.

3.4 Experimental psycholinguistics

The use of psycholinguistic experiments has many advantages: the experiments provide a large amount of data for investigation. The product-oriented data and especially the process-oriented data allow interpretations concerning the processes inherent in language production and comprehension. Methods, such as EEG, ERP, MRI and tomography scans, offer insight into the brain, thus allowing the neuro-physical investigation of the processes. The manipulation of the experiments allowing the control and separation of each parameter (including stimuli) of the experimental design is another plus point.
Nonetheless, there are also some problems involved in the use of psycholinguistic experiments, in the sense that most of the experiments focus on single problems and phenomena and can thus only be used in rather restricted research areas. Gearing them to other research purposes is quite problematic. Another problem is that some of the methods are rather expensive (e.g., eye-tracking, EEG, ERP, MRI and tomographic scans) and require special hardware. As to the interpretation of the results, gaps frequently occur between hypothesis generation and hypothesis testing. Some interpretations are problematic because they are too dependent on personal interpretation, memory or behavior of the subjects. These are factors which cannot be controlled in the experimental design (e.g., in the recognition method, the lexical decision task or the naming method). This argument is strengthened looking at the fact that the experiment is not a natural situation at all. For example in TAPs, where the translator has to verbalize every thought and every process, the verbalization task renders the translation situation rather unnatural and distracting, forcing the translator to think about what is happening and to verbalize it immediately, thus disturbing the translation process. With experiments of this kind, however, unconscious processes are totally neglected. The unconscious processes taking place during a translation task can only be detected by combining TAPs with other experimental designs.

In the context of this thesis, a psycholinguistic experiment is used to study the specific properties of translated text which result from the translation process. Since none of the psycholinguistic experiments described in Sec-

tion 2.4 meets the requirements of such a process-oriented investigation, experiments have to be chosen which, in combination, take into account the cognitive processes relevant to the translation tasks a translator is engaged in (i.e. text understanding as well as text production) without, however, disturbing the translation process. Thus, TAP, sentence completion task, gating task and lexical decision task are combined in order to investigate the translation process, the translator's verbalization immediately after the translation process and the translated text itself. The lexical decision task makes it possible to analyze the translation process under different constraints, within the sentence completion and the gating task, text production is focused on and TAPs allow the investigation of the verbalization.

3.5 Summary and conclusion

Having discussed all advantages and disadvantages of the empirical (cf. Section 3.2), cross-linguistic (cf. Section 3.3) and psycholinguistic (cf. Section 3.4) methods, it can clearly be seen that a comprehensive methodology for the investigation of the nature of translated text cannot be carried out on the basis of only one of these approaches, since they all fail to meet the requirements specified in Section 1.3. Thus, a combined approach on the basis of all these methods is needed which exploits the advantages of every method and avoids the disadvantages. For this purpose, the notions of normalization and anti-normalization were discussed in Section 3.2. In order to explain these phenomena, the concepts of grammatical metaphoricity and structural or cohesive dis-/ambiguation were derived from the SF model and coined in Section 3.3. Section 3.4 showed which psycholinguistic methods are to be combined with a view to investigating the translation process. Chapter 4 discusses the manner in which these theoretical concepts are employed for the product- and process-oriented analysis of translation properties.

Chapter 4

Methodology

4.1 Introduction

This chapter presents the methodology for the investigation of translated text developed on the basis of the empirical, cross-linguistic and psycholinguistic methods presented in Chapters 2 and 3. Thus, this methodology reflects the different parts of the analysis to be carried out (cf. Chapter 5): the empirical approach is discussed in Section 4.2, the cross-linguistic investigation is presented in Section 4.3, and the psycholinguistic experiment is introduced in Section 4.4.

Since Section 4.2 is based on the analysis of a large English comparable corpus and Section 4.3 is based on the analysis of a smaller multilingual corpus, the subsections of each of these sections are concerned with the same issues: first, the different analysis scenarios are to be introduced (cf. Sections 4.2.1 and 4.3.1). Then, the corpus designs of the different corpora are presented (cf. Sections 4.2.2 and 4.3.2). After the linguistic interpretation of the corpora (cf. Sections 4.2.3 and 4.3.3), their formal representation is discussed (cf. Sections 4.2.4 and 4.3.4). Finally, the different methods of corpus querying are presented (cf. Sections 4.2.5 and 4.3.5). Additionally, significance tests are introduced for the first corpus analysis (cf. Section 4.2.6), with a view to checking the statistical validity of the results.

In connection with the psycholinguistic approach (cf. Section 4.4), the psycholinguistic analysis scenario (cf. Section 4.4.1) as well as the experimental design (cf. Section 4.4.2) are described.

Section 4.5 summarizes the methodology regarding its eligibility for the present thesis.

A combination of the empirical approach, the cross-linguistic approach and the psycholinguistic approach is necessary to compensate the disadvantages and to exploit the advantages of each method. Thus, the dependent and independent variables vary according to the different approaches (cf. Chapter 1). Register, typology and cognitive processes process constitute possible sources of translation properties. Being interacting sources of the specific properties of translations, they are not explicitly separated as variables for this study. This means that it cannot be investigated on an empirical basis which source is responsible for which translation property. Within the framework of an example-based discussion, however, the empirical approach rather deals with register-specific translation properties since a large comparable corpus controlled in terms of register is investigated; the cross-linguistic approach mainly discusses typological differences between the languages involved since a parallel corpus of SL and TL texts is subject of investigation; and the psycholinguistic approach looks at translation properties owing to the translation process itself. Moreover, this combination of product- and process-oriented research allows the description as well as the explanation of the specific properties of translations (cf. Chapter 5).

4.2 Empirical approach

This section presents the empirical approach of the present thesis, i.e. the methodology for investigating a large English comparable corpus is discussed in detail. This includes decisions and proposals in view of the following corpus-related matters: analysis scenario (cf. Section 4.2.1), corpus design (cf. Section 4.2.2), corpus annotation (cf. Section 4.2.3), corpus representation (cf. Section 4.2.4), corpus querying (cf. Section 4.2.5) and significance test (cf. Section 4.2.6).

4.2.1 Analysis scenario

This section deals with the investigation of the relation between English translations and English originals, i.e. the investigation of translated text as special kind of text type or register. This analysis is based on Baker's normalization hypothesis (cf. [Baker, 1996]), as the notion of register is used

to substantiate the definition of what is "normal" (cf. [Teich, 2001a]). With this in view, Biber's lexico-grammatical register features (cf. [Biber, 1995]) are taken into account. Since the corpus consists of English fiction texts (cf. Section 4.2.2), the following functional dimensions are of relevance (cf. Section 2.2.2): Dimension 2 (narrative vs. non-narrative discourse), Dimension 3 (situation-dependent vs. elaborated reference), Dimension 5 (abstract vs. non-abstract style) and Dimension 6 (on-line informational elaboration marking stance). Taking together all the sub-registers of fiction (general fiction, mystery fiction, science fiction, adventure fiction and romantic fiction), fiction can be characterized as narrative (i.e. the positive features of Dimension 2 are typical, while the negative ones are untypical), situation-dependent (i.e. the positive features of Dimension 3 are typical, whereas the negative ones are untypical), non-abstract (i.e. the positive features of Dimension 5 are typical, whereas the negative ones are untypical) and edited (i.e. the positive features of Dimension 6 are untypical, whereas the negative ones are typical). The combination of the functional dimensions relevant to fiction results in the following list of typical and untypical fiction features[1]:

- typical features:
  - past tense verbs,
  - third person pronoun,
  - perfect aspect,
  - public verbs,
  - synthetic negation,
  - present participle clauses,
  - time adverbials,
  - place adverbials,
  - adverbs,
  - phrasal coordination;

[1]From the perspective of Dimension 6, phrasal coordination is a typical feature, but an untypical feature for Dimension 3. Since its statistical value in Dimension 6 is too low to be significant it is interpreted as an untypical feature (belonging to Dimension 3). The fact that it is the only typical feature in Dimension 6, but co-occurring with other untypical features in Dimension 3 supports this decision.

- untypical features:
  - present tense verbs,
  - attributive adjectives,
  - *wh*-relative clauses,
  - pied piping,
  - phrasal coordination,
  - nominalizations,
  - conjuncts,
  - agentless passives,
  - *by*-passives,
  - past participle clauses,
  - subordinators,
  - *that*-clauses,
  - demonstratives,
  - final prepositions,
  - existential *there*.

According to this list of lexico-grammatical features, a similar use (or overuse) of typical features as well as similar use (or underuse) of untypical features in English translated fiction texts (compared to English originals) would support Baker's normalization hypothesis since the TL texts would conform to (or even exaggerate) the norm of the TL. However, the notion of normalization has to be extended for this kind of analysis since the contrary tendency is investigated as well. The overuse of untypical features as well as the underuse of typical features in English translated fiction texts would therefore be seen as an indicator of anti-normalization since the translated texts would not conform to the norms of the TL (see also Sections 2.2.1 and 3.2).

The analysis of normalization and anti-normalization on the basis of the typical and untypical fiction features requires the following steps: the annotation of a comparable corpus (cf. Sections 4.2.2 and 4.2.4) with part-of-speech tags (cf. Section 4.2.3), the extraction of text instances carrying

the relevant lexico-grammatical features on the basis of these tags (cf. Section 4.2.5) and finally, the norming and significance testing of the results (cf. Section 4.2.6).

4.2.2 Corpus design

Since Baker's normalization hypothesis serves as basis for the investigation in this section, her criteria concerning corpus design are also taken into account (cf. [Baker, 1996]). Thus, the corpus under investigation in this section consists of an English comparable corpus comprising of a sub-corpus of English translations and a sub-corpus of English originals. The sub-corpus of English translations is taken from the fiction part of the Translational English Corpus (TEC), whereas the sub-corpus of English originals is extracted from the fiction part of the British National Corpus (BNC). Thus, both sub-corpora belong to the register of fiction, as mentioned in Section 4.2.1. The files which are included in the sub-corpora are listed in Appendix 7.1. The translational sub-corpus is made up of translations from several languages into English. Information on the SLs, the titles of the texts, the authors and the translators can also be found in Appendix 7.1. The translational sub-corpus consists of 4,843,763 words and the original sub-corpus includes 4,741,500 words (9,585,263 words in total). From this it can be seen that the corpora are compiled in such a way as to make them as comparable as possible, both in terms of register and in terms of size.

4.2.3 Corpus annotation

Since the corpus under investigation in this section is quite large (approximately 10 million words) and the features to be analyzed are rather shallow, linguistic annotation can be carried out automatically. Part-of-speech tagging is a fairly reliable method of annotation, either using a rule-based or a statistical approach. Recently, however, statistical approaches have become more popular. For this reason, the tagger which has been employed, the TnT tagger, is a statistical part-of-speech tagger that analyzes trigrams, incorporating several methods of smoothing and of handling unknown words (cf. [Brants, 1999] and [Brants, 2000]). The system can be trained to deal with different languages and comes with the Susanne tagset for English

(cf. [Sampson, 1995]) and the Stuttgart-Tübingen tagset (STTS) for German (cf. [Schiller et al., 1999]). It includes a tool for tokenization, which is a preparatory step in the tagging process. In basic mode, not only does the tagger provide each token with a part-of-speech tag, but it omits alternative tags and also performs probability calculations. It analyzes between 30,000 and 60,000 tokens per second and has an accuracy of about 97%.

	TEC	BNC
tokens	5,730,790	4,872,694
unknown tokens	549,613	445,083
unknown tokens	9.59%	9.13%
cardinals/ordinals	2,741	1,761
avg. tags/token	3.97	3.87
avg. tags/unknown token	2.27	2.20

Table 4.1: Descriptive statistics of the TnT output for TEC and BNC

During the tagging process of the comparable corpus looked at in the empirical investigation, TnT produced the descriptive statistics listed in Table 4.1. The number of tokens[2] and the number of unknown tokens are the most important results taken from the descriptive statistics displayed in Table 4.1. Both figures are higher for TEC which corresponds to the higher word count for TEC (cf. Section 4.2.2). The percentage of unknown tokens amounts to 9.59% for TEC and 9.13% for BNC, i.e. the tagger cannot find these words in its lexicon but has to assign tags to these words by itself[3]. A TEC sample output of TnT in tab separated vector (TSV) format is shown in Figure 4.1. The part-of-speech tags used for tagging TEC and BNC are based on the Susanne tagset (cf. [Sampson, 1995]) and can be found in Appendix 7.2.

4.2.4 Corpus representation

There are two TEI-conformant headers for the files in TEC: a header for single volumes (see Figure 4.2) and a header for collected works (see Figure 4.3). Each element in the headers has a start tag and an end tag. A start tag at the beginning of an element is represented by a balanced pair

[2]Note that this number consists of a word count plus punctuation marks.

[3]Note that the percentage of unknown words has nothing to do with the tagger's percentage of accuracy since it is able to handle unknown words using its probabilistic trigram analysis.

Figure 4.1: TnT sample output of TEC

of angle brackets containing annotation strings, while a slash preceding the annotation strings indicates an end tag.
As can be seen in Figure 4.2, the TEC header for single volumes includes the following information:

- title of the book,
- translator (where status refers to the question whether the translators have a full-time or part-time job and whether they work on a free-lance or in-house basis),
- translation (where extent means the number of words),
- translation process (where direction is relevant to whether or not the SL text is translated into the translator's mother tongue and type means a full translation in contrast to a summary, gist or excerpt),
- author of the original,
- SL text (where status refers to the problem whether the text is an original or a translation).

The TEC header for collected works comprises an additional element called *section*. The section is repeated for each article or story contained in the collection. It includes information on the translator, the translation, the translation process, the author and the SL text of each article, story or

```
<Header>
     <title>
          <filename></filename>
          <subcorpus></subcorpus>
          <collection></collection>
          <editor></editor>
     </title>
     <translator>
          <name></name>
          <gender></gender>
          <sexualOrientation></sexualOrientation>
          <Nationality></Nationality>
          <employment></employment>
          <status></status>
     </translator>
     <translation>
          <mode></mode>
          <extent></extent>
          <publisher></publisher>
          <pubPlace></pubPlace>
          <date></date>
          <copyright></copyright>
          <sponsor></sponsor>
          <reviews></reviews>
          <comments></comments>
     </translation>
     <translationProcess>
          <direction></direction>
          <mode></mode>
          <type></type>
     </translationProcess>
     <author>
          <name></name>
          <gender></gender>
          <sexualOrientation></sexualOrientation>
          <Nationality></Nationality>
     </author>
     <sourceText>
          <language></language>
          <mode></mode>
          <status></status>
          <publisher></publisher>
          <pubPlace></pubPlace>
          <date></date>
          <comments></comments>
     </sourceText>
</Header>
```

Figure 4.2: Header for single volumes in TEC

```
<Header>
     <title>
          <filename></filename>
          <subcorpus></subcorpus>
          <collection></collection>
          <editor></editor>
     </title>
     <translation>
          <mode></mode>
          <extent></extent>
          <publisher></publisher>
          <pubPlace></pubPlace>
          <date></date>
          <copyright></copyright>
          <sponsor></sponsor>
          <reviews></reviews>
          <comments></comments>
     </translation>
     <section>
          <sectionid></sectionid>
          <text></text>
          <translator>
               <name></name>
               <gender></gender>
               <sexualOrientation></sexualOrientation>
               <Nationality></Nationality>
               <employment></employment>
               <status></status>
          </translator>
          <translation>
               <mode></mode>
               <extent></extent>
               <copyright></copyright>
               <comments></comments>
          </translation>
          <translationProcess>
               <direction></direction>
               <mode></mode>
               <type></type>
          </translationProcess>
          <author>
               <name></name>
               <gender></gender>
               <sexualOrientation></sexualOrientation>
               <Nationality></Nationality>
          </author>
          <sourceText>
               <language></language>
               <mode></mode>
               <status></status>
               <publisher></publisher>
               <pubPlace></pubPlace>
               <date></date>
               <comments></comments>
          </sourceText>
     </section>
</Header>
```

Figure 4.3: Header for collected works in TEC

```
<title></title>
<head></head>
<sbhead></sbhead>
<chapter n="...">
<p>
<frontmatter></frontmatter>
<backmatter></backmatter>
<footnote></footnote>
<endnote></endnote>
<caption></caption>
```

Figure 4.4: Textual mark-up in TEC

paper in the collection.
In addition to the header, the body of each file contains meta-information on the text structure represented in a modified version of the Standard Generalized Mark-Up Language (SGML) (see Figure 4.4 for the most important tags). For example, the following mark-up is included in each file of TEC: information on the (sub-)headings of the chapters or sections, the chapter number, the frontmatter (including introduction, preface etc.), the backmatter (including afterword, bibliography etc.), footnotes, endnotes, captions (e.g., for pictures or tables) and paragraphs[4].

The TEI-conformant header of the BNC includes the main elements displayed in Figure 4.5. The BNC header consists of a file description, an encoding description, a profile description and a revision description. The file description contains information on the title, the edition, the extent (i.e. size), the publication and the bibliographic source. The purpose of the coding project, the sampling criteria, the editorial principles, the linguistic annotation, the structure of the canonical references and the classification codes used for the texts within the corpus are spelled out in the encoding description. The profile description provides insight into the creation of the text, the language usage, the participants as well as their interaction, the settings of the communicative situation and the classification scheme with which the texts are categorized. The revision description explains the major changes which have taken place during the revision process.
The mark-up scheme of the BNC is an standardized SGML application (ISO 8879). The main elements of the textual mark-up displayed in Figure 4.6 describe the use of headings, segments, words, punctuation, texts,

[4]For further information on TEC refer to the following URL: http://ceylon.ccl.umist.ac.uk/.

```
<teiHeader>
     <fileDesc>
          <titleStmt></titleStmt>
          <editionStmt></editionStmt>
          <extent></extent>
          <publicationStmt></publicationStmt>
          <sourceDesc></sourceDesc>
     </fileDesc>
     <encodingDesc>
          <projectDesc></projectDesc>
          <samplingDecl></samplingDecl>
          <editorialDecl></editorialDecl>
          <tagsDecl></tagsDecl>
          <refsDecl></refsDecl>
          <classDecl></classDecl>
     </encodingDesc>
     <profileDesc>
          <creation></creation>
          <langUsage></langUsage>
          <particDesc></particDesc>
          <settingDesc></settingDesc>
          <textClass></textClass>
     </profileDesc>
     <revisionDesc>
          <change>
               <date></date>
               <respStmt></respStmt>
               <para></para>
          </change>
     </revisionDesc>
</teiHeader>
```

Figure 4.5: Header in BNC

```
<head></head>
<s></s>
<w></w>
<c></c>
<text></text>
<stext></stext>
<p></p>
```

Figure 4.6: Textual mark-up in BNC

```
<body>
<w AT> The </w> <w NP1> Sheikh </w>
<w NN1> ’ </w> <w FO> s </w> <w NN1> moustache </w>
<w VBDZ> was </w> <w ICS> like </w> <w AT1> a </w>
<w NN1> thistle </w> <w II> in </w> <w APPG> his </w>
<w NN1> bed </w> <w YC> , </w> <w AT> no </w>
<w NN1> matter </w> <w RRQ> how </w>
<w PPHS1> he </w> <w VVD> tossed </w>
<w CC> and </w> <w VVD> turned </w> <w YC> , </w>
<w PPHS1> he </w> <w RR> only </w>
<w VVD> rolled </w> <w AT> the </w> <w DAR> more </w>
<w II> on </w> <w II> to </w> <w PPH1> it </w>
<w YF> . </w> <w RR> Yet </w> <w PPHS1> he </w>
<w VBDZ> was </w> <w II> at </w> <w AT> the </w>
<w NN1> height </w> <w IO> of </w> <w APPG> his </w>
<w NN1> power </w> <w YF> . </w>
</body>
```

Figure 4.7: Part-of-speech tagging in TEC

spoken texts and paragraphs. There are, however, other elements and attributes which describe the header and the markup of the BNC in greater detail[5]. In the BNC, all special characters are presented by SGML entity references, which take the form of an ampersand followed by a mnemonic for the character and terminated by a semicolon (e.g., the representation *été* for the word *été*). In TEC, the transcription of these characters had to be carried out manually to make the data processable for annotation and querying tools. The use of international standards for the specification of application-independent document grammars (such as TEI for headers or SGML for mark-up) makes the corpora processable for computers as well as exchangeable and usable for researchers.

Since TEC had not been tagged before, the part-of-speech tags produced by TnT were added to the corpus. The BNC was tagged according to the CLAWS tagging scheme (cf. [Garside, 1987]). However, in order to make the corpora as comparable as possible in terms of their linguistic interpretation, the CLAWS tags were removed and the BNC was tagged once again using TnT. For the representation of the part-of-speech tags in the comparable corpus, the vertical TSV output format (cf. Section 4.2.3) was transformed into a horizontal format and added to the text, as can be seen from Figure 4.7. In the horizontal tagging format, the part-of-speech tags

[5]More information on the BNC can be found under the following URL: http://www.hcu.ox.ac.uk/BNC/.

are encoded as attributes of tokens.

4.2.5 Corpus querying

In order to find particular kinds of linguistic information in the corpus annotated in the ways described above, tools for querying the corpus for the features annotated are needed. For this purpose, the IMS Corpus Workbench [Christ, 1994] can be used. This concordance tool, with which it is possible to query for words and/or part-of-speech tags on the basis of regular expressions, consists of two modules: the corpus query processor (CQP) and the user interface (Xkwic). Importing TnT output to the workbench

Figure 4.8: Passive query with Xkwic (monolingual)

is a straightforward step since the preparatory steps were all carried out by the part-of-speech tagger TnT (cf. Section 4.2.3). These steps include character set normalization, tokenization and sentence boundary detection, such that the input format of the tagged corpus is TSV. The corpus is then encoded in such a way that it can be queried by the system. After the encoding of the corpus, all attributes (words, part-of-speech tags etc.) are declared in a registry file, which is a crucial element for all operations of the corpus maintenance.

The required information can then be queried using CQP, which implements a query language on the basis of the following regular expressions:

concatenation, disjunction, negation, Kleene star, the plus and the interval operator. The results of the query are displayed using Xkwic. Figure 4.8 shows Xkwic with a query for passive. The query is based on the part-of-speech tags VB.* (forms of the verb *be*) followed by VVN.* (past participle) and zero to one words in between. The results are displayed in the KWIC (keyword in context) list indicating the number of matches (see Figure 4.8). Xkwic offers the usual functionalities of a concordance program: concordances can, for example, be saved (all at once or in different sub-corpora), deleted, sorted and printed. The query history can be viewed, saved and loaded, and sub-corpora can be saved as well. Furthermore, an extended view on the KWIC concordances, a window for messages and an alignment window for parallel concordances can be displayed, too. Xkwic also calculates a frequency distribution for the first word or part-of-speech tag of the matches. Figure 4.9 shows a frequency distribution for the passive query presented in Figure 4.8. The information thus provided makes it possible to analyze, for example, the use of aspect within passive constructions.
As to the query of instances of complex syntactic constructions, typically several different queries need to be made to obtain all (satisfactory recall) and only the relevant matches (satisfactory precision). In some cases irrelevant matches have to be removed from the list manually. In Appendix 7.3, the different queries and sample concordances (taken from TEC) for the typical and untypical fiction features introduced in Section 4.2.1 are presented. There are three groups of queries: queries based on words, queries based on part-of-speech tags and queries based on both words and part-of speech tags. It can be seen in Appendix 7.3 that the queries can become quite complex (as in the case of final prepositions)[6]. Since the two comparable fiction parts in TEC and BNC analyzed in this thesis are both in English, the queries can be used on both sub-corpora. This procedure ensures consistent querying even in cases where the values for precision and recall do not attain 100%[7].

[6] As mentioned in Section 4.2.2, the part-of-speech tags used for English are based on the Susanne tagset (cf. [Sampson, 1995]). Refer to Appendix 7.2 for further information on the tagset.

[7] Precision and recall have not been counted for the queries, since the corpus under investigation in this section is too large to be evaluated manually. Another possibility to count precision and recall would have been to extract a small sub-corpus, on the basis of which the queries could have been evaluated. However, since some of the features occurred rather rarely, the number of matches would not have been representative in a sub-corpus of a small size.

Figure 4.9: Frequency distribution for passive in Xkwic

4.2.6 Significance test

In order to attain comparable results, all frequencies of occurring features were normed using the following formula:

$$\frac{frequency\ of\ feature\ occurrences\ x\ basis\ of\ norming}{word\ count} = normed\ frequency$$

The basis of norming is 5 million for TEC and for BNC, this being the approximate size of the two sub-corpora.

As only the statistically significant results are relevant, all the frequencies of occurring features have to be subjected to a statistical test. The chi square test (cf. [Oakes, 1998]) was employed in this thesis. This test is a non-parametric statistical procedure for testing whether or not the distribution of feature occurrences is accidental. The first step of the test is to define the null hypothesis, according to which there is no statistically significant difference between the frequencies of feature occurrences found in the sub-corpora. If so, all the frequencies of occurrences would be the same as the sum of all frequencies divided by the number of categories. This value, referred to as the expected value (E), is determined using the

following formula:

$$\frac{raw\ x\ column\ total}{grand\ total\ of\ items} = E$$

In contrast to the expected value, the actual frequency is called the observed value (O). With these two values the chi square value can be calculated as follows:

$$\sum \frac{(O-E)^2}{E} = X^2$$

For the significance tests carried out in this thesis, the two-tailed/non-directional chi square test was employed and, additionally, Yates's correction was calculated given that the amount of data under investigation is rather large. For the same reason, the level of significance was set to 0.001, which means that the results have to reach 99.9% to be statistically significant.

4.3 Cross-linguistic approach

This section presents an approach to the cross-linguistic analysis and description of metaphoricity and ambiguity in TL texts compared to SL texts. This methodology includes the description of the theoretical basis which provides the foundation of this approach (cf. Section 4.3.1) and of the corpus design of the parallel corpora used for the analysis of metaphoricity and ambiguity (cf. Section 4.3.2). Moreover, it provides insight into the linguistic annotation needed for this multilingual investigation (cf. Section 4.3.3), the meta-linguistic mark-up (cf. Section 4.3.4) and the different methods of corpus querying for exploiting the multilingual corpus (cf. Section 4.3.5).

4.3.1 Analysis scenario

This section deals with the comparison of English translations and their German and French SL texts on the basis of the changes in grammatical metaphoricity as well as structural and cohesive dis-/ambiguity. For this reason, the analysis of metaphoricity and ambiguity is based on aligned German-English and French-English parallel corpora and their bilingual concordances of typical and untypical fiction features. These concordances

have to be annotated in view of the translation techniques employed for these features. Yet, since the translation techniques introduced in Section 2.3.2 are not sufficiently comprehensive for describing the changes between the TL and the SL in detail, the notions of grammatical metaphor and texture, as introduced in Sections 2.3.3 and 3.3, are used to explain the relation of SL and TL texts. Thus, the following categories for the description of changes between SL and TL sentences are of importance:

- de-metaphorization,
- metaphorization,
- re-metaphorization,
- disambiguation-cohesion,
- ambiguation-cohesion,
- disambiguation-structure,
- ambiguation-structure.

This categorization indicates whether a TL construction is metaphorized or de-metaphorized compared to the SL construction or whether it has the same degree of grammatical metaphor (using Steiner's multilingual classification of grammatical metaphor (cf. [Steiner, 2001b]) introduced in Section 2.3.3). Besides grammatical metaphor, textual cohesion and textual structure are analyzed, with a view to determining whether or not one of both is disambiguated or ambiguated. The translation techniques are then counted with regard to the typical or untypical fiction features in which they are realized. This serves as a basis for the explanation of instances of normalization and anti-normalization which are investigated on the basis of the empirical analysis.

4.3.2 Corpus design

Since the analysis scenario is mainly based on Steiner's multilingual classification of grammatical metaphor (cf. [Steiner, 2001b]), which is used to describe how SL texts and the corresponding TL texts are related, the corpus under investigation in this section is a parallel corpus. Here, the corpus

compiled consists of German and French original texts and their translations into English. Since the automatic analysis presented in Section 4.2.3 is not applicable in this context, the analysis has to be carried out manually. For this reason, the parallel corpus is compiled so as to keep its size operationalizable: the German-English translation consists of 4,033 words, the French-English translation of 4,034 words, the German original contains 6,237 words and the French original 4,196 words (18,500 words in total). The sub-corpus of English translations is extracted from the fiction part of TEC, i.e. the German and French originals are fiction texts, too[8]. As can be seen, the different sub-corpora are intended to be as comparable as possible. Since this analysis describes the relationship between SL and TL texts, the comparable corpus of English originals is not taken into account.

Figure 4.10: Corpus alignment with Déjà Vu

4.3.3 Corpus annotation

As part of the analysis relating to translations and their SL originals, the first step in linguistic annotation is to identify the translation units, which means that the parallel corpus needs to be aligned. For this purpose, the alignment program of the translation memory Déjà Vu is used (cf. [Atril,

[8]The texts of the English sub-corpus are taken from the fn000019.txt and fn000076.txt files (see Appendix 7.1) and the corresponding German and French originals. Information on the English, German and French texts used in the parallel corpus (e.g., author, title, translator etc.) can be found in Appendix 7.1.

2000]). With Déjà Vu, the translation units are defined by sentence boundaries, the text and its translation are aligned sentence by sentence. Figure 4.10 shows the alignment of the French-German parallel corpus performed with this tool.

Figure 4.11: Déjà Vu sample output of the French-English parallel corpus

Déjà Vu makes interactive corpus alignment possible, enabling the user to join, split and erase the SL and the TL sentences manually. After the alignment process, the aligned texts can be stored in one file or in two separate files, depending on the requirements of the querying tool used later in the analysis. Files can be exported to translation workbenches and to Microsoft Excel and Access. Figure 4.11 shows a Déjà Vu sample output of the French-German parallel corpus in TSV format.

If more abstract features are to be coded (as the translation techniques introduced in Section 4.3.1), annotation has to be carried out manually. Tools for such an annotation support the definition of annotation schemes and manual coding with the help of graphical user interfaces (GUIs). One such tool is Coder (cf. [O'Donnell, 1995]). Coder has five functions: chunking-up texts into units for coding (see Figure 4.12), defining coding schemes (see Figure 4.13), annotating texts with the pre-defined coding scheme (see Figure 4.14), calculating basic descriptive statistics for a coded corpus and outputting concordances (for the last two described functionalities see Section 4.3.5). The chunking mechanism divides the text into sentences and paragraphs. The annotation of units smaller than sentences

Figure 4.12: Corpus segmentation with Coder

Figure 4.13: Definition of an annotation scheme in Coder

requires manual chunking. From Figure 4.12, it can be seen that the segmentation needed to annotate the parallel corpus is based on the translation unit. Each English concordance line which includes a typical or untypical feature of fiction (delimited by angle brackets) and the respective German or French equivalent form one segment. Within such a segment, the SL text comes first and is followed by the TL text, separated by dashes. The definition of a coding scheme is supported by another GUI, additions and changes to a scheme are straightforward. Figure 4.13 shows that the annotation scheme used to annotate translation techniques is based on a system network called *tt* (for *translation technique*). This system network is split up into two sub-systems for metaphoricity, on the one hand, and texture, on the other. The *metaphoricity* system includes the features *re-metaphorization*, *metaphorization* and *de-metaphorization* (cf. Section 4.3.1). The *texture* system is composed of the *cohesion* sub-system, including the features *ambiguation-cohesion* and *disambiguation-cohesion*,

and the *structure* sub-system, comprising the features *ambiguation-structure* and *disambiguation-structure* (cf. Section 4.3.1). Coding is supported by another GUI that highlights the unit being coded at a time and presents the coding options. Figure 4.14 exemplifies the annotation of a segment for which the *tt* system and the *metaphoricity* system have already been selected and the feature *de-metaphorization* is being chosen. Texts annotated with the help of Coder are represented in SGML format, as shown in Figure 4.15, where the actual annotation choice is the value of the element *segment features*.

Figure 4.14: Corpus annotation with Coder

Figure 4.15: Coder sample output of the German-English parallel corpus

4.3.4 Corpus representation

With a view to maintaining the multilingual corpus, each text is encoded in terms of a header that provides meta-information (referring to the title, au-

thor, publication, translator, SL text etc.), register information (field, tenor, mode) as well as an abstract. The header developed for the multilingual corpus is a combination of the TEC header and the BNC header introduced in Section 4.3.4. Each file is encoded in the extensible Mark-up Language (XML), using a modified version of TEI (see Figure 4.16) and employing a standard XML editor[9]. The text body is provided with annotations containing information on headings, sentences, paragraphs etc. (in an XML modification of the mark-up introduced in Section 4.2.4).

As to linguistic annotation, a representation format is required for the body of each text with which different kinds of corpus annotation (here: part-of-speech, typical and untypical fiction features, translation techniques and alignment information) can be displayed and queried together. XML was chosen for this purpose (see Figure 4.17). In this integrated XML corpus representation, each *segment (seg)* is assigned its own *identification number (id)* as well as an attribute *translation technique (tt)*, where the changes between the SL and TL text are annotated. The element *language (lang)* encodes the TL or the SL in its attribute. In the case of the SL text, the sentence is not annotated in greater detail since part-of-speech tagging is only necessary for the querying of the English translations. As to the TL text, the elements *translation unit (tu)*, in which the typical and untypical fiction features are identified, and *token*, in which the part-of-speech tags are included, are annotated. The different units of annotation are defined in a document type definition (DTD), i.e. a formal annotation grammar (see Figure 4.18). The well-formedness and validity of the integrated XML-encoded corpus can be checked against the DTD with the help of an XML editor[10]. In the DTD presented in Figure 4.18, the units of annotation are specified as elements (*body*, *segment (seg)*, *language (lang)*, *translation unit (tu)* and *token*). Since XML obeys a hierarchical order, each element consists of other elements or of parsable character data, that is the raw text[11]. Thus, the *body* consists of *segments (seg)* and the *segments (seg)* of *language (lang)* elements. The *language (lang)* elements are composed of *translation*

[9]Here, XML Spy has been employed (URL: http://www.xml-spy.com).

[10]Here, again, XML Spy can be employed (URL: http://www.xml-spy.com).

[11]Parsable character data (PCDATA) is opposed to character data (CDATA), which is not parsed in a well-formedness or validity check. The plus denotes *one or more occurrences of* and, additionally, it is indicated whether or not the attributes are required.

units (tu), *tokens* or PCDATA, the *translation units (tu)* of *tokens* and the *tokens* of PCDATA. The *segment (seg)* includes the *id* and the *translation technique (tt)* as attributes. In the latter, de-metaphorization, metaphorization, re-metaphorization, disambiguation-cohesion, ambiguation-cohesion, disambiguation-structure and ambiguation-structure (cf. Section 4.3.1) are encoded. The attributes of the element *language (lang)* indicate the SL and TL. In the element *translation unit (tu)*, the typical and untypical fiction features are encoded as attributes, whereas *part-of-speech (pos)* tags, in the form of attributes, are assigned to the element *token*. The automatic generation of new DTDs makes updating the DTD a straightforward task as more data are annotated.

4.3.5 Corpus querying

For extracting parallel text instances tagged with TnT (cf. Section 4.2.3), which serves as basis for the more abstract annotation explained in Section 4.3.3, querying tools like the IMS Corpus Workbench (cf. [Christ, 1994]) can be employed. As described in Section 4.2.5, the corpus query processor (CQP) of this system makes it possible to query for words and/or annotation tags on the basis of regular expressions. The fact that its CQP caters for the querying of parallel corpora (i.e. aligned SL and TL texts) benefits multilingual application. For an example of a query executed on a parallel German-English corpus see Figure 4.19. The query discussed in Section 4.2.5 is implemented here. The matching English passive constructions are put in angle brackets and are followed by the German source.

For the extraction of abstract features annotated with Coder (e.g., translation techniques), the review (see Figure 4.20) and statistics functions (see Figure 4.21) of Coder can be used for further processing. Figure 4.20 illustrates a sample concordance of the feature *de-metaphorization*[12], and in Figure 4.21, a descriptive statistics of the corpus is shown. This descriptive statistics presents absolute figures and the corresponding percentages for the different sub-systems as well as the features.

In order to combine queries for lexico-grammatical features, as annotated with TnT, and queries for translation techniques, annotated with Coder,

[12]Note that the user has to scroll through the window in order to get the SL sentences displayed.

```
<tei.2>
     <teiHeader>
          <fileDesc>
               <filename>fn000020.txt</filename>
               <subcorpus>fiction (trans_en)</subcorpus>
               <language>English</language>
               <titleStmt>
                    <title>Infanta</title>
                    <author>
                         <name>J. M. Brownjohn</name>
                    </author>
               <titleStmt>
               <publicationStmt>
                    <publisher>Viking Press</publisher>
                    <pubPlace>Great Britain</pubPlace>
                    <date>1992</date>
               </publicationStmt>
               <translation>
                    <direction>German-English</direction>
               </translation>
               <sourceText>
                    <title>Infanta</title>
                    <language>German</language>
                    <author>
                         <name>B. Kirchhoff</name>
                    </author>
                    <publisher>Suhrkamp</publisher>
                    <pubPlace>Germany</pubPlace>
                    <date>1992</date>
               </sourceText>
               <registerAnalysis>
                    <register>fiction</register>
                    <field>
                         <experientialDomain>love story</experientialDomain>
                         <goalOrientation>story-telling</goalOrientation>
                         <socialActivity>narration</socialActivity>
                    </field>
                    <tenor>
                         <agentiveRole>narrator to reader</agentiveRole>
                         <socialRole>equal</socialRole>
                         <socialDistance>minimal</socialDistance>
                    </tenor>
                    <mode>
                         <languageRole>constitutive</languageRole>
                         <channel>graphic</channel>
                         <medium>written</medium>
                    </mode>
               </registerAnalysis>
          </fileDesc>
          <encodingDesc>Modified TEI</encodingDesc>
     <abstract>
     The center of attention is Kurt Lukas, a German fashion model
     who comes to rest on a suitably godforsaken Philippine island
     full of European priests. For reasons that are never fully
     hinted at -- much less made clear -- Kurt has decided to drop
     out of things for a while, and the mission fathers (who seem
     only marginally more balanced than he is) take him in gladly
     and do their best to keep him. They are helped along by Mayla,
     their orphaned housekeeper and surrogate daughter, who wastes
     no time in seducing Kurt with her virginity. The impending
     revolution that hovers in the background makes the seclusion
     of Kurt and Mayla's pastoral all the weirder--as does the
     fact that every single one of the priests is in love with her
     and makes a point of confessing his sexual transgressions to
     Kurt at the earliest opportunity.
     </abstract>
     </teiHeader>
</tei.2>
```

Figure 4.16: Sample XML header of the multilingual corpus

```
<body>
     <seg id="1" tt="de-met">
          <lang sl="ge">
          Wahrscheinlich kenne ich Mayla genauer als andere,
          weil ich, wie Mayla, nach meiner Geburt von einer
          Stadt in die naechste zog.
          </lang>
          <lang tl="en">
                    <token pos="CS">because</token>
                    <token pos="YC">,</token>
                    <token pos="ICS">like</token>
                    <token pos="APPG">her</token>
                    <token pos="YC">,</token>
                    <token pos="ICS">after</token>
                    <token pos="PPIS1">I</token>
               <tu feat="passive" status="untypical">
                    <token pos="VBDZ">was</token>
                    <token pos="VVN">born</token>
               </tu>
                    <token pos="YC">,</token>
                    <token pos="PPIS1">I</token>
                    <token pos="VVD">moved</token>
                    <token pos="II">from</token>
                    <token pos="MC1">one</token>
                    <token pos="NNL1">city</token>
                    <token pos="II">to</token>
          </lang>
     </seg>
     <seg id="2" tt="de-met">
          <lang sl="ge">
          Er verschwieg eine Schwerverletzte und den Tod eines
          Mannes; jede Hilfe fuer sie kaeme zu spaet, hatte
          ihm Romulus leise gemeldet.
          </lang>
          <lang tl="en">
                    <token pos="DD1">that</token>
                    <token pos="MC1">one</token>
                    <token pos="NN1">person</token>
                    <token pos="VHD">had</token>
               <tu feat="passive" status="untypical">
                    <token pos="VBN">been</token>
                    <token pos="VVN">killed</token>
               </tu>
                    <token pos="CC">and</token>
                    <token pos="DD1">another</token>
                    <token pos="RR">badly</token>
                    <token pos="VVN">wounded</token>
          </lang>
     </seg>
</body>
```

Figure 4.17: Integrated XML representation for multiply annotated corpus

```
<?xml version="1.0" encoding="UTF-8"?>
<!--DTD generated by XML Spy v4.3 (http://www.xmlspy.com)-->
<!ELEMENT body (seg+)>
<!ELEMENT seg (lang+)>
<!ATTLIST seg
     id ID #REQUIRED
     tt (de-met | met | re-met | disam-coh | am-coh |
     disam-str | am-str) #REQUIRED>
<!ELEMENT lang (tu | token | #PCDATA)+>
<!ATTLIST lang
     sl (de | fr | en) #REQUIRED
     tl (de | fr | en) #REQUIRED>
<!ELEMENT tu (token+)>
<!ATTLIST tu
     feat (past-tense-verbs | third-person-pronoun |
           perfect-aspect | public-verbs | synthetic-negation |
           present-participle-clauses | time-adverbials |
           place-adverbials | adverbs | phrasal-coordination |
           present-tense-verbs | attributive-adjectives |
           wh-relative-clauses | pied-piping |
           phrasal-coordination | nominalizations | conjuncts |
           passive | by-passives | past-participle-clauses |
           subordinators | that-clauses | demonstratives |
           final-prepositions | existential-there) #REQUIRED
     status (typical | untypical) #REQUIRED>
<!ELEMENT token (#PCDATA)>
<!ATTLIST token
     pos (APPG | CC | CS | DD1 | ICS | II | MC1 | NN1 |
     NNL1 | PPIS1 | RR | VBDZ | VBN | VHD | VVD | VVN |
     YC) #REQUIRED>
```

Figure 4.18: XML DTD for multi-layer multilingual corpus annotation

Figure 4.19: Passive query with CQP (multilingual)

Figure 4.20: Review function of Coder

(cf. Section 4.3.3), a tool is needed which allows querying the corpus with reference to the different layers of corpus annotation. Since the corpus is represented in XML and validated against a DTD (cf. Section 4.3.4), the MATE system with its query mechanism Q4M (cf. [Mengel, 1999] and [Mengel & Lezius, 2000]) can be used to fulfill this rather complex task. Figure 4.22 shows a query for the translation technique *de-metaphorization* for the untypical fiction feature *passive* on the basis of the TL. First of all, the elements which are to be included in the query have to be chosen (here: segment, language, translation unit and token). The next step is to define that segment governs language, language governs translation unit and translation unit governs token. Furthermore, the following requirements have to be met:

- the translation technique specified for the segment has to be de-metaphorization,
- the TL specified in language has to be English,
- the feature of the translation unit has to be passive,
- the part-of-speech tag has to be VVN for past participle.

The result is displayed in Figure 4.23, where de-metaphorization is found in the words “born” and “killed” (for explanations of these instances of de-

Figure 4.21: Statistics function of Coder

metaphorization see Examples 5.5 and 5.6 in Section 5.3.3)[13].
The fact that MATE does not display parallel concordances, i.e. it is not possible to view the corresponding SL construction, represents a shortcoming of this type of corpus querying. A solution to this problem can be found in Figure 4.24, where a query for those SL constructions is shown which triggered de-metaphorization in the translations. In this query, again, all the elements which are to be included in the query have to be selected (here: segment and language). Moreover, a further condition is that segment governs language. Additionally, the segment has to be de-metaphorization and the SL has to be German. The outcome is represented in Figure 4.25, in which the German sentences can be found which are de-metaphorized in the TL text.
Another problem with MATE is that it is not yet fully developed and thus not fully functional, i.e. its use in combination with large corpora is rather problematic. It can only be tested in connection with a very small sample corpus (as can be seen from Figures 4.23 and 4.25)[14].

[13]Note that the user has to scroll through the window to find the first match.
[14]Note that there is a follow-up project, called NITE, about to start. The aim of this project is to further develop MATE, so that a more fully functional system can be expected in due course (URL: http://mate.nis.sdu.dk/).

Figure 4.22: Query for de-metaphorization in passive with MATE (TL)

4.4 Psycholinguistic approach

This section presents the psycholinguistic approach of the present thesis and describes the methodology for investigating the cognitive processes taking place during translation. This methodology includes the theoretical framework for the analysis of the translation process (cf. Section 4.4.1). Furthermore, the experimental design used for the analysis is introduced in Section 4.4.2.

4.4.1 Analysis scenario

The psycholinguistic approach serves the explanation of the cognitive processes taking place during the translation process. Again, this section is based on the notion of grammatical metaphor introduced in Section 2.3.3. Starting from the multilingual classification of grammatical metaphor, [Steiner, 2001b] developed the following hypothesis (cf. [Steiner, 2001b, 12]):

> "Understanding involves, among other processes, the unpacking of grammatical metaphor, at least in many cases, and at least to a certain extent. At some point of 'depth of understanding', re-production in the target language sets in, and here the process of re-metaphorization is cut short below the degree to which it might otherwise go. The reasons could be language-

Figure 4.23: Results for de-metaphorization in passive with MATE (TL)

> specific (i.e. because of typological-contrastive properties of the languages involved), they could be register specific (i.e. in cases where the target language and context suggest a lower degree of metaphoricity), and/or they could have to do with a lack of effort or ability on the part of the translator – or, indeed, with some as yet unknown factor."

This hypothesis, summarized in Figure 4.26, can be applied to another phenomenon occurring in translation as well, namely the disambiguation of SL constructions in the TL. This means that understanding involves the unpacking of grammatical metaphor, on the one hand, and the disambiguation of ambiguous structures, on the other. Due to register-specific or typological reasons and depending on the individual translator and the translation process, the re-metaphorization of constructions is not necessarily completed in the course of the production of the TL text. Thus, the TL constructions remain disambiguated.

The psycholinguistic experiment, however, focuses on the notion of grammatical metaphoricity since clear experimental conditions have to be formulated, which is a rather difficult task in connection with the ambiguation hypothesis. Examples 4.1 and 4.2 (taken from [Steiner, 2001b, 9]) show the different degrees of grammatical metaphoricity, which serve as a basis for the formulation of the experimental conditions.

Figure 4.24: Query for de-metaphorization in passive with MATE (SL)

Figure 4.25: Results for de-metaphorization in passive with MATE (SL)

Figure 4.26: Incomplete re-metaphorization in TL

(4.1) a. A rising number of people smoke.

b. The number of people dying from lung cancer increases.

(4.2) a. A rising number of people smoke. As a consequence, the number of people dying from lung cancer increases.

b. Because a rising number of people smoke, the number of people dying from lung cancer increases.

c. The increase in smoking leads to an increase in death rates from lung cancer.

d. Increasing lung cancer death rates and the causally related increase in smoking ...

e. The cause of increasing lung cancer death rates in increased smoking ...

relator	(equal) so	a happened; so x happened
	(unequal) because	x happened, because a happened
minor process	because of	x happened because of a
process	cause	that a happened caused x to happen
quality	causal	happening a was (in a) causal (relation) to happening x
thing	cause	happening a was the cause of happening x

Table 4.2: Degrees of grammatical metaphor

These examples show that the relationship of any kind of meaning, can be rendered into different degrees of grammatical metaphor. Example 4.2 shows that the causal meaning implicit in Example 4.1 can be expressed in the form of a *relator*, a *minor process*, a *process*, a *quality* or a *thing*. This classification of the different degrees in which grammatical metaphor

can be expressed is summarized in Table 4.2. These different degrees of grammatical metaphor are helpful in view of formulating conditions in the psycholinguistic experiment which will be described in the following section.

4.4.2 Experimental design

Having analyzed the specific properties of translations with the help of product-oriented research, the psycholinguistic experiment is used to investigate the nature of translated text on a process-oriented basis. It combines the following methods in experimental psycholinguistics, which were introduced in Section 2.4, namely TAP, the sentence completion task, the gating task and the lexical decision task (see also Section 3.4). TAP offers the possibility of recording and transcribing the verbal part of the experiment. The sentence completion task allows the completion of text on a written basis, whereas the gating task is performed on a verbal basis. The lexical decision task provides the framework for the subject to take lexical decisions under time constraint. This combination of psycholinguistic methods has been chosen to focus more on the process of translation rather than on the product.

The experiment consists of two different parts, both of which include a German-English translation, on the one hand, and a French-English translation, on the other. Altogether, four different translation tasks have to be performed. The first part is a verbal translation task, where every sentence has to be read out, and its translation has to be verbalized immediately. The translator is not allowed any time for consideration. The sentences are to be translated in the given order (first, the German, then the French sentences). The second part consists of a written translation task, where the German and French sentences have to be translated in the given order. The translators are allowed to write down the translations without any time constraint and with the help of provided dictionaries (a German-English, a French-English, a monolingual German and a monolingual French dictionary). The idea behind posing both a verbal and a written translation task is to investigate whether the cut in re-metaphorization takes place earlier in the verbal translation or whether the lack of time causes the translator to directly transfer the structures. The verbal translation task, however, is not intended to be an interpreting task. The translator has to read aloud each

sentence, so that there is enough time to think about it. The only purpose of the verbalization is that the post-edition of the translations is prevented and the translation process is somehow cut.

The SL text consists of 5 different conditions, i.e. one for each degree of grammatical metaphor (see Table 4.2). Each condition is realized through 8 sample sentences. Since the sentences are presented in German and in French for the written and the verbal translation task, 4 x 40 sentences have to be translated (see Appendix 7.5). The sentences in block A of Appendix 7.5 represent the metaphorical degree of *relator*, B represents *minor process*, C stands for *process*, D means *quality* and E is *thing*. The important thing about the sample sentences is that they should not trigger typologically or register-specific translations, i.e. the realization in all degrees of grammatical metaphor should be possible for every sample sentence in the SL as well as in the TL. Thus, the conditions should be cross-lingually comparable and transferable in terms of the sample sentences. The sentences are presented to the translators in random order and together with an instruction sheet (see Appendix 7.4).

The translators start with the German verbal translation task, followed by the French verbal translation task, both under time constraint. This part of the experiment is recorded on tape. For the German and French written translation task, the subjects are allowed to produce translations under almost real working conditions (i.e. on a written basis, no immediate time limit and with the help of dictionaries). The translations should be as idiomatical as possible – the English audience should not recognize them as translations, but as texts originally produced in English. After the completion of the experiment, the recorded tape is transcribed and the degrees of grammatical metaphor in the SL and their realizations in the TL are determined. Errors or lacking words are not taken into account.

Only two translators participated in the experiment. The translators being English native speakers, they translated into their mother tongue. The first subject is a professional translator who teaches translation studies as well. In the following, he is referred to as *the teacher*. The second subject is a student in translation studies and is referred to as *the student*.

On the basis of this experimental design, the differences between the verbal and the written translation task, between the student and the teacher and between the German-English translation process and the French-English

translation process in terms of grammatical metaphorization can be investigated. This experiment was not conducted on an empirical basis with a statistically significant number of translators. It should rather be seen as a pilot study which allows to interpret some tendencies, but which does not claim to deliver statistically significant results.

4.5 Summary and conclusions

The aim of this chapter was to present the methodology which integrates the advantages of the different methods presented in Sections 2 and 3. This methodology includes the analysis of a large comparable corpus (see Section 4.2), the investigation of a smaller parallel corpus (see Section 4.3) and a psycholinguistic experiment (see Section 4.4). For each approach, the following sections were discussed: the different analysis scenarios (cf. Sections 4.2.1, 4.3.1 and 4.4.1), the different analysis designs (cf. Sections 4.2.2, 4.3.2 and 4.4.2), the linguistic annotation of the corpora (cf. Sections 4.2.3 and 4.3.3), their formal representations (cf. Sections 4.2.4 and 4.3.4), the different methods of corpus querying (cf. Sections 4.2.5 and 4.3.5) and a method of significance testing (cf. Section 4.2.6).

This chapter showed that an interdisciplinary methodology combining the empirical approach, the cross-linguistic and the psycholinguistic approach compensates the disadvantages and exploits the advantages of the methods introduced in Chapter 2 and assessed in Chapter 3. This means that the dependent and independent variables vary according to the different approaches (cf. Chapter 1). Not only does the combination of product- and process-oriented research allow the description of the translation-specific properties, it also makes it possible to investigate the causes underlying them (cf. Chapter 5).

Chapter 5

Analysis

5.1 Introduction

This chapter is concerned with the implementation of the methodology developed in Chapter 4. First, the empirical analysis is introduced (cf. Section 5.2). The second step in the interdisciplinary methodology for the investigation of translated text is the analysis of metaphoricity and ambiguity in SL and TL texts (cf. Section 5.3). Finally, Section 5.4 looks at the results of the psycholinguistic experiment.
For each analysis, the starting hypotheses (cf. Sections 5.2.1, 5.3.1 and 5.4.1), hypothesis testing (cf. Sections 5.2.2, 5.3.2 and 5.4.2) and the interpretation of results (cf. Sections 5.2.3, 5.3.3 and 5.4.3) are described.
Section 5.5 summarizes and interprets the results on the basis of the interdisciplinary analysis of the nature of translated text.

One advantage of combining empirical, cross-linguistic and psycholinguistic approaches is that the results of the three analyses cannot only be described but also explained: the cross-linguistic investigation clarifies the empirical approach because the influence of the SL on the translations is analyzed and the results of the cross-linguistic analysis can be explained with the help of the psycholinguistic approach because the cognitive processes during the translation process are investigated. Another novelty of this interdisciplinary methodology is that the specific properties of translations can be investigated on a product-oriented (in the course of the empirical and the cross-linguistic analyses) as well as a process-oriented (in the course of the psycholinguistic analysis) basis, which guarantees the description as well as

the explanation of the results gained from the investigation of the nature of translated text.

5.2 Empirical analysis

This section introduces the hypotheses which can be tested in the course of the empirical analysis (cf. Section 5.2.1). In Section 5.2.2 the results of the analysis are presented. Section 5.2.3 is concerned with the interpretation of the results.

5.2.1 Hypotheses

For the analysis of normalization and anti-normalization in the comparable corpus consisting of English translated fiction (TEC) and English original fiction (BNC), the following hypotheses are tested:

> (H1) In translations into English, one would expect that, if they show normalization, there will be a higher frequency of typical features in English translations than in English original texts.

> (H2) In translations into English, one would expect that, if they show normalization, there will be a lower frequency of untypical features in English translations than in English original texts.

The frequent use of typical register features is an indicator of normalization in translated texts. A comparison with English originals makes it possible to determine whether or not the use of fiction features in English translations obeys the usage norms (H1). Of course the same holds true for the untypical fiction features, so that, if the English translations show instances of normalization, the untypical fiction features are expected to be underused (H2).

> (H3) In translations into English, one would expect that, if they show anti-normalization, the frequen-

cy of typical features will be higher in English original texts than in English translations.

(H4) In translations into English, one would expect that, if they show anti-normalization, the frequency of untypical features will be higher in English translations than in English original texts.

If the English translations do not obey the English usage norms, the typical features are underused (H3), whereas the untypical features occur more frequently in the English translations than in the English originals (H4).

5.2.2 Testing of hypotheses

The normed frequencies for the typical features found in TEC and BNC are listed in Table 5.1. The frequencies are all statistically significant except for place adverbials, which are put in brackets. Table 5.2 contains the normed frequencies for the untypical features found in TEC and BNC. All frequencies are statistically significant. In Table 5.3, the differences between the frequencies found in TEC and BNC are calculated and normed (as percentages). Thus, the positive percentages refer to the degree of normalization in connection with the typical features, whereas the negative percentages relate to the degree of anti-normalization for the typical features. The degrees of anti-normalization (represented in the negative percentages) and normalization (reflected by the positive percentages for the untypical features) are summarized in Table 5.4. Figure 5.1 illustrates the results of Table 5.3, whereas the results of Table 5.4 can be found in Figure 5.2.

5.2.3 Interpretation of results

As regards the hypotheses, the higher frequency of typical fiction features to be found in TEC, as compared to the BNC, means that the translations show a trend towards normalization for typical fiction features, which supports (H1). Untypical fiction features occur more frequently in TEC than

	BNC	TEC
adverbs	269,430	306,471
present participles	30,046	41,520
perfect aspect	37,810	44,429
synthetic negation	13,903	17,631
time adverbials	25,668	29,145
(place adverbials)	34,918	35,142
third person pronouns	251,266	246,252
public verbs	37,129	29,225
past tense verbs	205,529	192,369

Table 5.1: Normed frequencies for typical features in the comparable corpus

	BNC	TEC
conjuncts	49,024	42,461
existential *there*	11,545	10,948
final prepositions	196	299
by-passives	3,204	4,701
past participles	2,082	3,730
pied piping	1,599	3,626
agentless passives	27,690	34,684
nominalizations	12,815	23,118
phrasal coordination	29,672	40,133
that-clauses	30,589	42,233
subordinations	40,705	54,608
wh-relative clauses	13,881	28,753
demonstratives	33,258	48,990
present tense verbs	12,861	39,749
attributive adjectives	117,827	158,934

Table 5.2: Normed frequencies for untypical features in the comparable corpus

in BNC, that is the translations show a trend towards anti-normalization for the untypical fiction features, which substantiates (H4). Example 5.1 is a sample concordance taken from TEC which exemplifies the normalized use of the typical fiction feature *place adverbial*.

(5.1) a. Marden is about to enter the house when he turns **around**: "You must leave the gun, naturally;

b. there wasn't enough air to breathe. After he went **back** outside, on the road, he noticed that

c. The bad weather will return. It was **there**, near the magazine (he can see it clearly from

present participles	38.19%
negation	26.81%
perfect aspect	17.51%
adverbs	13.75%
time adverbials	13.55%
(place adverbials)	0.64%
third person pronouns	-2.04%
past tense verbs	-6.84%
public verbs	-27.05%

Table 5.3: Degree of anti-/normalization for typical features

conjuncts	15.46%
existential *there*	5.45%
agentless passives	-25.26%
subordination	-34.16%
attributive adjectives	-34.39%
phrasal coordination	-35.26%
that-clauses	-38.07%
by-passives	-46.72%
demonstratives	-47.30%
final prepositions	-52.55%
past participles	-79.15%
nominalizations	-80.40%
wh-relative clauses	-107.14%
pied piping	-126.77%
present tense verbs	-209.07

Table 5.4: Degree of anti-/normalization for untypical features

Here, the *place adverbials* "around", "back" and "there" are instances of normalization, showing the overuse of this typical fiction feature. Example 5.2 contains another sample concordance from TEC and illustrates the anti-normalized use of the untypical fiction feature *by-passive*.

(5.2) a. asked when it was all going to start, he **was pacified by** the chairman of the discussion:

b. a black man's head horribly shrunken. Rebecca **was told all this and more by** Nell about a week after her arrival at Broom House.

c. because I could hear voices in the corridor and I **was distracted by** the noise of the lift. It didn't

In this example, the *by-passives* "was pacified by", "was told all this and more by" and "was distracted by" contribute to anti-normalization, meaning that this untypical fiction feature is overused. More instances of the use of typical and untypical fiction features can be found in Appendix 7.3.

Figure 5.1: Degree of anti-/normalization for typical features

In summary, it can be observed that the typical fiction features are normalized and the untypical fiction features are anti-normalized in English translated fiction compared to English original fiction. Thus, English translated fiction is more narrative and situation-dependent than English original fiction texts are (since an overuse of the typical features can be attested). At the same time, English translated fiction is more abstract and less edited than original fiction (since an overuse of the untypical features can be attested). This leads to the conclusion that translators tend to conform to the typical patterns of the TL in terms of typical fiction features, the use thereof being exaggerated, whereas they tend to use the untypical fiction features less often. These observations show that a register shift between English original narrative texts and English translated narrative texts takes place in the sense that, owing to the extensive overuse of typical fiction features, English translated narrative texts are even more typical of their register. However, they point towards a more neutral register through the extensive overuse of untypical fiction features. Reasons for normalization or anti-normalization regarding the typical and untypical fiction features are discussed in Section 5.3.

Figure 5.2: Degree of anti-/normalization for untypical features

5.3 Cross-linguistic analysis

This section is intended as an introduction of the hypotheses concerning metaphoricity and ambiguity on a cross-linguistic basis (cf. Section 5.3.1). In Section 5.3.2 the results of the analysis are presented. Section 5.3.3 is concerned with the interpretation of the results.

5.3.1 Hypotheses

For the analysis of grammatical metaphoricity and cohesive or structural ambiguity in the German-English and French-English parallel fiction corpus, the following hypotheses are tested:

(H5) In translations into English, one would expect that, if they show de-metaphorization, there will be a higher frequency of congruent units in English translated fiction features than in their German or French SL equivalents.

(H6) In translations into English, one would expect that, if they show re-metaphorization, there will be exactly the same degree of metaphoricity in the English translated fiction features compared to their German or French SL equivalents.

(H7) In translations into English, one would expect that, if they show metaphorization, the frequency of congruent units will be lower in the English translated fiction features compared to their German or French SL equivalents.

As regards grammatical metaphoricity, re-metaphorization can be found if the TL and SL texts show the same degree of metaphoricity encoded in the fiction features (H6). The occurrence of more congruent structures in the SL than in the TL texts encoded in the fiction features indicates metaphorization (H7). De-metaphorization is identified through the use of more congruent structures in the TL texts than in the SL texts encoded in the fiction features (H5).

(H8) In translations into English, one would expect that, if they show ambiguation in terms of cohesion, the cohesive means in the fiction features of the English translations will be more ambiguous than the cohesive means in the German and French SL equivalents.

(H9) In translations into English, one would expect that, if they show disambiguation in terms of cohesion, the cohesive means in the fiction features of the English translations will be less ambiguous than the cohesive means in the German and French SL equivalents.

Concerning ambiguation in terms of cohesion, the TL texts are more ambiguous than the SL texts in the encoded fiction features (H8), whereas disambiguation can be found if the TL fiction features are less ambiguous than their corresponding constructions in the SL as regards of the realization of their cohesive means (H9).

(H10) In translations into English, one would expect that, if they show ambiguation in terms of structure, the structural means in the fiction features of the

English translations will be more ambiguous than the structural means in the German and French SL equivalents.

(H11) In translations into English, one would expect that, if they show disambiguation in terms of structure, the structural means in the fiction features of the English translations will be less ambiguous than the structural means in the German and French SL equivalents.

Similarly to ambiguation and disambiguation regarding cohesion, ambiguation in relation to structure is indicated if the TL fiction features are more ambiguous in terms of structure than the corresponding constructions in the SL (H10). Whereas structurally disambiguated fiction features, as compared to the SL texts, imply disambiguation related to structure.

5.3.2 Testing of hypotheses

Table 5.5 contains the results concerning re-metaphorization, de-metaphorization and metaphorization as well as a co-occurring pattern of de-metaphorization and disambiguation in terms of cohesion for the German-English fiction features (normed as percentages). Table 5.6 presents the results regarding disambiguation in terms of cohesion and structure as well as ambiguation in terms of cohesion and structure for the German-English fiction features (normed as percentages). In Table 5.7, the results concerning re-metaphorization, de-metaphorization, metaphorization and a co-occurring pattern of de-metaphorization in combination with disambiguation in terms of cohesion are introduced for the French-English fiction features (normed as percentages). Table 5.8 shows the results concerning disambiguation in terms of cohesion and structure as well as ambiguation in terms of cohesion and structure for the French-English fiction features (normed as percentages).

In order to show the relationship between anti-/normalization and grammatical metaphoricity as well as dis-/ambiguity, an overview of the fiction features displayed in Tables 5.5 to 5.8 concerning their behavior regarding

normalization and anti-normalization found in Section 5.2 in combination with grammatical metaphoricity and dis-/ambiguation is displayed in Tables 5.9 and 5.10. In Table 5.9, the frequencies for the co-occurrence of anti-/normalization and grammatical metaphoricity are listed for the following German-English and French-English features:

- all fiction features,
- fiction features showing normalization,
- fiction features showing anti-normalization,
- typical fiction features showing normalization,
- untypical fiction features showing normalization,
- typical fiction features showing anti-normalization,
- untypical fiction features showing anti-normalization.

Table 5.10 contains the frequencies for the same features regarding the co-occurrence of anti-/normalization and dis-/ambiguation. Additionally, the co-occurrences of anti-/normalization and grammatical metaphoricity or dis-/ambiguation found in German-English and French-English fiction features are illustrated in Figure 5.3.

5.3.3 Interpretation of results

As to the hypotheses formulated on the subject of grammatical metaphoricity, (H6) is strongly supported because in 60% to 85% of all cases exactly the same degree of metaphoricity was found in the English translated fiction features compared to the German or French SL structures. (H6) is followed by (H5) and (H7), i.e. in second place de-metaphorization and in third place metaphorization was found. No hypothesis on ambiguity is supported since the figures and the differences found in view of structural and cohesive ambiguity and disambiguity are too low. Therefore, the results for

	re-meta.	de-meta.	meta.	de./dis.
present participles	5.88%	17.65%	76.47%	
negation	76.47%			
perfect aspect	55.17%	17.24%		
adverbs	67.90%	32.10%		
time adverbials	81.67%	16.67%	1.67%	
(place adverbials)	40.30%	49.25%	10.45%	
third person pronouns	75.12%			18.31%
past tense verbs	85.65%	7.14%		
public verbs	65.22%	8.70%		
conjuncts	82.35%			
existential *there*	100.00%			
agentless passives	18.52%	25.93%		
subordination	66.67%	33.33%		
attributive adjectives	65.08%	30.16%	4.76%	
phrasal coordination	69.57%	30.43		
that-clauses	73.68%	26.32%		
by-passives	100.00%			
demonstratives	20%			36%
final prepositions				
past participles		50%	50%	
nominalizations	75.86%		24.14%	
wh-relative clauses	62.50%			
pied piping				
present tense verbs	100.00%			

Table 5.5: Grammatical metaphoricity of the German-English fiction features

ambiguity are left aside in what follows.

As regards the co-occurrence of anti-/normalization and grammatical metaphoricity, it can be said that re-metaphorization is strongly correlated to the French fiction features which are anti-normalized, followed by the French fiction features which are normalized in second place, the German fiction features which are anti-normalized in third place and the German fiction features which are normalized in fourth place. This means that English fiction features are more influenced by the French SL text than the German SL text and that the anti-normalized fiction features translated from French and German into English are more influenced by the French and German SL texts than the normalized features translated from French and German into the English language. In the following, some characteristic examples of the use of metaphorization, de-metaphorization and re-metaphorization are discussed.

	disam.-coh.	am.-coh.	disam.-str.	am.-str.
present participles				
negation				23.58%
perfect aspect		27.59%		
adverbs				
time adverbials				
(place adverbials)				
third person pronouns	1.41%	5.16%		
past tense verbs		6.94%		
public verbs	26.09%			
conjuncts			17.65%	
existential *there*				
agentless passives				55.56%
subordination				
attributive adjectives				
phrasal coordination				
that-clauses				
by-passives				
demonstratives	40.00%	4.00%		
final prepositions				
past participles				
nominalizations				
wh-relative clauses	37.50%			
pied piping				
present tense verbs				

Table 5.6: Dis-/ambiguation of the German-English fiction features

(5.3) a. German: nicht mehr die einzige **Frau, die im Ort gefeiert wurde**.

b. English: was no longer the only **celebrity in town**.

Example 5.3 shows the metaphorization of the German relative clause "die im Ort gefeiert wurde" into the English NP "celebrity in town". This is an example in which metaphorization triggers the anti-normalized use of the untypical fiction feature *nominalization*.

(5.4) a. French: que l'élève **travailla dur** et qu'il était

b. English: that he was one of the **most diligent** students

Example 5.4 shows how the French VP "travailla dur" is metaphorized into the English adjective "diligent", illustrating that metaphorization causes,

	re-meta.	de-meta.	meta.	de./dis.
present participles	26.67%	6.67%	66.67%	
negation	90.00%			
perfect aspect	50.00%	4.17%		
adverbs	61.80%	38.20%		
time adverbials	72.00%	24.00%	4.00%	
(place adverbials)	18.92%	67.57%	13.51%	
third person pronouns	91.10%			4.71%
past tense verbs	93.38%	5.88%		
public verbs	95.45%			
conjuncts	100.00%			
existential *there*	100.00%			
agentless passives	42.31%	7.69%		
subordination	88.89%	11.11%		
attributive adjectives	73.74%	21.21%	5.05%	
phrasal coordination	100.00%			
that-clauses	100.00%			
by-passives	100.00%			
demonstratives	87.88%			6.06%
final prepositions				
past participles	66.67%	33.33%		
nominalizations	83.33%		16.67%	
wh-relative clauses	95.45%			
pied piping				
present tense verbs	90.00%	10.00%		

Table 5.7: Grammatical metaphoricity of the French-English fiction features

similarly to Example 5.3, the anti-normalized use of the untypical fiction feature *attributive adjective*.

(5.5) a. German: Er verschwieg eine **Schwerverletzte** und **den Tod eines Mannes**; jede Hilfe für sie käme zu spät

b. English: that one person **had been killed** and another **badly wounded**

In Example 5.5, the German NPs “Schwerverletzte” and “den Tod eines Mannes” are rendered into the English de-metaphorized subordinate clause “that one person had been killed and another badly wounded”. Here, de-metaphorization triggers the anti-normalization of the untypical fiction feature *agentless passive*.

	disam.-coh.	am.-coh.	disam.-str.	am.-str.
present participles				
negation				10.00%
perfect aspect		45.83%		
adverbs				
time adverbials				
(place adverbials)				
third person pronouns	2.62%	1.57%		
past tense verbs		1.47%		
public verbs	4.55%			
conjuncts				
existential *there*				
agentless passives				50.00%
subordination				
attributive adjectives				
phrasal coordination				
that-clauses				
by-passives				
demonstratives	6.06%			
final prepositions				
past participles				
nominalizations				
wh-relative clauses		4.55%		
pied piping				
present tense verbs				

Table 5.8: Dis-/ambiguation of the German-English fiction features

(5.6) a. French: **Au début**, on crut à une bouderie d'enfant.

b. English: **When they started**, it was thought to be a fit of childish

In Example 5.6, the French PP "Au début" is de-metaphorized into the English subordinate clause "When they started", which shows that de-metaphorization can cause the anti-normalized use of the typical fiction feature *past tense verb*.
In the following, examples of re-metaphorization in combination with German-English and French-English anti-/normalization are discussed.

(5.7) a. German: **Dann** sah er den Blick der Frau, und alles Menschliche wich einem Chaos.

b. English: in this monstrous way. **Then** he saw her expression,

	re-meta.	de-meta.	meta.	de./dis.
all-ge	63.07%	15.68%	7.61%	2.47%
norm.-ge	61.49%	16.53%	9.84%	2.03%
anti-norm.-ge	64.17%	15.09%	6.07%	2.77%
norm.-typ.-ge	54.57%	22.15%	14.77%	
norm.-untyp.-ge	75.33%	5.25%		6.10%
anti-norm.-typ.-ge	91.18%			
anti-norm.-untyp.-ge	59.26%	17.83%	7.17%	3.27%
all-fr	78.53%	10.45%	4.81%	0.49%
norm.-fr	66.59%	16.28%	9.35%	0.52%
anti-norm.-fr	86.79%	6.41%	1.67%	0.47%
norm.-typ.-fr	53.23%	23.44%	14.03%	
norm.-untyp.-fr	93.31%	1.96%		1.57%
anti-norm.-typ.-fr	100.00%			
anti-norm.-untyp.-fr	84.39%	7.58%	1.97%	0.55%

Table 5.9: Co-occurrence of anti-/normalization and grammatical metaphoricity

(5.8) a. French: puis, **tirant** de sa ceinture un couteau, il se mit à cueillir les plus belles fleurs

b. English: coming into flower; then, **taking** his knife from his belt,

(5.9) a. German: 'Es ist der einzige Ort, an dem wir **ein anständiges Frühstück** bekommen.

b. English: place where we'll get **a decent breakfast**. 'He shouted orders

(5.10) a. French: De plus, elle disposait d'un vaste terrain clos où les élèves pouvaient se récréer, et où, des années plus tard, **seraient construits**, pour des raisons fort louables - la venue d'un bon millier d'élèves -, des bâtiments bien moins coquets, hélas.

b. English: new buildings **would be built**, which were, alas,

In Example 5.7, re-metaphorization causes normalization for the typical fiction feature *time adverbial* in the German-English sub-corpus (using the same constructions in German ("Dann") and English ("Then")). In Example 5.8, the same phenomenon is shown in connection with the typical

	disam.-coh.	am.-coh.	disam.-str.	am.-str.
all-ge	4.77%	1.99%	0.80%	3.60%
norm.-ge	3.06%	4.41%		2.62%
anti-norm.-ge	5.96%	0.31%	1.36%	4.27%
norm.-typ.-ge		4.60%		3.93%
norm.-untyp.-ge	9.17%	4.03%		
anti-norm.-typ.-ge			8.83%	
anti-norm.-untyp.-ge	7.05%	0.36%		5.05%
all-fr	0.60%	2.64%		2.73%
norm.-fr	0.80%	5.43%		1.11%
anti-norm.-fr	0.47%	0.35%		3.85%
norm.-typ.-fr		7.64%		1.67%
norm.-untyp.-fr	2.39%	1.01%		
anti-norm.-typ.-fr				
anti-norm.-untyp.-fr	0.55%	0.41%		4.55%

Table 5.10: Co-occurrence of anti-/normalization and dis-/ambiguation

fiction feature *present participle* in French ("tirant") and English ("taking"). Example 5.9 illustrates that re-metaphorization triggers anti-normalization for the untypical fiction feature *attributive adjective* in the German-English translations (the same constructions being used in German ("ein anständiges Frühstück") and English ("a decent breakfast")). In Example 5.10, the same phenomenon is again illustrated in connection with the untypical fiction feature *agentless passive* for French ("seraient construits") and English ("would be built").
From Examples 5.7 to 5.10, it can be seen how the choice of TL constructions is influenced by the constructions of the SL texts. Another advantage of this kind of analysis is that the differences due to register or typology can be identified (see Examples 5.11 and 5.12).

(5.11) a. German: Während **seiner Wanderung** durch den Wald dachte er an

b. English: **he walked** through the forest and thought about

(5.12) a. German: Die Wahrheit über das dunkle Geheimnis fand man damals **in ihrem Tagebuch**.

b. English: At that time, **her diary** explained the dark secret

Figure 5.3: Co-occurrence of anti-/normalization and grammatical metaphoricity or dis-/ambiguation

Example 5.11, which is an example of de-metaphorization into the typical fiction feature *past tense verb*, shows that the German nominalization "Wanderung" is translated into the English verb "walked". *Nominalization* being a feature untypical of English fiction, it was transferred into the typical fiction feature *past tense verbs*. A similar phenomenon can be found in Example 5.12, which exemplifies re-metaphorization of the typical fiction feature *past tense verb*. Here, the typical fiction feature *past tense verb* is used both in the English and the German sentence, the semantic roles, however, are distributed differently as the German typology allows the frequent use of constructions like the impersonal passive alternative "man" (here: in combination with the PP "in ihrem Tagebuch"), whereas the English typology provides other lexico-grammatical means of expressing the same meaning. In this case, the non-agentive NP subject "her diary" is used.
Those properties of translated text which are not register- or typology-

related are inherent in the translation process itself. Evidence of this is contained in Example 5.3, in which the German relative clause "die im Ort gefeiert wurde" is rendered into the English NP "celebrity in town". The English translation is less explicit than the SL equivalent because the fact that a woman is the only celebrity in town (and not a man) is implied by the context and not spelled out in the English text. This example clearly shows how metaphorization entails a loss of explicitation. In contrast to this, Example 5.6 illustrates the co-occurrence of de-metaphorization with simplification as well as explicitation (cf. [Baker, 1996]). In this case, the French PP "Au début" is translated with the English subordinate clause "When they started", which is a de-metaphorized variant. Furthermore, the use of the subordinate clause makes the English translation easier to read and more explicit, spelling out the meaning of the French PP. Example 5.8 indicates how re-metaphorization goes hand in hand with shining-through (cf. [Teich, 2001a]) since a typical construction of the SL (i.e. the French present participle "tirant") can be found in the English translation (i.e. the English present participle "taking").
These observations show that metaphorization co-occurs with a loss of explicitation, whereas de-metaphorization tends to co-occur with explicitation and simplification. Re-metaphorization, however, can be triggered through shining-through.

To summarize these findings, it can be attested that anti-normalization and normalization are, first of all, due to re-metaphorization and, in second and third place, due to de-metaphorization and metaphorization. The sources of re-metaphorization, de-metaphorization and metaphorization can partially be found in register- and typology-specific language use. The translation process, which triggers not only normalization and anti-normalization, but also other translation properties such as explicitation, simplification and shining-through, can be seen as a further explanation. This source of explanation, which takes into account the processes inherent in translation, is discussed in greater detail in the following section.

5.4 Psycholinguistic analysis

This section introduces the hypotheses tested on the basis of a psycholinguistic experiment (cf. Section 5.4.1). In Section 5.4.2 the results of the analysis are presented. Section 5.4.3 is concerned with the interpretation of the results.

5.4.1 Hypotheses

The following hypotheses are tested with a view to the analysis of the different degrees of grammatical metaphor in the psycholinguistic experiment:

> (H12) In translations from German or French into English, one would expect that the frequency of (re-)metaphorization will be higher in the written translation task than in the verbal translation task.

This hypothesis is based on Steiner's assumption according to which understanding within the translation process means the unpacking of grammatical metaphor and re-metaphorization is not fully completed during the production of the TL text (cf. [Steiner, 2001b]). Hence, more instances of re-metaphorization or even metaphorization can be expected to occur in the written translations compared to the verbal translations, given that the translators have more time for re-metaphorization or even metaphorization (H12).

> (H13) In translations from German or French into English, one would expect that the frequency of (re-)metaphorization will be higher in the teacher's translation than in the student's translation.

Assuming that the teacher has more experience in translation and thus translates faster than the student, the number of occurrences of re-metaphorization or even metaphorization is expected to be higher in the translations produced by the teacher (verbal and written) than in the translations

produced by the student (verbal and written), since, with the teacher, the understanding and production processes take less time (H13).

> (H14) In translations from German or French into English, one would expect that the frequency of re-metaphorization will be higher in the German-English than in the French-English translations.

Since German and English belong to the same language family, presumably, they are more similar to each other than a Germanic and a Romance language. Consequently, there should be relatively more occurrences of re-metaphorization in the German-English translations than in the French-English translations (H14)[1]. This assumption is based on the different types of grammatical metaphoricity as a classificatory feature between languages: it can be expected that the more commonalities two different languages share, the higher the degree of re-metaphorization.

5.4.2 Testing of hypotheses

Table 5.11 summarizes the results of the psycholinguistic experiment for the teacher. They are split up into the verbal and written translation tasks for the German-English and French English translations. The features *de-metaphorization*, *re-metaphorization* and *metaphorization* were counted. Table 5.12 contains the results of the psycholinguistic experiment for the student. Again, the results are split up into the verbal and written translation tasks for the German-English and French-English translations. The features *de-metaphorization*, *re-metaphorization* and *metaphorization* were also counted for the student. Figures 5.4 and 5.5 illustrate the results for the teacher and the student. The German-English and French-English verbal and written translations produced by the teacher and the student can be found in Appendix 7.6.

[1]Note that this hypothesis can be formulated although the experimental design is intended to prevent typology-specific language use in translations (through the formulation of cross-lingually comparable and transferable conditions, as explained in Section 4.4.2). However, since the SL as well as the relation between the SL and the TL also have an influence on the translation process, the testing of the hypothesis formulated in (H14) also makes sense.

		de-meta.	re-meta.	meta.
GE-EN	verbal	10.00%	77.50%	12.50%
	written	12.50%	85.00%	2.50%
FR-EN	verbal	2.50%	95.00%	2.50%
	written	15.00%	82.50%	2.50%

Table 5.11: Results for the teacher

		de-meta.	re-meta.	meta.
GE-EN	verbal	42.50%	57.50%	
	written	25.00%	72.50%	2.50%
FR-EN	verbal	15.00%	85.00%	
	written	30.00%	67.50%	2.50%

Table 5.12: Results for the student

5.4.3 Interpretation of results

Concerning the hypotheses, it can be said that (H12) does not apply to the teacher at all since the degree of de-metaphorization is higher in the written translation than in the verbal translation for German-English as well as French-English. However, the difference for German-English is quite small and can be neglected. The results gained from the translations produced by the student support (H12) for German-English as there is a high degree of de-metaphorization in the verbal translation task. Yet, (H12) is refuted for French-English since a higher degree of de-metaphorization was found in the written translation than in the verbal one. The data regarding metaphorization obtained from the German-English and French-English verbal translation task supports (H13). This, in turn, backs up the presumption that, thanks to his experience, it takes the teacher less time to deal with the translation task. The few cases of metaphorization produced by the student are contained in the written translation, which supports the presumption according to which the student requires more time to metaphorize a given construction. Concerning the teacher, the results for metaphorization are higher in the verbal translations than in the written ones (especially for German-English). This confirms the assumption that experienced translators directly transfer structures into the TL without unpacking them. (H13) is also supported in view of re-metaphorization since, compared to the student's performance, the degree of re-metaphorization is higher in connection with all the tasks the teacher performed. This result substantiates the as-

Figure 5.4: Results for the teacher

sumption according to which the student unpacks grammatical metaphors during the understanding process, which he fails to re-pack completely in the course of the production process. (H14) is refuted since the average degree of re-metaphorization is higher for the teacher's and the student's French-English translations than the German-English ones. The reason for this phenomenon could be that the language contact between English and French in the past resulted in commonalities of the two languages. For example, French word order is similar to English word order (both obey SVO), which probably triggers many instances of re-metaphorization by allowing the translator to use the surface structure as transfer level without the semantic unpacking of information. This observation suggests that the traditional notion of language family is motivated historically rather than typologically.

The following examples provide an outline of how the student and the teacher de-metaphorized, metaphorized or re-metaphorized SL structures into the TL.

(5.13) a. German: Das Auftreten der Prinzen steht **in temporalem Zusammenhang** mit dem Schrei der Prinzessin.

b. English: **At the same time** the prince entered, there was a cry by the princess.

Figure 5.5: Results for the student

(5.14) a. German: Die Erläuterungen des Vaters sind **hilfreich** für das Auffinden des Kindes.

b. English: The father's statements was **assistance** in finding the child.

(5.15) a. French: La mauvaise performance da la voiture **cause** le terrible accident.

b. English: The car's bad performance **causes** the terrible accident.

(5.16) a. French: L'initiative de la mère avait **une influence négative** pour le développement du fils.

b. English: The mother's initiative had **a negative influence** on the son's development.

Example 5.13 is a typical example of the student's tendency to de-metaphorize in the German-English verbal translation task since the German *quality* "in temporalem Zusammenhang" is translated into the English *relator* "at the same time". This translation exemplifies the co-occurrence of de-metaphorization and explicitation as well as simplification. The English *relator* is easier to understand than the German *quality* and explicitly spells

out the meaning conveyed by the German *quality*. Example 5.14 provides evidence of the teacher's tendency to metaphorize in the German-English verbal translation, the German *quality* "hilfreich" being translated into the English *thing* "assistance". Examples 5.15 and 5.16 are typical examples of the student's and the teacher's use of re-metaphorization, which occurs particularly frequently in connection with the French-English verbal translation task: the student renders the French *process* "cause" into the English *process* "causes". The teacher translates the French *quality* "une influence négative" into the English *quality* "a negative influence". More examples of the teacher's and the student's translations can be found in Appendix 7.6.

In summary, it can be said that the translation process has a strong impact on the properties of translated text: depending on the ability of the translator and the external constraints imposed on the translation process, information is, partly or completely, unpacked in the understanding process and, partly or completely, unpacked in the production process. As a result, the translations can, for instance, be more or less normalized, easier or harder to read and understand as well as more or less explicit.

5.5 Summary and conclusions

This chapter presented the analysis of the specific properties of translated texts. The empirical investigation was the first analysis step (cf. Section 5.2), the second one was the cross-linguistic analysis of metaphoricity and ambiguity (cf. Section 5.3) and the third one the psycholinguistic experiment (cf. Section 5.4). For each analysis, hypotheses, results and interpretations were discussed. The aim of the cross-linguistic analysis was to explain the results gained in the empirical analysis from the investigation of the influence the SL has on the translation. Furthermore, the psycholinguistic experiment helped to clarify the results of the cross-linguistic analysis since the cognitive processes taking place during the translation process were investigated. This combination of empirical, cross-linguistic and psycholinguistic research looked at the nature of translated text both from a product- and process-oriented view.

As a result of the empirical analysis, it can be said that the typical as well as the untypical fiction features are used more frequently in English translations than in English originals. Thus, instances of normalization for the typical features and anti-normalization for the untypical ones are found in the translations. This implies that English translated fiction is more narrative and situation-dependent than English original fiction texts, while being more abstract and less edited than the originals. Thus, English translated fiction conforms to the typical patterns of the TL as regards the typical fiction features, whereas it does not in terms of untypical fiction features. This leads to a register shift between English original narrative texts and English translated narrative texts since English translated narrative texts are even more typical of their register owing to the extensive overuse of typical fiction features, on the one hand. However, they point towards a more neutral register due to the extensive overuse of untypical fiction features, on the other.

While the empirical analysis simply served to detect normalization and anti-normalization, it could not explain why they occur. This was the goal of the investigation of cross-linguistic metaphoricity: most of the occurrences of normalization and anti-normalization in both translation directions can be explained through a high degree of re-metaphorization, i.e. the majority of the cases in which a lexico-grammatical feature is normalized or anti-normalized is due to the direct transfer of the SL structures into the TL text. However, this explanation shows how difficult it is to investigate normalization without taking the SL text into account. In cases in which normalization, which was analyzed on the basis of a comparable corpus, is due to re-metaphorization, which was investigated on the basis of a comparable corpus, it is difficult to speak of "real" TL normalization in the narrow sense of the term. In these cases, the typical or untypical fiction features conform to the norms of the TL because of the use of similar lexico-grammatical features in the SL. Thus, in the context of this thesis, normalization can be triggered by re-metaphorization, which means, on a more abstract level, by the influence of the SL. Another interesting observation is that the highest degree of re-metaphorization can be found in French-English translations. The second source of normalization and anti-normalization is de-metaphorization and the third one metaphoriza-

tion, which shows that besides re-metaphorization, de-metaphorization is preferred over metaphorization. This supports the assumption that translation includes the unpacking of information during the understanding process and the incomplete re-packing of information in the production process. Furthermore, it was shown that metaphorization co-occurs with a loss of explicitation, whereas de-metaphorization goes hand in hand with explicitation and simplification. Re-metaphorization can be combined with shining-through.

Like the cross-linguistic approach, the psycholinguistic experiment was based on the same assumption according to which information is not re-packed completely in the translation process. Again, this assumption is substantiated by the results gained in the experiment because de-metaphorization is more frequent than metaphorization. Nevertheless, re-metaphorization is preferred over de-metaphorization and metaphorization, which again underpins the results gained in the cross-linguistic analysis. Another observation which is similar to the findings of the cross-linguistic analysis is that re-metaphorization occurs most frequently in French-English translations. When comparing the verbal and the written translation task it can be said that the teacher and the student use different strategies to deal with the different translation situations. The student tends to de-metaphorize in the verbal translation task and metaphorizes only in the written translation task, whereas the teacher de-metaphorizes more frequently in the written translation task and metaphorizes more often in the verbal translation task. The results of the student strongly support the assumption regarding the incomplete re-packing, whereas the teacher seems to transfer the structures directly when under time pressure. Without time pressure, the teacher tries to render the structures less metaphorical (and thus probably easier to understand) for the reader by de-metaphorizing them. This assumption is supported by the fact that de-metaphorization goes hand in hand with explicitation and simplification.

Additionally, it was shown that the different sources which underlie the specific properties of translations can be identified with the help of the empirical, cross-linguistic and psycholinguistic analyses. Since the three different sources of specific properties of translations, i.e. register, typology

and cognitive processes, were not considered separately as variables in the different analyses, their influence on translated text was merely discussed on the basis of examples.

The importance of the results gained in this chapter for the nature of translated text is discussed in Chapter 6.

Chapter 6

Summary and conclusions

6.1 The nature of translated text

In this thesis, two approaches to empirical linguistics, relevant approaches for cross-linguistic description and an overview of psycholinguistic methods were presented (cf. Chapter 2). On this basis, a methodology for the investigation of the nature of translated text was developed by combining product- as well as process-oriented research (cf. Chapter 4). Finally, the analysis was conducted (cf. Chapter 5): in connection with the empirical approach, the relation between English translations and English originals was investigated on the basis of a comparable corpus. In the context of the cross-linguistic approach, the relation between English translations and their German and French SL texts was analyzed on the basis of a parallel corpus, while the cognitive processes taking place during the translation process were investigated on the basis of a psycholinguistic experiment within the framework of the psycholinguistic approach.

The investigation started from the assumption that translated text constitutes a special kind of text type or register. In order to define the norm of a register and determine how the translations differ, English translated narrative texts were compared to narrative texts originally produced in English. The analysis of English translations in contrast to their German and French SL texts clarified the influence the SL has on the nature of translated texts. Since no clear line could be drawn between the sources of the translation properties, that is to say register, typology and cognitive processes, they were not considered in terms of variables for the investigation.

Their influence on the nature of translated text was merely analyzed on the basis of examples. Translation properties owing to the translation process were investigated in the course of the psycholinguistic experiment, but also on the basis of examples in the context of the cross-linguistic and the psycholinguistic analyses. This leads to the following description of the nature of translated text, or more precisely, the nature of translated English fiction:

- Since normalization of the typical fiction features and anti-normalization of the untypical fiction features can be found in English translated fiction, as compared to English original fiction, the following can be attested: English translated fiction is more narrative and situation-dependent than English original fiction texts (the typical features being overused). At the same time, English translated fiction is more abstract and less edited than original texts (the untypical features being overused). Thus, English translated narrative texts are more typical than English original narrative texts as regards the use of typical features, whereas they point towards a more neutral register owing to the extensive overuse of untypical fiction features.

- Anti-normalization of the untypical fiction features and normalization of the typical fiction features are mainly due to re-metaphorization and, in second and third place, due to de-metaphorization and metaphorization. This means that in most cases the typical and untypical fiction features are re-metaphorized since, to realize the register of fiction, the same constructions are used in the SL texts and the TL texts. There are fewer cases in which the typical and untypical fiction features are de-metaphorized in the TL text compared to the SL text, and even fewer cases of metaphorization of the typical and untypical fiction features. Re-metaphorization, de-metaphorization and metaphorization partly result from register-specific language use or are due to typological factors.

- The translation process provides a further explanation: depending on the ability of the translator and the external constraints which are imposed on the translation process, information is, partly or completely, unpacked in the understanding process and, partly or completely, unpacked in the production process. Furthermore, it was shown on the basis of examples that re-metaphorization co-occurs with

shining-through, metaphorization goes hand in hand with a loss of explicitation and de-metaphorization can be related to explicitation and simplification.

The chart in Figure 6.1 illustrates the interdisciplinary investigation of the nature of translated texts as well as the results gained from the analysis.

6.2 Assessment of the methodology

As described in the beginning of this thesis, an interdisciplinary methodology for investigating the nature of translated text was to be geared to the following requirements (cf. Section 1.3).

- The differences between English translations and English originals were to be described in terms of quantitative patterns.
- The differences between English translations and their German and French originals were to be described and explained in terms of qualitative patterns.
- The translation process was to be investigated in order to explain the nature of translated text.

The first requirement was fulfilled by the empirical analysis of the English comparable corpus in Section 5.2. This approach was based on Baker's prototypical hypotheses (cf. [Baker, 1996]), but profited from linguistically enriched data. This means that lexico-grammatical features such as Biber's register features (cf. [Biber, 1995]) could be investigated automatically. The second requirement was dealt with in the investigation of the German-English and French-English parallel corpus in Section 5.3. For the investigation of the parallel corpus, a method for the cross-linguistic description of the data, as proposed by [Hawkins, 1986], [Newmark, 1988] or [Vinay & Darbelnet, 1995], was needed. However, in order to provide sufficient linguistic interpretation, the SF model was chosen for the cross-linguistic approach. The analysis of the corpus was, basically, a text-based contrastive linguistic analysis for which the SF categories were annotated semi-automatically such that the linguistically interpreted corpus could easily be queried. With

the help of XML, the annotations of the empirical approach and the cross-linguistic approach could be combined to form one integrated version of the corpus, which allowed the extraction of shallow and/or more abstract linguistic features. Finally, the last requirement was covered by the psycholinguistic experiment in Section 5.4. This experiment was especially geared to the cognitive processes during the translation process by exploiting and combining the advantages of TAP, sentence completion task, gating task and lexical decision task.

In summary, the independent variables were a large comparable corpus for the empirical approach vs. a smaller parallel corpus for the cross-linguistic approach vs. a psycholinguistic experiment for the psycholinguistic approach. The relation between English translations and English originals in terms of quantitative patterns vs. the relation between English translations and their German and French originals in terms of qualitative patterns vs. the cognitive processes during the translation process were the dependent variables. Through this variation of variables, the shortcomings of one approach were compensated through the advantages of the others. For instance, the cross-linguistic analysis proved that it is not sufficient to analyze English translations and compare them to English originals in order to detect specific properties of translations (as it is done, for example, by [Baker & Olohan, 2000] or [Laviosa-Braithwaite, 1996b]) since the influence of the SL text is too important for the nature of translated text to be neglected. The empirical investigation, however, showed that it is important to test the statistical significance of the results on the basis of large amounts of data in order to show whether or not they are sufficiently significant to refute or confirm linguistic hypotheses. In contrast to these product-based methods of investigation, the psycholinguistic experiment made possible a process-oriented view regarding the nature of translated text. Furthermore, the combination of methods from translation studies, corpus linguistics, contrastive linguistics, comparative linguistics, computational linguistics, psycholinguistics and linguistics in general guaranteed the fulfillment of the posed requirements for the investigation of the nature of translated text conducted on the basis of product- and process-oriented research.

One drawback of this approach is that the empirical analysis did not take SL texts into account and the cross-linguistic investigation did not deal with comparable texts in the TL. This, however, can be justified for this study since the goals of the respective analyses were clearly defined in such a way that an extension of the sub-corpora was not prerequisite. Another problem of this methodology is that no clear line was drawn between the variables register, typology and cognitive processes since it is difficult to isolate the interacting influences the three possible sources have on the nature of translated text. However, the empirical approach takes into account register-specific language use since a large comparable corpus controlled on the basis of its register is investigated. The cross-linguistic approach focuses on the typological differences between the languages involved since a parallel corpus of SL and TL texts is investigated; and the psycholinguistic approach mainly deals with the cognitive processes taking place during the translation process. Another problem of the analyses carried out in Sections 5.3 and 5.4 is that the statistical significance has not yet been tested in view of the cross-linguistic and psycholinguistic approaches. The reasons for this are the corpus size of the cross-linguistic approach, on the on hand, and the number of test subjects in the psycholinguistic experiment, on the other. It would not have been possible to carry out the analyses on a broader basis in the framework of the present study. As regards the cross-linguistic analysis, the size of the parallel corpus has to be extended by adding further SL texts to the corpus. For the experiment, the number of test subjects would have to be increased to provide an empirical basis for the psycholinguistic analysis. Finally, this study is devoted to the nature of English translated fiction, which means that hypotheses concerning other languages, registers or even translation universals cannot been taken into account.

6.3 Future work

The research which could (and should) immediately follow this thesis can be split up into the following steps:

- **Extension of analyses.** The empirical analysis has to be extended by taking into consideration the SL texts of the translations. This

would allow the analysis of the influence of the SL on a broader basis and with empirical methods. Additionally, a sub-corpus of comparable texts in the TL could be added to the cross-linguistic approach. As a result, the translations could be compared to originals in the TL on the basis of a more profound linguistic analysis. The parallel corpus investigated in Section 5.3 should be extended in order to find more occurrences of grammatical metaphoricity as well as cohesive and structural ambiguity. The psycholinguistic experiment carried out in Section 5.4 should be repeated with more test subjects to provide an empirical basis for significance testing. Significance tests have to be employed on the results gained from the extended cross-linguistic analysis and the psycholinguistic experiment. Moreover, the interdisciplinary methodology developed in this thesis should be exploited for the analysis of translated texts in other languages and registers. This would allow a comprehensive description of the nature of translated text as well as assumptions regarding translation universals.

- **Improvements of conceptual weaknesses.** The main conceptual weakness of this study is that the possible sources of translation properties, i.e. register, typology and cognitive processes, were not dealt with separately as variables in the course of the investivation of the specific properties of translated text. The three different approaches consider different sources respectively, i.e. the empirical analysis focuses on register as possible source of translation properties, the cross-linguistic investigation mainly deals with typology as possible source and the psycholinguistic experiment considers the cognitive processes during the translation process as possible sources. However, the influences of the three sources interact in each of the three investigations. Thus, the separation of these three sources would be an improvement in the sense that their influence on the specific nature of translated text could be analyzed independently for register, typology and cognitive processes. For this purpose, corpus analyses, which could be similar to this study, should be carried out in which the register-specific language use as well as the typological differences between the languages involved are better controlled than in this study. This means that the register should not be defined as broadly as it was in the context of this thesis. Furthermore, a comprehensive description of

register use and typological differences should be provided with a view to analyzing how translations differ in these respects. Such a corpus design is proposed in [Teich, 2001a] for analyzing cross-linguistic variation in translations and multilingually comparable texts. As a result of employing such a corpus design for the investigation of the nature of translated text, translation properties due to register-specific language use can be distinguished from translation phenomena on the basis of typological differences between the languages involved. Moreover, "real" or "pure" properties of translation which are due to the translation processes itself can clearly be identified through such an conceptual improvement.

- **Addition of other computational techniques.** Other computational techniques could be added to the methodology proposed in this thesis in order to broaden the view on the nature of translated text. For example, eye-tracking methods (cf. Section 2.4) could be used to give more detailed insight into the process of understanding taking place when a translator reads an SL text. In such an experiment, the eye movements of readers who are not translators could be compared to those of translators while they read a text before translating it. One would assume that, with translators, possible translation problems or the recognition of cohesive chains result in longer gaze durations or regressions. Another psycholinguistic experiment which could be employed is the window method (cf. Section 2.4). With this computational technique, the reading times of the other readers could be compared to the time it takes the translators to read an SL text. One would assume that longer reading times are recorded for translators because, consciously or unconsciously, they carry out a text analysis as a preparatory step for the translation. As a result of the eye-tracking or window methods, the understanding process of translators could be investigated in more detail. Additionally, the use of EEGs, ERPs, MRIs or tomography scans (cf. Section 2.4) would make it possible to analyze which parts of the brain are generally activated during a translation process and which parts are necessary to understand an SL text, transfer it into another language and produce a TL text, respectively.

  As to computational corpus linguistics, large linguistically interpreted

corpora could serve as comparable corpora for the interdisciplinary investigation of the nature of translated text proposed in this thesis. For instance, the Penn Treebank (cf. [Marcus et al., 1993]) could be used as an English reference corpus and the TIGER Treebank (cf. [Brants & Hansen, 2002]) could be employed for German.

The linguistic annotation of multiple layers within a corpus is another important issue concerning corpus-based research. An integrated XML corpus annotation, as it was proposed in Chapter 4, is appropriate if there is a hierarchical relation between the linguistic units annotated in the corpus (as it is, for example, the case with syntactic annotation co-occurring with part-of-speech tagging). However, for the annotation of overlapping linguistic segments (as it is, for example, the case with syntactic annotation in combination with phonological annotation), it would be more appropriate to keep the annotated layers apart by using XML-based *stand-off mark-up* (cf. [Thompson & McKelvie, 1997]). In the internal representation of such a multi-layer corpus annotation, the intersecting hierarchies are represented as *ordered directed acyclic graphs (ODAGs)* (cf. [Sperberg-McQueen & Huitfeldt, 2001]). An alternative to ODAGs would be the use of *annotation graphs (AGs)* (cf. [Bird & Liberman, 2001]). AGs allow multiple perspectives on a corpus identifying possible intersecting segmentations on the basis of a given time line. In [Teich et al., 2001], the representation of multi-layer annotated corpora is discussed in further detail.

The issue of querying a multi-layer annotated corpus is related to the linguistic annotation of multiple layers within a corpus. In Section 4, Q4M (cf. [Mengel, 1999]) was proposed for querying a multi-layer annotated corpus on the basis of an integrated XML representation. For the extraction of annotated data represented as ODAGs or AGs special purpose query languages, such as the ones proposed by [McKelvie et al., 2001] for ODAGs and by [Bird et al., 2000] for AGs, can be used. As an alternative to special purpose query languages, a standardized query language for XML, XQuery (cf. [Teich et al., 2001]), can be employed.

However, querying a corpus does not only encompass the technical perspective, but also the linguistic one: linguists are interested in rather complex research questions they wish to test on the basis of a linguisti-

cally interpreted corpus. These research questions frequently concern the interaction of different annotated layers within a corpus, such as mutual constraints between the properties of grammatical units and phonological units, which are rather complex to query. For this reason, from a linguistics point of view, it would be interesting to classify the research questions as well as the appropriate queries in order to build up a query library for the most common forms of representation (e.g., ODAGs and AGs). Thus, the annotation of a linguistically interpreted corpus should include a query library which covers the major linguistic research questions in order to provide a comprehensive and user-friendly resource for linguists. A good example of this is the treebank project TIGER (cf. [Brants & Hansen, 2002]), in which the creation of such a query library is part of the second project phase.

Another important aspect for future work is the application of the interdisciplinary method for the investigation of the nature of translated text to other linguistic contexts. The following contexts are of special interest for applying the interdisciplinary methodology proposed in this study:

- **Foreign language acquisition.** The interdisciplinary methodology proposed in this thesis could be used, for example, for the investigation of foreign language acquisition. For this purpose, a corpus composed of texts written by language learners has to be analyzed and compared to texts produced by native speakers. Furthermore, psycholinguistic tests, similarly to the one suggested in this thesis, have to be carried out in order to obtain insight into the cognitive processes involved in learning a language. As a result, language teaching could be better geared to the requirements of foreign language learners.

- **Bilingual language acquisition.** In a similar way, the language of bilingually raised children could be investigated. It could be tested whether they have similar knowledge of both languages or whether one language is predominant. Additionally, their language use can be compared to monolingually raised children, which could reveal possible weaknesses or problems the bilingually raised children have. Such findings could improve the way of (language) teaching, dealing with these children's needs.

- **Non-native scientific English.** The investigation of scientific English produced by non-native speakers represents a further field of application. Through the use of the interdisciplinary methodology proposed in this thesis, non-native scientific English could be compared to native scientific English with the help of product- as well as process-oriented analyses. On the basis of the gained results, style guides for non-native English speakers could be provided in which relevant problems are pointed out and solutions are proposed so as to enhance the quality of non-native scientific English.

As can be seen in the present thesis, for my purposes, I exploited the resources provided by translation studies, corpus linguistics, contrastive linguistics, comparative linguistics, computational linguistics, psycholinguistics and linguistics in general. Thus, in my opinion, another important issue for future work is that neighboring research disciplines should make an effort to profit from each other and benefit from each other's resources.

Figure 6.1: An interdisciplinary investigation of the nature of translated text

Bibliography

[Aaronson & Ferres, 1984] Aaronson, D., & Ferres, S. 1984. The word-by-word reading paradigm: An experimental and theoretical approach. *Pages 31–68 of:* Kieras, D. E., & Just, M. A. (eds), *New methods in reading comprehension research.* Hillsdale: Erlbaum.

[Albrecht & O'Brien, 1991] Albrecht, J. E., & O'Brien, E. J. 1991. Effects of centrality on retrieval of text-based concepts. *Pages 932–939 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 17.

[Atril, 2000] Atril (ed). 2000. *Déjà Vu. Productivity system for translators.* Development SL. Software Manual.

[Baker, 1993] Baker, M. 1993. Corpus Linguistics and Translation Studies: Implications and Applications. *Pages 233–250 of:* Baker, M., Francis, G., & Tognini-Bonelli, E. (eds), *Text and Technology: In Honour of John Sinclair.* Amsterdam and Philadelphia: John Benjamins.

[Baker, 1995] Baker, M. 1995. Corpora in translation studies: An overview and some suggestions for future research. *Pages 223–243 of: Target*, vol. 7. no. 2.

[Baker, 1996] Baker, M. 1996. Corpus-based translation studies: The challenges that lie ahead. *Pages 175–186 of:* Somers, H. (ed), *Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager.* Amsterdam: John Benjamins.

[Baker, 1998] Baker, M. 1998. Rexplorer la langue de la traduction: une approche par corpus (Investigating the Language of Translation: A Corpus-based Approach). *Pages 480–485 of: Meta*, vol. 43. no. 4.

[Baker, 1999] Baker, M. 1999. The role of corpora in investigating the linguistic behaviour of professional translators. *Pages 281–298 of: International Journal of Corpus Linguistics*, vol. 4. no. 2.

[Baker, 2000a] Baker, M. 2000a. Linguistic Perspectives on Translation. *Pages 20–26 of:* France, P. (ed), *The Oxford Guide to Literature in English Translation.* Oxford: Oxford University Press.

[Baker, 2000b] Baker, M. 2000b. Towards a Methodology for Investigating the Style of a Literary Translator. *Pages 241–266 of: Target*, vol. 12. no. 2.

[Baker & Olohan, 2000] Baker, M., & Olohan, M. 2000. Reporting that in Translated English: Evidence for Subconscious Processes of Explicitation? *Pages 141–158 of: Across Languages and Cultures*, vol. 1. no. 2.

[Baluch & Besner, 1991] Baluch, B., & Besner, D. 1991. Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Pages 644–652 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 17.

[Bateman et al., 1991] Bateman, J., Matthiessen, C. M. I. M., Nanri, K., & Zeng, L. 1991. The re-use of linguistic resources across languages in multilingual generation. *Pages 966–971 of: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI).*

[Beach, 1991] Beach, C. M. 1991. The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Pages 644–663 of: Journal of Memory and Language*, vol. 30.

[Berman, 2000] Berman, A. 2000. Translation and the trials of the foreign. *Pages 284–297 of:* Venuti, L. (ed), *The translation studies reader.* London and New York: Routledge.

[Biber, 1988] Biber, D. 1988. *Variation across Speech and Writing.* Cambridge: Cambridge University Press.

[Biber, 1990] Biber, D. 1990. Methodological issues regarding corpus-based analyses of linguistic variation. *Pages 257–269 of: Literary and Linguistic Computing*, vol. 5.

[Biber, 1992] Biber, D. 1992. On the complexity of discourse complexity: a multidimensional analysis. *Pages 133–163 of: Discourse Processes*, vol. 15.

[Biber, 1993] Biber, D. 1993. Using register-diversified corpora for general language studies. *Pages 219–242 of: Computational Linguistics*, vol. 19. no. 2.

[Biber, 1995] Biber, D. 1995. *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.

[Biber & Finegan, 1993] Biber, D., & Finegan, E. (eds). 1993. *Perspectives on register: situating register variation within sociolinguistics*. Oxford University Press: Oxford.

[Biber & Hared, 1992] Biber, D., & Hared, M. 1992. Dimensions of register variation in Somali. *Pages 41–75 of: Language variation and change*, vol. 4.

[Biber et al., 1998] Biber, D., Conrad, S., & Reppen, R. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

[Biber et al., 1999] Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

[Bird & Harrington, 2001] Bird, S., & Harrington, J. (eds). 2001. *Speech Annotation and Corpus Tools. A special issue of Speech Communication*. vol. 33. no. 1–2.

[Bird & Liberman, 2001] Bird, S., & Liberman, M. 2001. A formal framework for linguistic annotation. *Pages 23–60 of: Speech Communication*, vol. 33. no. 1–2.

[Bird et al., 2000] Bird, S., Bunemann, P., & Tan, W.-C. 2000. Towards a query language for annotation graphs. *In: Proceedings of LREC*.

[Bloom et al., 1990] Bloom, C. P., Fletcher, C. R., van den Broek, P., Reitz, L., & Shapiro, B. P. 1990. An on-line assessment of causal reasoning during comprehension. *Pages 65–71 of: Memory and Cognition*, vol. 18.

[Blum-Kulka et al., 1989] Blum-Kulka, S., House, J., & Kasper, G. (eds). 1989. *Cross-cultural Pragmatics: requests and apologies.* Norwood: Ablex.

[Blutner & Sommer, 1988] Blutner, R., & Sommer, R. 1988. Sentence processing and lexical access: The influence of the focus-identifying task. *Pages 359–367 of: Journal of Memory and Language*, vol. 27.

[Branigan et al., 1995] Branigan, H. P., Pickering, M. J., Liversedge, S. P., Stewart, A. P., & Urbach, T. P. 1995. Syntactic priming: Investigating the mental representation of language. *Pages 489–506 of: Journal of Psycholinguistic Research*, vol. 24.

[Brants & Hansen, 2002] Brants, S., & Hansen, S. 2002. Developments in the TIGER Annotation Scheme and their Realization in the Corpus. *Pages 1643–1649 of: Proceedings of LREC.*

[Brants, 1999] Brants, T. 1999. *Tagging and Parsing with Cascaded Markov Models - Automation of Corpus Annotation.* Ph.D. thesis, German Research Center for Artificial Intelligence and Saarland University, Saarbrücken. Saarbrücken Dissertations in Computational Linguistics and Language Technology. vol. 6.

[Brants, 2000] Brants, T. 2000. TnT – A Statistical Part-of-Speech Tagger. *In: Proceedings of the Sixth Applied Natural Language Processing Conference ANLP-2000.*

[Bryson, 1991] Bryson, B. 1991. *Mother Tongue.* London: Penguin Books.

[Caffarel, 1996] Caffarel, A. 1996. *Prolegomena to a systemic functional interpretation of French grammar: from discourse to grammar and back.* Ph.D. thesis, Sydney University, Sydney.

[Chomsky, 1970] Chomsky, N. 1970. Remarks on nominalization. *In:* Jacobs, R. A., & Rosenbaum, P. S. (eds), *Readings in English Transformational Grammar.* Waltham: Ginn.

[Christ, 1994] Christ, O. 1994. A modular and flexible architecture for an integrated corpus query system. *Pages 23–32 of: Proceedings of COMPLEX 94. 3rd Conference on Computational Lexicography and Text Research.*

[Claparède, 1932] Claparède, E. 1932. Die Entdeckung der Hypothese. *Pages 109–115 of:* Graumann, C. F. (ed), *Denken.* Köln: Kiepenhauer and Witsch.

[Clark & Chase, 1972] Clark, H. H., & Chase, W. G. 1972. On the process of comparing sentences against pictures. *Pages 472–517 of: Cognitive Psychology*, vol. 3.

[Clark & Chase, 1974] Clark, H. H., & Chase, W. G. 1974. Perceptual coding strategies in the formation and verification of descriptions. *Pages 101–111 of: Memory and Cognition*, vol. 2.

[Clifton et al., 1991] Clifton, C., Speer, S., & Abney, S. P. 1991. Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Pages 251–272 of: Journal of Memory and Language*, vol. 30.

[Cutler et al., 1986] Cutler, A., Mehler, J., Norris, D., & Segui, J. 1986. The syllable's differing role in the segmentation of French and English. *Pages 385–400 of: Journal of Memory and Language*, vol. 25.

[Cutler et al., 1987a] Cutler, A., Butterfield, S., & Williams, J. N. 1987a. The perceptual integrity of syllabic onsets. *Pages 406–418 of: Journal of Memory and Language*, vol. 26.

[Cutler et al., 1987b] Cutler, A., Mehler, J., Norris, D., & Segui, J. 1987b. Phoneme identification and the lexicon. *Pages 141–177 of: Cognitive Psychology*, vol. 19.

[Dell et al., 1983] Dell, G. S., McKoon, G., & Ratcliff, R. 1983. The activation of antecedent information during the processing of anaphoric reference in reading. *Pages 121–132 of: Journal of Verbal Learning and Verbal Behavior*, vol. 22.

[Doherty, 1991] Doherty, M. 1991. Informationelle Holzwege. *Pages 30–49 of: LiLi. Zeitschrift für Literaturwissenschaft und Linguistik*, vol. 84.

[Doherty, 1993] Doherty, M. 1993. Parametrisierte Perspektive. *Pages 3–38 of: Zeitschrift für Sprachwissenschaft*, vol. 12. Göttingen: Vandenhoeck and Ruprecht. no. 1.

[Doherty, 1996] Doherty, M. 1996. Passive perspectives: different preferences in English and German – a result of parametrized processing. *Pages 591–644 of: Linguistics*, vol. 34. no. 3.

[Doherty, 1997] Doherty, M. 1997. Acceptability and Language Specific Preferences in the Distribution of Information. *Pages 1–24 of: Target*, vol. 9. no. 1.

[Doherty, 1998a] Doherty, M. 1998a. Clauses or phrases: a principled account of *when*-clauses in translations between English and German. *Pages 235–254 of:* Johansson, S., & Oksefjell, S. (eds), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies.* Amsterdam: Rodopi.

[Doherty, 1998b] Doherty, M. 1998b. Processing and Directionality in German and English. *Pages 23–43 of: Languages in Contrast*, vol. 1. no. 1.

[Doherty, 1999a] Doherty, M. 1999a. Clefts in Translations between English and German. *Pages 289–315 of: Target*, vol. 11. no. 2.

[Doherty, 1999b] Doherty, M. 1999b. Position and explicitness – language specific conditions for the use of adverbial clauses in translations between German and English. *Pages 112–148 of:* Doherty, M. (ed), *Sprachspezifische Aspekte der Informationsverteilung. studia grammatica*, vol. 47. Berlin: Akademie-Verlag.

[Duncker, 1935] Duncker, K. 1935. *Zur Psychologie des produktiven Denkens.* Berlin: Springer.

[Dupoux & Mehler, 1990] Dupoux, E., & Mehler, J. 1990. Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. *Pages 316–335 of: Journal of Memory and Language*, vol. 29.

[Ehrich & Rapp, 2000] Ehrich, V., & Rapp, I. 2000. Sortale Bedeutung und Argumentationsstruktur: ung-Nominalisierungen im Deutschen. *Pages 245–303 of: Zeitschrift für Sprachwissenschaft*, vol. 19. no. 2.

[Ehrlichman & Wiener, 1980] Ehrlichman, H., & Wiener, M. 1980. EEG asymmetry during covert mental activity. *Pages 228–236 of: Psychophysiology*, vol. 17.

[Eysenck & Keane, 1990] Eysenck, M. W., & Keane, M. T. 1990. *Cognitive psychology - a student's handbook*. Hove: Erlbaum.

[Fabricius-Hansen, 1998] Fabricius-Hansen, C. 1998. Information density and translation, with special reference to German – Norwegian – English. *Pages 197–234 of:* Johansson, S., & Oksefjell, S. (eds), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies*. Amsterdam: Rodopi.

[Fabricius-Hansen, 1999a] Fabricius-Hansen, C. 1999a. *Bei dieser Gelegenheit – on this occasion – ved denne leilighet*. German *bei* – a puzzle in translational perspective. *In:* Hasselgård, H., & Oksefjell, S. (eds), *Out of Corpora. Festschrift in honour of Stig Johansson*. Amsterdam: Rodopi.

[Fabricius-Hansen, 1999b] Fabricius-Hansen, C. 1999b. Information packaging and translation: Aspects of translational sentence splitting (German - English/Norwegian). *Pages 175–214 of:* Doherty, M. (ed), *Sprachspezifische Aspekte der Informationsverteilung. studia grammatica*, vol. 47. Berlin: Akademie-Verlag.

[Fawcett, 1997] Fawcett, P. 1997. *Translation and Language. Linguistic Theories Explained*. Manchester: St. Jerome.

[Ferreira, 1991] Ferreira, F. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances. *Pages 210–233 of: Journal of Memory and Language*, vol. 30.

[Ferreira & Anes, 1994] Ferreira, F., & Anes, M. 1994. Why Study Spoken Language? *Pages 33–56 of:* Gernsbacher, M. A. (ed), *Handbook of Psycholinguistics*. San Diego: Academic Press.

[Firth, 1957] Firth, J. 1957. *Papers in Linguistics 1934-51*. London: Oxford University Press.

[Fodor & Bever, 1965] Fodor, J. A., & Bever, T. G. 1965. The psychological reality of linguistic segments. *Pages 414–420 of: Journal of Verbal Learning and Verbal Behavior*, vol. 4.

[Fodor et al., 1974] Fodor, J. A., Bever, T. G., & Garrett, M. 1974. *The psychology of language*. New York: McGraw Hill.

[Forster, 1981] Forster, K. I. 1981. Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Pages 465–495 of: Quarterly Journal of Experimental Psychology*, vol. 33A.

[Frauenfelder et al., 1990] Frauenfelder, U. H., Segui, J., & Dijkstra, T. 1990. Lexical effects in phonemic processing: Facilitatory or inhibitory? *Pages 77–91 of: Journal of Experimental Psychology: Human Perception and Performance*, vol. 16.

[Galin & Ornstein, 1972] Galin, D., & Ornstein, R. E. 1972. Lateral specialization of cognitive mode. II. EEG frequency analysis. *Pages 412–418 of: Psychophysiology*, vol. 9.

[Garrett et al., 1965] Garrett, M., Bever, T. G., & Fodor, J. A. 1965. The active use of grammar in speech perception. *Pages 30–32 of: Perception and Psychophysics*, vol. 1.

[Garrod et al., 1990] Garrod, S., O'Brien, E. J., Morris, R. K., & Rayner, K. 1990. Elaborative inferencing as an active or passive process. *Pages 250–269 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 16.

[Garside, 1987] Garside, R. 1987. The CLAWS Word-tagging System. *In:* Garside, R., Leech, G., & Sampson, G. (eds), *The Computational Analysis of English: A Corpus-based Approach.* London: Longman.

[Geers, 1978] Geers, A. E. 1978. Intonation contour and syntactic structure as predictors of apparent segmentation. *Pages 273–283 of: Journal of Experimental Psychology: Human Perception and Performance*, vol. 4.

[Gernsbacher, 1989] Gernsbacher, M. A. 1989. Mechanisms that improve referential access. *Pages 99–156 of: Cognition*, vol. 32.

[Gernsbacher & Faust, 1990] Gernsbacher, M. A., & Faust, M. 1990. The role of suppression in sentence comprehension. *Pages 97–128 of:* Simpson, G. B. (ed), *Understanding word and sentence.* Amsterdam: North-Holland.

[Gernsbacher & Shroyer, 1989] Gernsbacher, M. A., & Shroyer, S. 1989. The cataphoric use of the indefinite *this* in spoken narratives. *Pages 536–540 of: Memory and Cognition*, vol. 17.

[Ghadessy, 1988] Ghadessy, M. (ed). 1988. *Registers of written English: situational factors and linguistic features*. London: Pinter.

[Ghadessy, 1993] Ghadessy, M. (ed). 1993. *Register Analysis. Theory and Practice*. London: Pinter.

[Glenberg et al., 1987] Glenberg, A. M., Meyer, M., & Lindem, K. 1987. Mental models contribute to foregrounding during text comprehension. *Pages 69–83 of: Journal of Memory and Language*, vol. 26.

[Gordon, 1983] Gordon, B. 1983. Lexical access and lexical decision: Mechanism of frequency sensitivity. *Pages 24–44 of: Journal of Verbal Learning and Verbal Behavior*, vol. 22.

[Gough, 1965] Gough, P. B. 1965. Grammatical transformations and speed of understanding. *Pages 107–111 of: Journal of Verbal Learning and Verbal Behavior*, vol. 4.

[Gough, 1966] Gough, P. B. 1966. The verification of sentences: The effects of delay of evidence and sentence length. *Pages 492–496 of: Journal of Verbal Learning and Verbal Behavior*, vol. 5.

[Grosjean, 1980] Grosjean, F. 1980. Spoken word recognition processes and the gating paradigm. *Pages 299–310 of: Perception and Psychophysics*, vol. 28.

[Haberlandt, 1984] Haberlandt, K. 1984. Components of sentence and word reading times. *Pages 219–252 of:* Kieras, D. E., & Just, M. A. (eds), *New methods in reading comprehension research*. Hillsdale: Erlbaum.

[Haberlandt, 1994] Haberlandt, K. 1994. Methods in Reading Research. *Pages 1–31 of:* Gernsbacher, M. A. (ed), *Handbook of Psycholinguistics*. San Diego: Academic Press.

[Haberlandt & Graesser, 1985] Haberlandt, K., & Graesser, A. C. 1985. Component processes in text comprehension and some of their interactions. *Pages 357–374 of: Journal of Experimental Psychology: General*, vol. 114.

[Haberlandt & Graesser, 1989] Haberlandt, K., & Graesser, A. C. 1989. Processing of new arguments at clause boundaries. *Pages 186–193 of: Memory and Cognition*, vol. 17.

[Haberlandt & Graesser, 1990] Haberlandt, K., & Graesser, A. C. 1990. Integration and buffering of new information. *Pages 71–87 of:* Graesser, A. C., & Bower, G. (eds), *Inferences and text comprehension*. San Diego: Academic Press.

[Halliday, 1961] Halliday, M. A. K. 1961. Categories of the theory of grammar. *Pages 241–292 of: Word*, vol. 17.

[Halliday, 1963] Halliday, M. A. K. 1963. Class in relation to the axes of chain and choice in language. *Pages 5–15 of: Linguistics*, vol. 2.

[Halliday, 1967] Halliday, M. A. K. 1967. Notes on transitivity and theme in English – parts 1 and 2. *Pages 37–81 and 199–244 of: Linguistics*, vol. 3.

[Halliday, 1968] Halliday, M. A. K. 1968. Notes on transitivity and theme in English – part 3. *Pages 179–215 of: Linguistics*, vol. 4.

[Halliday, 1973] Halliday, M. A. K. 1973. *Explorations in the Functions of Language*. London: Edward Arnold.

[Halliday, 1976] Halliday, M. A. K. 1976. *System and Function in Language*. London: Oxford University Press.

[Halliday, 1978] Halliday, M. A. K. 1978. *Language as social semiotic*. London: Edward Arnold.

[Halliday, 1985] Halliday, M. A. K. 1985. *An introduction to Functional Grammar*. London: Edward Arnold.

[Halliday, 2001] Halliday, M. A. K. 2001. Towards a theory of good translation. *Pages 13–18 of:* Steiner, E., & Yallop, C. (eds), *Exploring Translation and Multilingual Text Production: Beyond Content. Series Text, Translation, Computational Processing*. Berlin and New York: Mouton de Gruyter.

[Halliday & Hasan, 1976] Halliday, M. A. K., & Hasan, R. 1976. *Cohesion in English*. London: Longman.

[Halliday & Hasan, 1989] Halliday, M. A. K., & Hasan, R. 1989. *Language, Context, and Text: Aspects of Language in a Social Semiotic Perspective.* Oxford: Oxford University Press.

[Halliday & Matthiessen, 1999] Halliday, M. A. K., & Matthiessen, C. M. I. M. 1999. *Construing experience through meaning. A language-based approach to cognition.* London and New York: Cassell.

[Hansen & Teich, 1999] Hansen, S., & Teich, E. 1999. Kontrastive Analyse von Übersetzungskorpora: ein funktionales Modell. *Pages 311–322 of:* Gippert, J. (ed), *Multilinguale Corpora. Codierung, Strukturierung und Analyse (Sammelband der Jahrestagung der GLDV 99).* Prag: enigma corporation.

[Hansen & Teich, 2001] Hansen, S., & Teich, E. 2001. Multi-layer analysis for translation corpora: methodological issues and practical implications. *Pages 44–55 of:* Cristea, D., Ide, N., Marcu, D., & Poesio, M. (eds), *Proceedings of EUROLAN 2001 Workshop on Multi-layer Corpus-based Analysis.*

[Haspelmath et al., 2001] Haspelmath, M., König, E., Oesterreicher, W., & Raible, W. (eds). 2001. *Sprachtypologie und sprachliche Universalien. Handbücher zur Sprach- und Kommunikationswissenschaft. vol. 20.* vol. 1 and 2. Berlin and New York: Mouton de Gruyter.

[Hasselgård, 1998] Hasselgård, H. 1998. Thematic structure in translation between English and Norwegian. *Pages 145–167 of:* Johansson, S., & Oksefjell, S. (eds), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies.* Amsterdam: Rodopi.

[Hatim & Mason, 1990] Hatim, B., & Mason, I. 1990. *Discourse and the translator.* London: Longman.

[Hawkins, 1981] Hawkins, J. A. 1981. The semantic diversity of basic grammatical relations between English and German. *Pages 1–25 of: Linguistische Berichte*, vol. 75.

[Hawkins, 1983] Hawkins, J. A. 1983. *Word order universals.* New York: Academic Press.

[Hawkins, 1984] Hawkins, J. A. 1984. Modifier-head or Function-argument relations in phrase structure? The evidence of some word-order universals. *Pages 107–138 of: Lingua*, vol. 63.

[Hawkins, 1986] Hawkins, J. A. 1986. *A comparative typology of English and German*. London and Sydney: Croom Helm.

[Hawkins, 1994] Hawkins, J. A. 1994. A performance theory of order and constituency. *In: Cambridge Studies in Linguistics*, vol. 73. Cambridge University Press.

[House, 1977] House, J. 1977. *A Model for Translation Quality Assessment*. Tübingen: Gunter Narr.

[House, 1997] House, J. 1997. *Translation Quality Assessment. A Model Revisited*. Tübingen: Gunter Narr.

[Inhoff, 1987] Inhoff, A. W. 1987. Parafoveal word perception during eye fixations in reading: Effects of visual salience and word structure. *Pages 403–418 of:* Coltheart, M. (ed), *Attention and performance XII: The psychology of reading*. Hove: Erlbaum.

[J. Doyle & Galin, 1974] J. Doyle, R. E. Ornstein, & Galin, D. 1974. Lateral specialization of cognitive mode: An EEG study. *Pages 247–252 of: Psychophysiology*, vol. 16.

[Jääskeläinen, 1989] Jääskeläinen, R. 1989. Translation Assignment in Professional vs. Non-Professional Translation: A Think-Aloud Protocal Study. *Pages 87–98 of:* Séguinot, C. (ed), *The Translation Process*. Toronto: H. G. Publications.

[Jääskeläinen, 1993] Jääskeläinen, R. 1993. Investigating Translation Strategies. *In:* Tirkkonen-Condit, S., & Laffling, J. (eds), *Recent Trends in Empirical Translation Research*. Joensuu: University of Joensuu, Faculty of Arts.

[Jääskeläinen & Tirkkonen-Condit, 1991] Jääskeläinen, R., & Tirkkonen-Condit, S. 1991. Automatised Processes in Professional vs. Non-Professional Translation: A Think-Aloud Protocal Study. *Pages 89–109 of:* Tirkkonen-Condit, S. (ed), *Empirical Research in Translation and Intercultural Studies*. Tübingen: Gunter Narr.

[Jakimik & Glenberg, 1990] Jakimik, J., & Glenberg, A. 1990. Verbal learning meets psycholinguistics: Modality effects in the comprehension of anaphora. *Pages 582–590 of: Journal of Memory and Language*, vol. 29.

[Johansson, 1995] Johansson, S. 1995. Mens Sana in corpore sano: On the role of corpora in linguistic research. *Pages 19–25 of: The European English Messenger*, vol. 4. no. 2.

[Johansson, 1997a] Johansson, S. 1997a. In search of the missing *not*: Some notes on negation in English and Norwegian. *Pages 197–214 of:* Fries, U., Müller, V., & Schneider, P. (eds), *From Ælfric to the New York Times: Studies in English Corpus Linguistics*.

[Johansson, 1997b] Johansson, S. 1997b. Using the English-Norwegian Parallel Corpus – a corpus for contrastive analysis and translation studies. *Pages 282–296 of:* Lewandowska-Tomaszczyk, B., & Melia, P. J. (eds), *Practical Applications in Language Corpora*. Lodz: Lodz University.

[Johansson, 1998a] Johansson, S. 1998a. On computer corpora in contrastive linguistics. *Pages 269–289 of:* Cooper, W. R. (ed), *Compare or contrast? Current issues in cross-language research. Tampere Studies in English*. Tampere: University of Tampere.

[Johansson, 1998b] Johansson, S. 1998b. On the role of corpora in cross-linguistic research. *Pages 1–24 of:* Johansson, S., & Oksefjell, S. (eds), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies*. Amsterdam: Rodopi.

[Johansson, 2001a] Johansson, S. 2001a. Contrastive linguistics and corpora. *In: SPRIKreport no. 3. Reports from the Project "Languages in Contrast"*. Oslo: University of Oslo.

[Johansson, 2001b] Johansson, S. 2001b. The English verb *seem* and its correspondences in Norwegian: What seems to be the problem? *Pages 221–245 of:* Aijmer, K. (ed), *A Wealth of English. Studies in Honour of Göran Kjellmer*. Göteborg: Acta Universitatis Gothoburgensis.

[Johansson, 2001c] Johansson, S. 2001c. The German and Norwegian correspondences to the English construction type *that's what*. *Pages 583–605 of: Linguistics*.

[Johansson, 2001d] Johansson, S. 2001d. Towards a multilingual corpus for contrastive analysis and translation studies. *In: SPRIKreport no. 1. Reports from the Project "Languages in Contrast"*. Oslo: University of Oslo.

[Johansson, 2001e] Johansson, S. 2001e. Translationese: Evidence from the English-Norwegian Parallel Corpus. *Pages 162–176 of:* Allén, S., Berg, S., Malmgren, S. G. R., Norén, K., & Ralph, B. (eds), *Gäller stam, suffix ach ord. Festskrift till Martin Gellerstam*. Göteborg: Elanders Novum.

[Johansson & Ebeling, 1996] Johansson, S., & Ebeling, J. 1996. Exploring the English-Norwegian parallel corpus. *Pages 3–15 of:* Percy, C. E., Meyer, Ch. F., & Lancashire, I. (eds), *Synchronic Corpus Linguistics*. Amsterdam: Rodopi.

[Johansson & Hasselgård, 1999] Johansson, S., & Hasselgård, H. 1999. Corpora and cross-linguistic research in the Nordic countries. *Pages 145–162 of:* Granger, S., Beheydt, L., & Colson, J. P. (eds), *Contrastive Linguistics and Translation. Special edition of Le Langage et l'Homme*, vol. XXXIV. no. 1.

[Johansson & Hofland, 1994] Johansson, S., & Hofland, K. 1994. Towards an English-Norwegian parallel corpus. *Pages 25–37 of:* Fries, U., Tottie, G., & Schneider, P. (eds), *Creating and Using English Language Corpora. Papers from the Fourteenth International Conference on English Language Research on Computerized Corpora*. Amsterdam: Rodopi.

[Johansson & Hofland, 2000] Johansson, S., & Hofland, K. 2000. The English-Norwegian Parallel Corpus: Current work and new directions. *In:* Botley, S. P., McEnery, A. M., & Wilson, A. (eds), *Multilingual Corpora in Teaching and Research*. Amsterdam: Rodopi.

[Johansson & Løken, 1997] Johansson, S., & Løken, B. 1997. Some Norwegian discourse particles and their English correspondences. *Pages 149–170 of:* Bache, C., & Klinge, A. (eds), *Sounds, Structures and Senses. Essays Presented to Niels Davidsen-Nielsen on the Occasion of His Sixtieth Birthday*. Odense: Odense University Press.

[Johansson et al., 1996] Johansson, S., Ebeling, J., & Hofland, K. 1996. Coding and aligning the English-Norwegian parallel corpus. *Pages 87–*

112 of: Aijmer, K., Altenberg, B., & Johansson, M. (eds), *Languages in Contrast. Papers from a symposium on text-based cross-linguistic studies.* Lund: Lund University Press.

[Just & Carpenter, 1980] Just, M. A., & Carpenter, P. A. 1980. A theory of reading: From eye fixations to comprehension. *Pages 329–354 of: Psychological Review*, vol. 87.

[Just et al., 1982] Just, M. A., Carpenter, P. A., & Woolley, J. D. 1982. Paradigms and processes in reading comprehension. *Pages 228–238 of: Journal of Experimental Psychology: General*, vol. 111.

[Kenny, 1997] Kenny, D. 1997. (Ab)normal Translations: a German-English Parallel Corpus for Investigating Normalization in Translation. *Pages 387–392 of:* Lewandowska-Tomaszczyk, B., & Melia, P. J. (eds), *Practical Applications in Language Corpora. PALC '97 Proceedings.*

[Kenny, 1998a] Kenny, D. 1998a. Corpora in Translation Studies. *Pages 50–53 of:* Baker, M. (ed), *Routledge Encyclopedia of Translation Studies.* London and New York: Routledge.

[Kenny, 1998b] Kenny, D. 1998b. Creatures of Habit? What Translators Usually Do with Words. *Pages 515–524 of: Meta*, vol. 43. no. 4.

[Kenny, 1999] Kenny, D. 1999. The German-English Parallel Corpus of Literary Texts (GEPCOLT): A Resource for Translation Scholars. *Pages 25–42 of: Teanga*, vol. 18.

[Kenny, 2000a] Kenny, D. 2000a. Lexical Hide-and-Seek: looking for creativity in a parallel corpus. *Pages 93–104 of:* Olohan, M. (ed), *Intercultural Faultlines: Research Models in Translation Studies I: Textual and Cognitive Aspects.* Manchester: St. Jerome.

[Kenny, 2000b] Kenny, D. 2000b. Translators at Play: Exploitations of Collocational Norms in German-English Translation. *Pages 143–160 of:* Dodd, B. (ed), *Working with German Corpora.* Birmingham: University of Birmingham Press.

[Kenny, 2001] Kenny, D. 2001. *Lexis and Creativity in Translation. A Corpus-based Study.* Manchester: St. Jerome.

[Kittredge & Lehrberger, 1982] Kittredge, R., & Lehrberger, J. (eds). 1982. *Sublanguage: Studies of language in restricted semantic domains.* Berlin and New York: Mouton de Gruyter.

[Koller, 1984] Koller, W. 1984. Übersetzungen ins Deutsche und ihre Bedeutung für die deutsche Sprachgeschichte. *Pages 112–129 of:* Besch, W., Reichmann, O., & Sonderegger, S. (eds), *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung.*

[Königs, 1987] Königs, F. G. 1987. Was beim Übersetzen passiert. Theoretische Aspekte, empirische Befunde und praktische Konsequenzen. *Pages 162–185 of: Die Neueren Sprachen*, vol. 86. no. 2.

[Königs, 2000] Königs, K. 2000. *Übersetzen Englisch-Deutsch: systemischer Ansatz.* München and Berlin: Oldenbourg.

[Krings, 1986a] Krings, H. P. 1986a. Translation problems and translation strategies of advanced German learners of French (L2). *Pages 257–270 of:* House, J., & Blum-Kulka, S. (eds), *Interlingual and intercultural communication. Discourse and cognition in translation and second language acquisition studies.* Tübingen: Gunter Narr.

[Krings, 1986b] Krings, H. P. 1986b. *Was in den Köpfen von Übersetzern vorgeht. Eine empirische Untersuchung zur Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern.* Tübingen: Gunter Narr.

[Krings, 1987] Krings, H. P. 1987. The use of introspective data in translation. *Pages 159–176 of:* Færch, C., & Kasper, G. (eds), *Introspection in second language research.* Clevedon: Multilingual Matters.

[Kurz, 2000] Kurz, D. 2000. A statistical account on word order variation in German. *In: Proceedings of the COLING Workshop on Linguistically Interpreted Corpora.*

[Kurz et al., 2000] Kurz, D., Skut, W., & Uszkoreit, H. 2000. German factors constraining word order variation. *In: Thirteenth Annual Conference on Human Sentence Processing CUNY 2000.*

[Kussmaul & Tirkkonen-Condit, 1995] Kussmaul, P., & Tirkkonen-Condit, S. 1995. Think-Aloud Protocol Analysis in Translation Studies. *Pages 177–199 of: TTR*, vol. 8. no. 1.

[Kutas & van Petten, 1994] Kutas, M., & van Petten, C. K. 1994. Psycholinguistics Electrified: Event-Related Brain Potential Investigations. *Pages 83–143 of:* Gernsbacher, M. A. (ed), *Handbook of Psycholinguistics*. San Diego: Academic Press.

[Lang & Zifonun, 1996] Lang, E., & Zifonun, G. (eds). 1996. *Deutsch – typologisch*. Berlin and New York: Mouton de Gruyter.

[Laviosa, 1997] Laviosa, S. 1997. How Comparable Can 'Comparable Corpora' Be? *Pages 289–319 of: Target*, vol. 9. no. 2.

[Laviosa, 1998a] Laviosa, S. 1998a. Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose. *Pages 557–570 of: Meta*, vol. 43. no. 4.

[Laviosa, 1998b] Laviosa, S. 1998b. The English Comparable Corpus (ECC): A Resource and a Methodology. *In:* Bowker, L., Cronin, M., Kenny, D., & Pearson, J. (eds), *Unity in Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome.

[Laviosa-Braithwaite, 1996a] Laviosa-Braithwaite, S. 1996a. Comparable corpora: towards a corpus linguistic methodology for the empirical study of translation. *Pages 153–163 of:* Thelen, M., & Lewandowska-Tomaszczyk, B. (eds), *Translation and Meaning Part 3. Proceedings of the Maastricht Session of the 2nd International Maastricht-Lodz Duo Colloquium on "Translation and Meaning"*. Maastricht: Universitaire Pers Maastricht.

[Laviosa-Braithwaite, 1996b] Laviosa-Braithwaite, S. 1996b. *The English Comparable Corpus (ECC): A Resource and a Methodology for the Empirical Study of Translation*. Ph.D. thesis, UMIST, Manchester.

[Laviosa-Braithwaite, 1997] Laviosa-Braithwaite, S. 1997. Investigating Simplification in an English Comparable Corpus of Newspaper Articles. *Pages 531–540 of:* Klaudy, K., & Kohn, J. (eds), *Transferre Necesse Est. Proceedings of the 2nd International Conference on Current Trends in Studies of Translation and Interpreting*. Budapest: Scholastica.

[Laviosa-Braithwaite, 1998] Laviosa-Braithwaite, S. 1998. Universals of translation. *Pages 288–291 of:* Baker, M. (ed), *Routledge Encyclopedia of Translation Studies*. London and New York: Routledge.

[Lee, to appear] Lee, D. to appear. *Modelling Variation in Spoken And Written English: the Multi-Dimensional Approach Revisited*. Ph.D. thesis, Lancaster University, Lancaster. to appear as Routledge monography.

[Lively et al., 1994] Lively, S. E., Pisoni, D. B., & Goldinger, S. D. 1994. Spoken Word Recognition: Research and Theory. *Pages 265–301 of:* Gernsbacher, M. A. (ed), *Handbook of Psycholinguistics*. San Diego: Academic Press.

[Lörscher, 1991a] Lörscher, W. 1991a. Investigating the translation process. *Pages 3–22 of: Interface*, vol. 6.

[Lörscher, 1991b] Lörscher, W. 1991b. Process-oriented research into translation and implications for translation teaching. *In: LAUD*, vol. 227. Duisburg: Linguistic Agency, University of Duisburg. series B.

[Lörscher, 1991c] Lörscher, W. 1991c. *Translation Performance, Translation Process, and Translation Strategies. A Psycholinguistic Investigation*. Tübingen: Gunter Narr.

[Lucas, 1987] Lucas, M. 1987. Frequency effects on the processing of ambiguous words in sentence context. *Pages 25–46 of: Language and Speech*, vol. 30.

[Luce, 1986] Luce, P. A. 1986. *Neighborhoods of words in the mental lexicon*. Ph.D. thesis, Indiana University, Bloomington.

[Malinowski, 1935] Malinowski, B. 1935. *Coral Gardens and their Magic*. London: Allen and Unwin.

[Marcus et al., 1993] Marcus, G. P., Santorini, B., & Marcinkiewicz, M. A. 1993. Building a large annotated corpus of English: The Penn Treebank. *Pages 313–330 of: Computational Linguistics*, vol. 19. no. 2.

[Marslen-Wilson, 1985] Marslen-Wilson, W. D. 1985. Speed shadowing and speech comprehension. *Pages 55–73 of: Speech Communication*, vol. 4.

[Marslen-Wilson & Tyler, 1980] Marslen-Wilson, W. D., & Tyler, L. K. 1980. The temporal structure of spoken language understanding. *Pages 1–71 of: Cognition*, vol. 8.

[Martin, 1992] Martin, J. R. 1992. *English text: systems and structure.* Amsterdam: Benjamins.

[Matthiessen, 1993] Matthiessen, C. M. I. M. 1993. Register in the round, or diversity in a unified theory of register. *Pages 221–292 of:* Ghadessy, M. (ed), *Register Analysis. Theory and Practice.* London: Pinter.

[Matthiessen, 1995] Matthiessen, C. M. I. M. 1995. *Lexicogrammatical cartography: English systems.* Tokyo, Taipei and Dallas: International Language Science Publishers.

[Matthiessen, 2001] Matthiessen, C. M. I. M. 2001. The environments of translation. *Pages 41–124 of:* Steiner, E., & Yallop, C. (eds), *Exploring Translation and Multilingual Text Production: Beyond Content. Series Text, Translation, Computational Processing.* Berlin and New York: Mouton de Gruyter.

[Matthiessen & Bateman, 1991] Matthiessen, C. M. I. M., & Bateman, J. 1991. *Text generation and systemic-functional linguistics: Experiences from English and Japanese.* London: Pinter.

[McKelvie et al., 2001] McKelvie, D., Isard, A., Mengel, A., Moller, M.B., Grosse, M., & Klein, M. 2001. The MATE workbench — An annotation tool for XML coded speech corpora. *Pages 97–112 of: Speech Communication*, vol. 33. no. 1–2.

[McKoon & Ratcliff, 1980] McKoon, G., & Ratcliff, R. 1980. Priming in item recognition: The organization of propositions in memory for text. *Pages 269–386 of: Journal of Verbal Learning and Verbal Behavior*, vol. 19.

[McKoon & Ratcliff, 1986] McKoon, G., & Ratcliff, R. 1986. Inferences about predictable events. *Pages 82–91 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 12.

[McKoon & Ratcliff, 1990] McKoon, G., & Ratcliff, R. 1990. Dimensions of inference. *Pages 313–328 of:* Graesser, A. C., & Bower, G. (eds), *Inferences and text comprehension.* San Diego: Academic Press.

[Melchuk, 1988] Melchuk, I. 1988. *Dependency Syntax: Theory and Practice.* Albany: State University of New York Press.

[Mengel, 1999] Mengel, A. 1999. Die integrierte Repräsentation linguistischer Daten. *Pages 115–121 of:* Gippert, J. (ed), *Multilinguale Corpora. Codierung, Strukturierung und Analyse (Sammelband der Jahrestagung der GLDV 99).* Prag: enigma corporation.

[Mengel & Lezius, 2000] Mengel, A., & Lezius, W. 2000. An XML-based representation format for syntactically annotated corpora. *Pages 121–126 of: Proceedings of LREC.*

[Mitchell, 1984] Mitchell, D. C. 1984. An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. *Pages 69–89 of:* Kieras, D. E., & Just, M. A. (eds), *New methods in reading comprehension research.* Hillsdale: Erlbaum.

[Morrow et al., 1990] Morrow, D. G., Bower, G. H., & Greenspan, S. L. 1990. Situation-based inferences during narrative comprehension. *Pages 123–135 of:* Graesser, A. C., & Bower, G. (eds), *Inferences and text comprehension.* San Diego: Academic Press.

[Newmark, 1988] Newmark, P. 1988. *A textbook of translation.* London: Prentice Hall.

[Nida, 1964] Nida, E. 1964. *Toward a Science of Translating.* Leiden: E. J. Brill.

[Oakes, 1998] Oakes, M. P. 1998. *Statistics for Corpus Linguistics.* Edinburgh: University of Tübingen. Edinburgh University Press.

[O'Brien, 1987] O'Brien, E. J. 1987. Antecedent search processes and the structure of text. *Pages 278–290 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 13.

[O'Brien & Albrecht, 1991] O'Brien, E. J., & Albrecht, J. E. 1991. The role of context in accessing antecedents in text. *Pages 94–102 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 17.

[O'Brien et al., 1986] O'Brien, E. J., Duffy, S. A., & Myers, J. L. 1986. Anaphoric inference during reading. *Pages 346–352 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 12.

[O'Brien et al., 1988] O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. 1988. Elaborative inferences during reading: Do they occur on-line? *Pages 410–420 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 14.

[O'Brien et al., 1990] O'Brien, E. J., Plewes, P. S., & Albrecht, J. E. 1990. Antecedent retrieval processes. *Pages 241–249 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 16.

[O'Donnell, 1995] O'Donnell, M. 1995. From Corpus to Codings: Semi-Automating the Acquisition of Linguistic Features. *Pages 120–124 of: Proceedings of the AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*. Standford: Stanford University.

[Olohan, 2001] Olohan, M. 2001. Spelling out the optionals in translation: A corpus study. *Pages 423–432 of:* Rayson, P., Wilson, A., McEnery, T., Hardie, A., & Khoja, S. (eds), *Proceedings of the Corpus Linguistics 2001 conference. UCREL Technical Paper number 13. Special issue.*

[Onifer & Swinney, 1981] Onifer, W., & Swinney, D. A. 1981. Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Pages 225–236 of: Memory and Cognition*, vol. 9.

[Pickering & Branigan, 1998] Pickering, M. J., & Branigan, H. P. 1998. The representation of verbs: Evidence from syntactic priming in language production. *Pages 633–651 of: Journal of Memory and Language*, vol. 39.

[Pöchhacker, 1994] Pöchhacker, F. 1994. *Simultandometschen als komplexes Handeln*. Tübingen: Gunter Narr.

[Polenz, 1988] Polenz, P. 1988. *Deutsche Satzsemantik*. Berlin: Mouton de Gruyter.

[Potts et al., 1985] Potts, G. R., Keenan, J. M., & Golding, J. M. 1985. Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Pages 399–415 of: Journal of Memory and Language*, vol. 27.

[Quirk et al., 1985] Quirk, R., Greenbaum, S., Leech, G., & Svartik, J. 1985. *A comprehensive grammar of the English language.* London: Longman.

[Ratcliff & McKoon, 1978] Ratcliff, R., & McKoon, G. 1978. Priming in item recognition: Evidence for the propositional structure of sentences. *Pages 403–417 of: Journal of Verbal Learning and Verbal Behavior*, vol. 17.

[Rayner et al., 1989] Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. J. 1989. Eye movements and on-line language comprehension processes. *Pages SI21–SI49 of: Language and Cognitive Processes*, vol. 4.

[Sager, 1984] Sager, J. C. 1984. Reflections on the didactic implications of an extended theory of translation. *Pages 333–343 of:* Wilss, W., & Thome, G. (eds), *Die Theorie des Übersetzens und ihr Aufschlußwert für die Übersetzungs- und Dolmetschdidaktik /Translation theory and its implementation in the teaching of translating and interpreting.* Tübingen: Gunter Narr.

[Sager, 1994] Sager, J. C. 1994. *Language engineering and translation: consequences of automation.* Amsterdam: John Benjamins.

[Sampson, 1995] Sampson, G. 1995. *English for the Computer.* Oxford: Oxford University Press.

[Savin, 1963] Savin, H. B. 1963. Word-frequency effect and errors in the perception of speech. *Pages 200–206 of: Journal of the Acoustical Society of America*, vol. 35.

[Scheepers & Corley, 2000] Scheepers, C., & Corley, M. 2000. Syntactic Priming in German Sentence Production. *Pages 435–440 of:* Gleitman, L. R., & Joshi, A K. (eds), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society.* Mahawah: Erlbaum.

[Schiller et al., 1999] Schiller, A., Teufel, S., & Stöckert, C. 1999. *Guidelines für das Tagging deutscher Textcorpora mit STTS*. Stuttgart and Tübingen: University of Stuttgart, University of Tübingen. Technical report.

[Schmidt et al., 1989] Schmidt, A. L., Arthur, D. L., Kutas, M., George, J., & Flynn, E. 1989. Neuromagnetic responses evoked during reading meaningful and meaningless sentences. *Page S6 of: Psychophysiology*, vol. 26.

[Schriefers et al., 1991] Schriefers, H., Zwitserlood, P., & Roelofs, A. 1991. The identification of morphologically complex spoken words: Continuous processing or decomposition? *Pages 26–47 of: Journal of Memory and Language*, vol. 30.

[Schuberth & Eimas, 1977] Schuberth, R. E., & Eimas, P. D. 1977. Effects of context on the classification of words and nonwords. *Pages 27–36 of: Journal of Experimental Psychology: Human Perception and Performance*, vol. 3.

[Schustack et al., 1987] Schustack, M., Ehrlich, S. F., & Rayner, K. 1987. The complexity of contextual facilitation in reading: Local and global influences. *Pages 322–340 of: Journal of Memory and Language*, vol. 26.

[Séguinot, 1991] Séguinot, C. 1991. A Study of Student Translation Strategies. *In:* Tirkkonen-Condit, S. (ed), *Empirical Research in Translation and Intercultural Studies*. Tübingen: Gunter Narr.

[Sharkey & Mitchell, 1985] Sharkey, N. E., & Mitchell, D. C. 1985. Word recognition in a functional context: The use of scripts in reading. *Pages 253–270 of: Journal of Memory and Language*, vol. 24.

[Simpson & Krueger, 1991] Simpson, G. B., & Krueger, M. A. 1991. Selective access of homograph meanings in sentence context. *Pages 627–643 of: Journal of Memory and Language*, vol. 30.

[Slobin, 1966] Slobin, D. I. 1966. Grammatical transformations and sentence comprehension in childhood and adulthood. *Pages 219–227 of: Journal of Verbal Learning and Verbal Behavior*, vol. 5.

[Slowiaczek, 1981] Slowiaczek, M. L. 1981. *Prosodic units as language processing units*. Ph.D. thesis, University of Massachusetts, Amherst.

[Snell-Hornby, 1988] Snell-Hornby, M. 1988. *Translation Studies – An Integrated Approach*. Amsterdam: John Benjamins.

[Solso, 1995] Solso, R. L. 1995. *Cognitive psychology*. Boston: Allyn and Bacon.

[Speelman & Kirsner, 1990] Speelman, C. P., & Kirsner, K. 1990. The representation of text-based and situation-based information in discourse comprehension. *Pages 119–132 of: Journal of Memory and Language*, vol. 29.

[Sperberg-McQueen & Huitfeldt, 2001] Sperberg-McQueen, C. M., & Huitfeldt, C. 2001. GODDAG: A Data Structure for Overlapping Hierarchies. *In: Proceedings of PODDP'00 and DDEP'00*.

[Steiner, 1997] Steiner, E. 1997. An extended register analysis as a form of text analysis for translation. *Pages 235–256 of:* Wotjak, G., & Schmidt, H. (eds), *Modelle der Translation – Models of Translation. Leipziger Schriften zur Kultur-, Literatur-, Sprach- und Übersetzungswissenschaft.*

[Steiner, 1998] Steiner, E. 1998. A register-based translation evaluation: An advertisement as a case in point. *Pages 291–318 of: Target*, vol. 10. no. 2.

[Steiner, 2001a] Steiner, E. 2001a. Intralingual and Interlingual Versions of a Text – How Specific is the Notion of Translation? *Pages 161–190 of:* Steiner, E., & Yallop, C. (eds), *Exploring Translation and Multilingual Text Production: Beyond Content. Series Text, Translation, Computational Processing*. Berlin and New York: Mouton de Gruyter.

[Steiner, 2001b] Steiner, E. 2001b. Translations English-German: investigating the relative importance of systemic contrasts and of the text-type "translation". *In: SPRIKreport no. 7. Reports from the project "Languages in Contrast"*. Oslo: University of Oslo.

[Steiner, to appear] Steiner, E. to appear. Grammatical metaphor in translation – some methods for corpus-based investigations. *In:* Behrens, B.,

Fabricius-Hansen, C., Hasselgård, H., & Johansson, S. (eds), *Information structure in a cross-linguistic perspective.* Amsterdam: Rodopi.

[Steiner & Teich, to appear] Steiner, E., & Teich, E. to appear. Metafunctional profile: German. *In:* Caffarel, A., Matthiessen, C. M. I. M., & Martin, J. R. (eds), *Systemic Functional Typology.* Amsterdam: Benjamins.

[Sturt et al., 1999] Sturt, P., Pickering, M. J., & Crocker, M. W. 1999. Structural Change and Reanalysis Difficulty in Language Comprehension. *Pages 136–150 of: Journal of Memory and Language*, vol. 40.

[Swinney, 1979] Swinney, D. A. 1979. Lexical access during sentence comprehension: (Re)consideration of context effects. *Pages 645–659 of: Journal of Verbal Learning and Verbal Behavior*, vol. 18.

[Tanenhaus et al., 1979] Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. 1979. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Pages 427–440 of: Journal of Verbal Learning and Verbal Behavior*, vol. 18.

[Teich, 1995a] Teich, E. 1995a. *A proposal for dependency in Systemic Functional Grammar: metasemiosis in Computational Systemic Functional Linguistics.* Ph.D. thesis, Saarland University, Saarbrücken.

[Teich, 1995b] Teich, E. 1995b. Towards a methodology for the construction of multilingual resources for multilingual generation. *Pages 136–148 of: Proceedings of the IJCAI workshop on multilingual generation. International Joint Conference on Artificial Intelligence (IJCAI).*

[Teich, 1999] Teich, E. 1999. Contrastive Linguistics and Translation Studies Revisited. *Pages 507–521 of:* Gil, A., Haller, J., Steiner, E., & Gerzymisch-Arbogast, H. (eds), *Modelle der Translation: Grundlagen für Methodik, Bewertung, Computermodellierung.* Frankfurt: Peter Lang.

[Teich, 2001a] Teich, E. 2001a. *English - German contrast and commonality in system and text. A methodology for the investigation of parallel and multilingually comparable texts. Habilitationsschrift.* Saarbrücken: Saarland University.

[Teich, 2001b] Teich, E. 2001b. Towards a model for the description of cross-linguistic divergence and commonality in translation. *In:* Steiner, E., & Yallop, C. (eds), *Exploring Translation and Multilingual Text Production: Beyond Content. Series Text, Translation, Computational Processing.* Berlin and New York: Mouton de Gruyter.

[Teich, to appear] Teich, E. to appear. System-oriented and text-oriented comparative linguistic research: cross-linguistic variation in translation. *In: Languages in Contrast.*

[Teich & Hansen, 2001a] Teich, E., & Hansen, S. 2001a. Methods and techniques for a multi-level analysis of multilingual corpora. *Pages 572–580 of:* Rayson, P., Wilson, A., McEnery, T., Hardie, A., & Khoja, S. (eds), *Proceedings of the Corpus Linguistics 2001 conference. UCREL Technical Paper number 13. Special issue.*

[Teich & Hansen, 2001b] Teich, E., & Hansen, S. 2001b. Towards an integrated representation of multiple layers of linguistic annotation in multilingual corpora. *In: Online Proceedings of Computing Arts 2001: Digital Resources for Research in the Humanities.*

[Teich et al., 1996] Teich, E., Degand, L., & Bateman, J. 1996. Multilingual textuality: Experiences from multilingual text generation. *Pages 331–349 of:* Adorni, G., & Zock, M. (eds), *Trends in Natural Language Generation: An artificial intelligence perspective.* Berlin and New York: Springer.

[Teich et al., 2001] Teich, E., Hansen, S., & Fankhauser, P. 2001. Representing and querying multi-layer annotated corpora. *In: Proceedings of IRCS Workshop on Linguistic Databases.*

[Tesnière, 1959] Tesnière, L. 1959. *Eléments de syntaxe structurale.* Paris: Klincksieck.

[Thompson & McKelvie, 1997] Thompson, H. S., & McKelvie, D. 1997. Hyperlink semantics for standoff markup of read-only documents. *In: Proceedings of SGML Europe.*

[Till et al., 1988] Till, R. E., Mross, E. F., & Kintsch, W. 1988. Time course of priming for associate and inference words in a discourse context. *Pages 283–299 of: Memory and Cognition*, vol. 16.

[Tirkkonen-Condit, 1989] Tirkkonen-Condit, S. 1989. Professional vs. Non-Professional Translation: A Think-Aloud Protocol Study. *In:* Séguinot, C. (ed), *The Translation Process*. Toronto: H. G. Publications.

[Tirkkonen-Condit, 1997] Tirkkonen-Condit, S. 1997. Who Verbalises What: A Linguistic Analysis of TAP Texts. *Pages 69–84 of: Target*, vol. 9. no. 1.

[Toury, 1995] Toury, G. 1995. *Descriptive translation studies and beyond*. Amsterdam: John Benjamins.

[Tyler & Warren, 1987] Tyler, L. K., & Warren, P. 1987. Local and global structure in spoken language comprehension. *Pages 638–657 of: Journal of Memory and Language*, vol. 26.

[Ure & Ellis, 1977] Ure, J. N., & Ellis, J. 1977. Register in descriptive linguistics and linguistic sociology. *In:* Uribe-Villegas, O. (ed), *Issues in Sociolinguistics*. The Hague: Mouton.

[Uszkoreit et al., 1998] Uszkoreit, H., Brants, T., Duchier, D., Krenn, B., Konieczny, L., Oepen, S., & Skut, W. 1998. Studien zur performanzorientierten Linguistik: Aspekte der Relativsatzextraposition. *Pages 125–133 of: Kognitionswissenschaft*, vol. 7.

[van Petten & Kutas, 1990] van Petten, C. K., & Kutas, M. 1990. Interactions between sentence context and word frequency in event-related brain potentials. *Pages 380–393 of: Memory and Cognition*, vol. 18.

[van Petten & Kutas, 1991] van Petten, C. K., & Kutas, M. 1991. Electrophysiological evidence for the flexibility of lexical processing. *Pages 129–184 of:* Simpson, G. (ed), *Word and sentence*. Amsterdam: North-Holland.

[Vinay & Darbelnet, 1995] Vinay, J.-P., & Darbelnet, J. 1995. *Comparative stylistics of French and English: a methodology for translation*. Amsterdam and Philadelphia: John Benjamins.

[Warren, 1977] Warren, R. E. 1977. Time and the spread of activation in memory. *Pages 458–466 of: Journal of Experimental Psychology: Human Learning and Memory*, vol. 3.

[Warren, 1970] Warren, R. M. 1970. Perceptual restoration of missing speech sounds. *Pages 392–393 of: Science*, vol. 176.

[West & Stanovich, 1982] West, R. F., & Stanovich, K. E. 1982. Source of inhibition in experiments on the effect of sentence context on word recognition. *Pages 385–399 of: Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 8.

[Whorf, 1956] Whorf, B. L. 1956. *Language, Thought and Reality. Selected Writings of B. L. Whorf.* Cambridge: MIT Press. Carroll, J. B. (ed).

[Wikberg et al., 1999] Wikberg, S. Johansson K., Stenström, A.-B., & Virtanen, T. 1999. Corpus studies in applied linguistics. *In:* Pietilä, P., & Salo, O.-P. (eds), *Multiple Languages – Multiple Perspectives.*

[Wilss, 1977] Wilss, W. 1977. *Übersetzungswissenschaftliche Probleme und Methoden.* Stuttgart: Klett.

[Zola, 1984] Zola, D. 1984. Redundancy and word perception during reading. *Pages 277–284 of: Perception and Psychophysics*, vol. 36.

Chapter 7

Appendix

7.1 Appendix 1: TEC and BNC files

TEC files included in the comparable corpus:

- fn000001.txt/Restless Nights/short stories/by Dino Buzzati/translated from Italian by Lawrence Venuti
- fn000002.txt/The Siren/short stories/by Dino Buzzati/translated from Italian by Lawrence Venuti
- fn000003.txt/Turbulence/by Chico Buarque/translated from Brazilian Portuguese by Peter Bush
- fn000004.txt/The Stone Raft/by José Saramago/translated from Portuguese by Giovanni Pontiero
- fn000005.txt/The History of the Siege of Lisbon/by José Saramago/translated from Portuguese by Giovanni Pontiero
- fn000006.txt/Discovering the World/by Clarice Lispector/translated from Brazilian Portuguese by Giovanni Pontiero
- fn000007.txt/The Gospel according to Jesus Christ/by José Saramago/translated from Portuguese by Giovanni Pontiero
- fn000008.txt/The Stone of Laughter/by Hoda Barakat/translated from Arabic by Sophie Bennett
- fn000009.txt/Memoirs of Leticia Valle/by Rosa Chacel/translated from Spanish by Carol Maier
- fn000010.txt/Passion/by I. U. Tarchetti/translated from Italian by Lawrence Venuti
- fn000011.txt/Lucio's Confession/by Mario De Sá Carneiro/translated from Portuguese by Margaret Jull Costa

- fn000012.txt/Fantastic Tales/short stories mainly/by I. U. Tarchetti/translated mainly from Italian by Lawrence Venuti
- fn000013.txt/The Dedalus Book of Surrealism/short stories/by different authors/translated mainly from French and edited by Michael Richardson
- fn000014.txt/Violetta/by Pieke Biermann/translated from German by I. Rieder and J. Hannum
- fn000015.txt/The Head of Vitus Bering/by Konrad Bayer/translated from German by W. Billeter
- fn000016.txt/Worlds of Difference/by Georges Goldshmi/translated from German by J. Kirkup
- fn000017.txt/Our Conquest/by Gert Hofmann/translated from German by C. Middleton
- fn000018.txt/Wonderful/Wonderful Times/by Elfriede Jelinek/translated from German by M. Hulse
- fn000019.txt/Lust/by Elfriede Jelinek/translated from German by M. Hulse
- fn000020.txt/Infanta/by Bodo Kirchhoff/translated from German by J. M. Brownjohn
- fn000021.txt/The Monument/by Erich Loest/translated from German by I. Mitchell
- fn000022.txt/The Terrors of Ice and Darkness/by Christoph Ransmay/translated from German by J. Woods
- fn000023.txt/Winner takes all/by Dieter Wellershoff/translated from German by P. Knight
- fn000024.txt/The Dedalus Book of French Horror/short stories/by different authors/translated from French by Terry Hale and Liz Heron
- fn000025.txt/Sacred Cow/by Diamela Eltit/translated from Spanish by Amanda Hopkinson
- fn000026.txt/The Compassion Protocol/by Hervé Guibert/translated from French by James Kirkup
- fn000027.txt/To The Friend Who Did Not Save My Life/by Hervé Guibert/translated from French by Linda Coverdale
- fn000028.txt/Quarantine/by Juan Goytisolo/translated from Spanish by Peter Bush
- fn000029.txt/For Every Sin/by Gabriel Josipovici/translated from Hebrew by Jeffrey M. Green
- fn000030.txt/A Mortal Sin/by Giovanni Verga/translated by Iain Halliday

- fn000031.txt/The Golden Chariot/by Salwa Bakr/translated from Arabic by Dinah Manisty
- fn000032.txt/Justice/by Bernard Evans/translated from Welsh by Marion Eames
- fn000033.txt/The Voice of the Turtle/short stories/by different authors/ translated from Cuban Spanish
- fn000034.txt/The Prostitute/by K. Surangkhanang/translated from Thai by David Smyth
- fn000035.txt/Strawberry and Chocolate/by Senel Paz/translated from Spanish by Peter Bush
- fn000036.txt/Manual of Painting and Calligraphy/by José Saramago/translated from Portuguese by Giovanni Pontiero
- fn000037.txt/The Hour of the Star/by Clarice Lispector/translated from Portuguese by Giovanni Pontiero
- fn000038.txt/The Dark Domain/by Stefan Grabinski/translated from Polish by Miroslaw Lipinski
- fn000039.txt/Mothballs/by Alia Mamdouh/translated from Arabic by Peter Theroux
- fn000040.txt/The Eye of the Mirror/by Liana Badr/translated from Arabic by Samira Kawar
- fn000041.txt/Jerome/by Franco Troiano/translated from Italian by Allan Riger-Brown
- fn000042.txt/Declares Pereira/by Antonio Tabucchi/translated from Italian by Patrick Creagh
- fn000043.txt/Eden, Eden, Eden/by Pierre Guyotat/translated from French by Graham Fox
- fn000044.txt/The Japanese Chronicles/by Nicolas Bouvier/translated from French by Anne Dickerson
- fn000045.txt/Feet in Chains/by Kate Roberts/translated from Welsh by Idwal Walters
- fn000046.txt/The Secret Room/by Marion Eames/translated from Welsh
- fn000047.txt/The Golden Road/by Marion Eames/translated from Welsh
- fn000048.txt/Dubai Tales/by Muhammad al Murr/translated from Arabic by Peter Clark
- fn000049.txt/Grandfather's Tale/by Muhammad al Murr/translated from Arabic by Peter Clark

- fn000050.txt/Sabriya/by Ulfat Idilbi/translated from Arabic by Peter Clark
- fn000051.txt/The Woman Watching/by Paola Capriolo/translated from Italian by Liz Heron
- fn000052.txt/The Emigrants/by W.G. Sebald/translated from German by Michael Hulse
- fn000053.txt/Artemisia/by Anna Barti/translated from Italian by Shirley D'Ardia Caracciolo
- fn000054.txt/Far from Madina/by Assia Djebar/translated from French by Dorothy S. Blair
- fn000055.txt/Women of Sand Myrrh/by Hanan al-Shaykh/translated from Arabic by Catherine Cobham
- fn000056.txt/City of Saffron/by Edwar al-Kharrat/translated from Arabic by Fraces Liardet
- fn000057.txt/State of Absence/by Tahar Ben Jelloun/translated from Arabic by James Kirkup
- fn000058.txt/Cutting Timber/by Thomas Bernhard/translated from German by Ewald Osers
- fn000059.txt/The Sand Child/by Tahar Ben Jelloun/translated from Arabic by Alan Sheridan
- fn000060.txt/Silent Day in Tangier/by Tahar Ben Jelloun/translated from Arabic by David Lobdell
- fn000061.txt/Respected Sir/by Naguib Mahfouz/translated from Arabic by Rasheed El-Enany
- fn000062.txt/A Sister to Scheherazade/by Assia Djebar/translated from French by Dorothy S. Blair
- fn000063.txt/Endings/by 'Abd al-Rahman Munif/translated from Arabic by Roger Allen
- fn000064.txt/Bedside Manners/by Luisa Valenzuela/translated from Spanish by Margaret Jull Costa
- fn000065.txt/Cousin Bazilio/by Ea de Queiroz/translated from Portuguese by Joe Staines
- fn000066.txt/Fantasia: An Algerian Cavalcade/by Assia Djebar/translated from French by Dorothy S. Blair
- fn000067.txt/The Flight of the Tiger/by Daniel Moyano/translated from Spanish by Norman Thomas di Giovanni

- fn000068.txt/The story of Zahra/by Hanan Al-Zhaykh/translated from Arabic by Peter Ford
- fn000069.txt/Season of Migration to the North - The Wedding of Zein/by Tayeb Salih/translated from Arabic by Denys Johnson-Davies
- fn000070.txt/My Grandmother's Cactus/Stories/by Egyptian Women/translated from Arabic by Marilyn Booth
- fn000071.txt/Beirut Nightmares/by Ghada Samman/translated from Arabic by Nancy N. Roberts
- fn000072.txt/Elissa/by Fawzi Mellah/translated from French by Howard Curtis
- fn000073.txt/Down to the Sea/by Gamil Atia Ibrahim/translated from Arabic by Frances Liardet
- fn000074.txt/The Honour of the Tribe/by Rachid Mimouni/translated from French by Joachim Neugroschel
- fn000075.txt/Milk for the Orange Tree/by Gisèle Halimi/introduced and translated from French by Dorothy S. Blair
- fn000076.txt/The rock of Tanios/by Amin Maalouf/translated from French by Dorothy S. Blair
- fn000077.txt/The Game of Forgetting/by Mohamed Berrada/translated from Arabic by Issa J. Boullata
- fn000078.txt/The Veil of Silence/by Djura/translated from French by Dorothy S. Blair

BNC files included in the comparable corpus:

- A0D/A classic English crime/Heald, Tim (ed)
- A0F/Part of the furniture/Falk, Michael
- A0L/Jay loves Lucy/Cooper, Fiona
- A0N King Cameron/Craig, David
- A0R/Nudists may be encountered/Scott, Mary
- A0U/So very English/Rowe, Martha (ed)
- A1C/Henry's leg/Pilling, Ann
- A6J/Three times table/Maitland, Sara
- A6N/Amongst women/McGahern, J
- A73/The best man to die/Rendell, Ruth
- A74/Billy Bayswater/Watts, Nigel
- A7A/Bury the dead/Carter, Peter
- A7J/The charnel house/McGrath, Eamonn
- AB9/Death of a partner/Neel, Janet
- ABS/ESQUIRE April 1991/The National Magazine Company Ltd(pub)
- ABW/Jane's Journey/Bow, Jean
- ABX/Jubilee wood/Hassall, Angela
- AC2/Man at the sharp end/Kilby, M
- AC3/Misfortunes of Nigel/Pitt-Kethley, Fiona
- AC4/On the edge/Cross, Gillian
- AC5/Paper faces/Anderson, Rachel
- AC6/A poet could not but be gay/Kirkup, James
- AC7/The reluctant Samaritan/Beechey, Winifred
- ACB/The lock/Gates, Susan
- ACE/Willoughby's phoney war/Fox, William
- ACK/Dandelion days/Howell, Bette
- ACV/The forest of the night/Kelly, Chris
- ACW/Frankie/Highsmith, Domini

- AD1/Gentleman and ladies/Hill, Susan
- AD9/Hermetech/Constantine, Storm
- ADA/Hide and seek/Potter, Dennis
- ADS/Lady's maid/Forster, Margaret
- ADY/Longshot/Francis, Dick
- AE8/Roads that move/Perrie, Walter
- AEA/Tomorrow/Taylor, Elizabeth Russell
- AEB/A twist of fate/Scobie, Pamela
- C85/The first of midnight/Darke, Marjorie
- C86/The five gates of hell/Thomson, Rupert
- C8D/Black justice/Shepherd, Stella
- C8E/The Buddha of suburbia/Kureishi, Hanif
- C8S/Destiny/Beauman, Sally
- C8T/Devices and desires/James, P D
- C98/Fields in the sun/Sunley, Margaret
- CA0/Polo/Cooper, Jilly
- CA3/Lee's ghost/Pulsford, Petronella
- CAM/The negotiator/Forsyth, Frederick
- CB5/Ruth Appleby/Rhodes, Elvi
- CB7/A sensible life/Wesley, Mary
- CBS/The fifth child/Lessing, Doris
- CCD/The child bride/Wiat, Philippa
- CCW/Crackdown/Cornwell, Bernard
- CD1/Distant relations/Conlon, Kathleen
- CD2/A dark star passing/Hill, Pamela
- CDE/Guilty parties/Street, Pamela
- CDN/The latchkey kid/Forrester, Helen
- CDY/No enemy but time/Anthony, Evelyn
- CEB/A season for murder/Granger, Ann

- CEC/Sons of heaven/Strong, Terence
- CEH/All the sweet promises/Elgin, Elizabeth
- CEU/Wings/Pratchett, Terry
- CEX/A woman of my age/Bawden, Nina
- CEY/A woman of style/McDowell, Colin
- CFY/My beloved son/Cookson, Catherine
- CH0/Krokodil tears/Yeovil, Jack
- CHG/The other side of paradise/Barber, Noel
- CJA/Take back plenty/Grafton Books(pub)
- CJF/A taste for death/James, P D
- CJJ/Warhammer 40,000: space marine/Watson, Ian
- CJW/Command of the king/Lide, Mary
- CJX/Death in springtime/Nabb, Magdalen
- CK8/Peach/Adler, Elizabeth
- CK9/The rag nymph/Cookson, Catherine
- CKB/The raven on the water/Taylor, Andrew
- CKC/Santorini/MacLean, Alistair
- CKD/The shoemaker's daughter/Gower, Iris
- CKE/Sergeant Joe/Staples, Mary Jane
- CKF/Stormy petrel/Stewart, Mary
- CL3/Cara's land/Rhodes, Elvi
- CLD/Condition black/Seymour, Gerald
- CLJ/The deceiver/Forsyth, Frederick
- CLK/Castle Drachenfels/Sargent, Carl
- CM1/High elves/King, Bill
- CM4/Inquisitor/Watson, Ian
- CM7/Lie down with lions/Follett, Ken
- CML/Ring of red roses/Shah, Eddy
- CMP/Sharpe's Waterloo/Cornwell, Bernard

- CN3/Whirlpool/Forbes, Colin
- CS4/Sudden death/Hamer, Malcolm
- EA5/The girl from cotton lane/Bowling, Harry
- ECK/Death train/MacNeill, Alastair
- ECP/Duncton tales/Horwood, W
- EDJ/Talking it over/Barnes, J
- EDN/King Solomon's carpet/Vine, Barbara
- EDV/A matter of honour/Archer, Jeffrey
- EE5/Mouthful of rocks/Jennings, C
- EEW/In sunshine or in shadow/Bingham, C
- EF1/Time of the assassins/MacNeill, Alastair
- EF7/A twist in the tale/Archer, Jeffrey
- EFJ/Carrie's War/Bawden, Nina
- EFP/Jerusalem the golden/Drabble, Margaret
- EFW/The siege Of Krishnapur/Farrell, J G
- EVC/Topaz/Bennetts, Pamela
- EVG/Another time, another season/Dillon, Anna
- EWC/Watership Down/Adams, Richard

7.2 Appendix 2: Susanne tagset

Susanne tagset for English (taken from [Sampson, 1995])

- APPGf: her as possessive
- APPGh1: its
- APPGh2: their
- APPGi1: my as possessive
- APPGi2: our
- APPGm: his except as pronoun
- APPGy: your
- AT: the
- AT1: indefinite article a, an
- AT1e: every
- ATn: no as determiner or qualifier
- BTO: in_order introducing infinitive
- BTO21: in[_order] introducing infinitive
- BTO22: [in_]order introducing infinitive
- CC: co-ordinating conjunction: and, and/or, plus, as_well_as etc.
- CC31: co-ordinating conjunction: as[_well_as]
- CC32: co-ordinating conjunction: [as_]well[_as].
- CC33: co-ordinating conjunction: [as_well_]as
- CCB: but as co-ordinating conjunction
- CCn: nor
- CCr: or
- CS: subordinating conjunction
- CS21: subordinating conjunction
- CS22: subordinating conjunction
- CS31: subordinating conjunction
- CS32: subordinating conjunction

- CS33: subordinating conjunction
- CSA: as as subordinating conjunction or as preposition in comparative sense
- CSN: than in all uses
- CST: that as subordinating conjunction including relative clauses; non-standard as_how
- CST21: non-standard as[_how]
- CST22: non-standard [as_]how
- CSW: whether in all uses
- CSf: for as conjunction
- CSg: though as subordinating conjunction
- CSi: if
- CSk: as_if, as_though
- CSk21: as[_if], as[_though]
- CSk22: [as_]if, as[_though]
- CSn: subordinating conjunction when (i.e. = "at the time at which")
- CSr: subordinating conjunction where (i.e. = "at the place at which")
- DA1: much, little
- DA2: many, few in all uses
- DA2R: fewer
- DA2q: several
- DA2T: fewest
- DAR: more, less in all uses except less II
- DAT: most, least in all uses
- DAg: own as part of genitive construction
- DAr: former, latter in all uses
- DAy: same, selfsame
- DAz: such in all uses
- DB2: both as determiner or pronoun
- DBa: all as determiner or pronoun
- DBh: half as determiner or pronoun

- DD: yon, yonder as determiner; somesuch, the_rest
- DD21: the[_rest]
- DD22: [the_]rest
- DD1a: determiner, demonstrative pronoun or qualifier that
- DD1b: a_bit
- DD1b21: a[_bit]
- DD1b22: [a_]bit
- DD1e: determiner or pronoun either
- DD1i: this in all uses including as qualifier
- DD1n: determiner or pronoun neither
- DD1q: another, each, one_and_the_same, as determiner or pronoun
- DD1q41: one[_and_the_same], as determiner or pronoun
- DD1q42: [one_]and[_the_same], as determiner or pronoun
- DD1q43: [one_and_]the[_same], as determiner or pronoun
- DD1q44: [one_and_the_]same, as determiner or pronoun
- DD1t: a_little
- DD1t21: a[_little]
- DD1t22: [a_]little
- DD2: a_few, a_good_few, a_good_many, a_great_many
- DD221: a[_few]
- DD222: [a_]few
- DD231: a[_good_few], a[_good_many], a[_great_many]
- DD232: [a_]good[_few], [a_]good[_many], [a_]great[_many]
- DD233: [a_good]_many, [a_great]_many, [a_good]_few
- DD2a: those
- DD2i: these
- DDQ: what
- DDQGq: whose in interrogative uses
- DDQGr: whose in relative uses

- DDQV: whichever, whatever, whichsoever, whatsoever, no_matter_which, no_matter_what
- DDQV31: no[_matter_which], no[_matter_what]
- DDQV32: [no_]matter[_which], [no]_matter[_what]
- DDQV33: [no_matter_]which, [no_matter_]what
- DDQq: which in interrogative uses
- DDQr: which in relative uses
- DDf: enough as pronoun or pre- or postmodifying a noun
- DDi: some as determiner or pronoun
- DDo: a_lot
- DDo21: a[_lot]
- DDo22: [a_]lot
- DDy: determiner or pronoun any
- EX: existential there
- FA: suffix (separately wordtagged)
- FB: prefix (separately wordtagged)
- FD: distorted word- used in analysing speech
- FO: indeterminate formula
- FOc: formula or acronym for chemical substance, molecule or subatomic particle
- FOp: London postal district, British post code, American "Zip code": W $\gg$ C $\gg$ 2, LA6 7AN
- FOqc: chemical equation, when analysed as a single word
- FOqx: algebraic equation, when analysed as a single word
- FOr: road name (M6 etc.)
- FOs: registration/reference/serial/model number
- FOt: telephone number ...
- FOx: algebraic expression with nominal as opposed to equative function (x, dx/dy etc.)
- FW: foreign word ...
- FWg: biological Latin name of genus or other rank higher than species

- FWs: biological Latin species (or lower-rank) name
- GG: Germanic genitive inflexion ...
- ICS: considering, notwithstanding
- ICSk: preposition or subordinating conjunction like ...
- ICSt: preposition (with complement) or subordinating conjunction after, before, ere, since, until, till
- ICSx: preposition or subordinating conjunction but, except, save
- IF: preposition for
- II: preposition including prepositional use of word that can function as adverb or preposition
- II21: preposition including prepositional use of word that can function as adverb or preposition
- II22: preposition including prepositional use of word that can function as adverb or preposition
- II31: preposition including prepositional use of word that can function as adverb or preposition
- II32: preposition including prepositional use of word that can function as adverb or preposition
- II33: preposition including prepositional use of word that can function as adverb or preposition
- II41: preposition including prepositional use of word that can function as adverb or preposition
- II42: preposition including prepositional use of word that can function as adverb or preposition
- II43: preposition including prepositional use of word that can function as adverb or preposition
- II44: preposition including prepositional use of word that can function as adverb or preposition
- IIa: preposition as used non-comparatively
- IIb: preposition by
- IIg: aged as pseudo-preposition
- IIp: per, solidus character representing “per” (as in 3 counts/minute)
- IIt: preposition to

- IIx: mathematical infix operator
- IO: of
- IW: with in all uses, preposition without; what_with
- IW21: what[_with]
- IW22: [what_]with
- JA: adjective used predicatively
- JA21: adjective used predicatively
- JA22: adjective used predicatively
- JB: adjective used attributively
- JBR: inner, lesser, nether, outer, upper
- JBT: utmost, uttermost
- JBo: other in all uses
- JBy: only as adjective
- JJ: general adjective
- JJ21: general adjective
- JJ22: general adjective
- JJR: comparative adjective including elder, further
- JJT: superlative adjective including eldest, furthest
- JJh: pseudo-adjective formed by suffixing -ed to the last word of a nominal compound ...
- JJj: abbreviated adjective appended to organization name to identify its legal status ...
- JJs: adjective in -most other than utmost, uttermost
- LE: pre-co-ordinator both, not_only
- LE21: pre-co-ordinator not[_only]
- LE22: pre-co-ordinator [not_]only
- LEe: pre-co-ordinator either
- LEn: pre-co-ordinator neither
- MC: cardinal numeral from zero upwards, spelled out; including umpteen
- MCb: “labels” used for cross-reference within and between texts ...

- MCd: numeral including decimal point
- MCe: number containing a non-decimal separator, other than a time-of-day expression
- MCn: cardinal numeral written in digits ...
- MCo: 0 written as digit
- MCr: roman numeral
- MCs: integer written digitally, with leading zero
- MCy: year name written digitally, in full or with apostrophe
- MC1: one as numeral, spelled out ...
- MC1n: 1 written as digit, including use as ordinal
- MC2: plural of cardinal, spelled out
- MC2n: plural of cardinal written in digits (1s, 3's etc.)
- MC2r: roman numeral pluralized
- MC2y: plural of year name
- MD: ordinal form used as ordinal adjective or adverb or as fraction
- MDn: ordinal numeral written digitally
- MDo: first second
- MDt: next last
- MFn: fraction written digitally
- ND1: direction: north, N etc.
- NN1c: C noun
- NN1c21: C noun
- NN1c22: C noun
- NN1m: noun beginning mid- other than time noun
- NN1n: M+C noun
- NN1u: M noun
- NN1u21: M noun
- NN1u22: M noun
- NN1ux: M noun in -ics
- NN2: P noun ...

- NN221: P noun
- NN222: P noun
- NNJ: J-only items (singular or plural)
- NNJ1c: J+C noun
- NNJ1n: J+M+C noun
- NNJ2: J+P noun
- NNL: L-only item
- NNLb: L-only noun with tendency to precede the specific name
- NNLc: L+C+P noun
- NNL1c: L+C noun
- NNL1cb: L+C noun with tendency to precede the specific name
- NNL1n: L+M+C noun
- NNL2: L+P noun
- NNOc: dozen, score, gross, hundred, thousand, million, billion, trillion (etc.), zillion
- NNOn: abbreviated NNOc word, ambiguous between singular and plural
- NNS: S-only item, singular
- NNS21: S-only item, singular
- NNS22: S-only item, singular
- NNS1c: S+C noun
- NNS1c21 S+C noun
- NNS1c22 S+C noun
- NNS1n: S+M+C noun
- NNS2: S+P noun
- NNSA: S item following name, with or without C use
- NNSS: S-only item, plural
- NNSj: S-only adjectival item
- NNT1c: singular time noun that can head a noun phrase functioning adverbially
- NNT1h: name of holiday or season

- NNT1h21 name of holiday or season
- NNT1h22 name of holiday or season
- NNT1m: point-of-time noun, which cannot head a noun phrase functioning adverbially
- NNT2: plural of any NNT1 ... noun
- NNU: U-only item (singular, plural or neutral)
- NNU21: U-only item (singular, plural or neutral)
- NNU22: U-only item (singular, plural or neutral)
- NNU1c: U+C noun
- NNU1n: U+M+C noun
- NNU2: U+P noun
- NNUb: unit symbol which precedes numeral
- NNUc: U+C+P noun
- NNUn: U+M+C+P noun
- NNUp: %, percent, per_cent
- NNUp21: %, percent, per_cent
- NNUp22: %, percent, per_cent
- NNa: time of day written digitally
- NNb: attributive common noun
- NNc: C+P noun
- NNm: singular abbreviation normally followed by a numeral functioning as a proper name
- NNmm: plural of NNm abbreviation
- NNn: M+C+P noun
- NNp: time of day written digitally in 24-hour notation from 13.00 on
- NNu: M+P noun
- NNux: M+P noun in -ics
- NP1c: country name
- NP1c21: country name
- NP1c22: country name

- NP1f: feminine forename
- NP1g: miscellaneous singular geographical proper name
- NP1i: organization name: initial of personal name
- NP1j: organization name
- NP1j31: organization name
- NP1j32: organization name
- NP1j33: organization name
- NP1m: masculine forename
- NP1p: “province” name: name of US state, English county etc.
- NP1p2: “province” name: name of US state, English county etc.
- NP1p22: “province” name: name of US state, English county etc.
- NP1s: surname
- NP1t: town name
- NP1t21: town name
- NP1t22: town name
- NP1x: miscellaneous singular proper name
- NP1z: “code name” functioning as countable noun ...
- NP2c: plural country name
- NP2f: plural feminine forename
- NP2g: miscellaneous plural geographical proper name
- NP2j: plural organization name
- NP2m: plural masculine forename
- NP2p: plural “province” name
- NP2s: plural surname, including company names by adding -s without apostrophe to a surname
- NP2t: plural town names
- NP2x: miscellaneous plural proper names
- NP2z: pluralized form of “code name” functioning as countable noun
- NPD: abbreviated day of the week
- NPD1: day of the week

- NPD2: plural of day of the week
- NPM: abbreviated month name
- NPM1: month name
- NPM2: plural of month name
- PN: none in all uses
- PN1: anybody, anyone, anything, everybody, everyone, everything, naught, nobody, no_one, nothing, somebody, someone, something
- PN121: no_[one]
- PN122: [no]_one
- PN1o: one as impersonal pronoun
- PN1z: so as a pro-form
- PNQOq: interrogative whom
- PNQOr: relative whom
- PNQSq: interrogative who
- PNQSr: relative who
- PNQVG: no_matter_whose, whosever
- PNQVO: no_matter_whom, whomever
- PNQVS: no_matter_who, whoever, whosoever
- PNX1: oneself
- PPGf: hers
- PPGh2: theirs
- PPGi1: mine as pronoun
- PPGi2: ours
- PPGm: his as pronoun
- PPGy: yours
- PPH1: it
- PPHO1f: her as pronoun
- PPHO1m: him
- PPHO2: them
- PPHS1f: she

- PPHS1m: he
- PPHS2: they
- PPIO1: me
- PPIO2: us
- PPIS1: I as personal pronoun
- PPIS2: we
- PPX1f: herself
- PPX1h: itself
- PPX1i: myself
- PPX1m: himself
- PPX1y: yourself
- PPX2h: themselves, each_other, one_another
- PPX2h21 [each]_other, [one]_another
- PPX2h22 [each]_other, [one]_another
- PPX2i: ourselves, royal ourself
- PPX2y: yourselves
- PPY: you
- RAa: ago, since as synonym for ago
- RAb: A.D., Anno_Domini and counterparts used in pre-date position in other calendars
- RAc: co-ordination-closing element
- RAc21: co-ordination-closing element
- RAc22: co-ordination-closing element
- RAc31: co-ordination-closing element
- RAc32: co-ordination-closing element
- RAc33: co-ordination-closing element
- RAe: else in all uses
- RAh: a.m., p.m., o'clock
- RAi: inst, ult etc.
- RAj: postnominal adjective

- RAn: whatever, whatsover, after negative or non-assertive (any ...) nominal head
- RAp: per_annum, pa, per_diem etc.
- RAp21: per[_annum], per[_diem]
- RAp22: [per_]annum, [per_]diem
- RAq: apiece, distributive uses of each
- RAx: mathematical postfix operator
- RAy: B.C. and counterparts following dates in other calendars
- RAz: or_so as in fifty or so
- REX: apposition-introducing element: for_example, for_instance, namely, eg, ie, that_is, viz
- REX21: apposition-introducing element: for[_example],for[_instance], e.[g.], that[_is]
- REX22: apposition-introducing element: [for_]example, [for_]instance, [e.]g., [that_]is,
- RG: qualifier having no other adverbial use
- RG21: qualifier having no other adverbial use
- RG22: qualifier having no other adverbial use
- RGa: as as qualifier
- RGb: quite as qualifier or before article
- RGf: too as qualifier
- RGi: about, around, circa, getting_on_for, over, some, under, up_to used with quantity or number
- RGi21: up[_to] used with quantity or number
- RGi22: [up_]to used with quantity or number
- RGr: rather as qualifier or before article, more_of before article
- RGr21r: more[_of] before article
- RGr22r: [more_of] before article
- RGz: so as qualifier
- RGA: indeed as qualifier (commonly following head)
- RGAf: enough as qualifier

- RGQq: how as qualifier
- RGQV: however, no_matter_how as qualifiers
- RGQV31: no[_matter_how] as qualifiers
- RGQV32: [no_]matter[_how] as qualifiers
- RGQV33: [no_matter_]how as qualifiers
- RL: adverb of place or direction
- RL21: adverb of place or direction
- RL22: adverb of place or direction
- RLe: elsewhere
- RLh: here; there as adverb of place
- RLn: downstairs, upstairs
- RLw: somewhere, someplace, anywhere, anyplace, everywhere, nowhere
- RP: adverbial uses of across, down, in, off, on, out, over, through, up
- RPK: about in adverbial and catenative (about to ...) uses
- RR: general adverb
- RR21: general adverb
- RR22: general adverb
- RR31: general adverb
- RR32: general adverb
- RR33: general adverb
- RR41: general adverb
- RR42: general adverb
- RR43: general adverb
- RR44: general adverb
- RRQV: wh-...-ever adverb: however, no_matter_how etc.
- RRQV31: adverb no[_matter_how]
- RRQV32: adverb [no_]matter[_how]
- RRQV33: adverb [no_matter_]how
- RRQq: interrogative wh- adverb

- RRQr: relative wh- adverb
- RRR: single-word comparative adverb other than more, less
- RRT: single-word superlative adverb other than most, least
- RRe: enough as clause verb
- RRf: far as adverb
- RRg: long as adverb
- RRs: otherwise, yet
- RRx: only as adverb
- RRy: any as qualifier with comparative
- RRz: so introducing main clause or clause of purpose or result, or as adverb of manner or degree
- RT: again, hereafter, overnight
- RTn: then
- RTo: now
- RTt: today, tomorrow, tonight, yesterday
- TO: infinitival tl
- UH: interjection
- UH21: interjection
- UH22: interjection
- VB0: be
- VBDR: were
- VBDZ: was
- VBG: being
- VBM: am
- VBN: been
- VBR: are
- VBZ: is
- VD0: do
- VDD: did
- VDG: doing

- VDN: done
- VDZ: does
- VH0: have
- VHD: had (past tense finite)
- VHG: having
- VHN: had (past participle)
- VHZ: has
- VMK: ought, used as modal catenative
- VMd: modal (past)
- VMo: modal (present)
- VV0i: intransitive verb base form
- VV0t: transitive verb base form
- VV0v: base form with transitive verb or intransitive verb
- VVDi: intransitive verb, past tense
- VVDt: transitive verb, past tense
- VVDv: transitive or intransitive verb, past tense
- VVGi: present participle of intransitive verb
- VVGt: present participle of transitive verb
- VVGv: present participle of verb having (in)transitive uses
- VVGK: going as catenative
- VVNi: past participle of intransitive verb
- VVNt: past participle of transitive verb
- VVNv: past participle of verb having (in)transitive uses
- VVNK: bound as catenative
- VVZi: third person singular of intransitive verb
- VVZt: third person singular: of transitive verb
- VVZv: third person of transitive or intransitive verb
- XX: not
- YB: text division of paragraph or higher rank, not having heading: ¡majbrk¿, ¡minbrk¿

- YBL: beginning of heading: ¡bmajhd¿, ¡bminhd¿
- YBR: end of heading: ¡emajhd¿, ¡eminhd¿
- YC: comma
- YD: dash (em-dash or larger), solidus character
- YE: ellipsis mark (three dots)
- YF: full stop (American "period")
- YG: logical position of transformationally moved/deleted item
- YH: hyphen or en-dash
- YIL: opening (single or double) inverted comma(s)
- YIR: closing (single or double) inverted comma(s)
- YM: filled pause- used in analysing speech
- YN: colon
- YND: colon-dash
- YO: bullet, pilcrow, or other special symbol visually marking a text division
- YP: silent pause- used in analysing speech
- YPL: opening bracket (round, square etc.)
- YPR: closing bracket (round, square etc.)
- YQ: question mark
- YR: interruption point - used in analysing speech
- YS: semicolon
- YTL: begin italics/bold-face
- YTR: end italics/bold-face
- YX: exclamation mark
- ZZ1: singular letter of the (roman or other) alphabet (including names or letters spelled out)
- YY2: letter of the alphabet with plural inflexion (including spelled-out cases)

7.3 Appendix 3: Queries for fiction features

Queries and sample concordances (taken from TEC) for typical fiction features:

past tense verbs:

[pos="VVD. * "];
minus [pos="VM. * "] [pos="VVD. * "];
minus [pos="VH. * "] [pos="VVD. * "];
minus [pos="VB. * "] [pos="VVD. * "];
* *
So many cobwebs ! I ≪ saw ≫ a millipede on the wall
protect you . " He ≪ began ≫ to speak in glowing terms
to her bosom and ≪ hugged ≫ us , and we buried our
nt a long way and ≪ emerged ≫ outside the city , where
oming of pumps . I ≪ heard ≫ them guffawing with
* *

third person pronoun:

[pos="PPX1 | PPX2 | PPHS1 | PPHO1 | PPG | PPH1 | PPHO2 | PPHS2"];
* *
language , who said ≪ they ≫ were my subjects . I
sed having seven of ≪ them ≫ was an utter extravagance
d not yet returned . ≪ He ≫ arrived on the tenth night
t proceeding alone , ≪ he ≫ could cover a distance
y quickly noted that ≪ it ≫ was sufficient to multiply
* *

perfect aspect:

[pos="VH. * "] [ ] {0,2} [pos="VVN. * "];
* *
ome haunted . It ≪ had been taken ≫ over by those
And then , when he ≪ has pulled ≫ the knife out of the
- who likewise ≪ had long ago given ≫ up plans of ever
n of day and night ≪ has begun ≫ again . But no , this
I realise that I ' ≪ ve used ≫ a rather trivial
* *

public verbs:

[word=”[Aa]ssert. ∗ | ASSERT. ∗ | [Dd]eclar. ∗ | DECLAR. ∗ | [Mm]ention. ∗ | MENTION. ∗ | [Pp]roclaim. ∗ | PROCLAIM. ∗ | [Rr]eport. ∗ | REPORT. ∗ | [Ss]ay. ∗ | SAY. ∗ | [Tt]ell. ∗ | TELL. ∗ | [Ss]aid | SAID | [Tt]old | TOLD” & pos=”V. ∗ ”];

⋆ ⋆
Judas the Galilean ≪ reported ≫ that one of the
, I ’ m here , ’ he ≪ said ≫ . At a leisurely pace
s spots . ’ ’ I ’ ll ≪ tell ≫ you one thing . I can
himm be . Should he ≪ mention ≫ all this to Danny ?
t in simple terms and ≪ say ≫ that this was something
⋆ ⋆

synthetic negation:

[word=”[Nn]o | NO | [Nn]or | NOR | [Nn]either | NEITHER”];

⋆ ⋆
he map , but there was ≪ no ≫ city of Anagoor . It
Therefore , there was ≪ no ≫ sensible reason for
the price . There was ≪ no ≫ hurry , he answered ,
now neither the cause ≪ nor ≫ the mechanics of the
the evening papers . ≪ Neither ≫ radio nor tele
⋆ ⋆

present participle clauses:

[pos=”N. ∗ | Y. ∗ ”] [pos=”VVG. ∗ ”];

⋆ ⋆
in this , ” he said ≪ , pointing ≫ to a dignified
n suddenly stop . A ≪ lad holding ≫ a cleaning rag
’ t take so many ≪ people flocking ≫ there from near
y as he did each day ≪ , knowing ≫ that he will no
hen my eyes met his ≪ . Standing ≫ alone by his desk
⋆ ⋆

time adverbials:

[pos=”RT. ∗ ”];

⋆ ⋆
d up leaving me . I had ≪ now ≫ grown up , and such
ir . I ’ m sorry ... ” ≪ Then ≫ I think : how stupid
a turned the sound off ≪ again ≫ and then switched
ty , I would see Clara ≪ again ≫ , I would spend for
of Ricardo that exist ≪ today ≫ he looks like an
⋆ ⋆

place adverbials:

[pos="RL. ∗ "];
⋆ ⋆
hich the firemen washed ≪ away ≫ before dawn . It is
he house when he turns ≪ around ≫ : " You must leave
breathe . After he went ≪ back ≫ outside , on the
r will return . It was ≪ there ≫ , near the magazine
narrative abandoned ≪ somewhere ≫ without name or
⋆ ⋆

adverbs:

[pos="RL. ∗ | RP. ∗ | RR. ∗ | RT. ∗ "];
⋆ ⋆
can ' t let it be , we ≪ always ≫ have to be setting
light fiction , was ≪ obsessively ≫ eager that he
st . He asked around ≪ discreetly ≫ and learned that
an . ' " The father was ≪ deeply ≫ disturbed . He did
s . He flew off and came ≪ back ≫ in a single day .
⋆ ⋆

Queries and sample concordances (taken from TEC) for untypical fiction features:

present tense verbs:

[pos="VVZ. ∗ "];
⋆ ⋆
is is not the time . He ≪ shakes ≫ Berton ' s hand
; his hands relax . He ≪ feels ≫ happy , amid the
lope by a young lad who ≪ says ≫ ' A tip from the
een him and someone who ≪ looks ≫ like his twin
ter looking at him . He ≪ begins ≫ with the more
⋆ ⋆

attribute adjectives:

[pos="J. ∗ "] [pos="N. ∗ "];
⋆ ⋆
spapers for whose ≪ illegal survival ≫ she had labour
lence . It was an ≪ endless struggle ≫ to make enough
dawn , in that ≪ artificial brightness ≫ there were
in winter , to the ≪ little room ≫ in which her
I wanted to know his ≪ true name ≫ , and what his
⋆ ⋆

wh relative clauses:

[pos="N. ∗ | DD2. ∗ "] [pos!="YX | YS | YQ | YN | YF"] {0,1} [word!="W. ∗ | [Ww]hat" & pos="DDQG | DDQ | PNQO | PNQS"];
⋆ ⋆
There were also some ≪ people who ≫ were simply
ant ' s theory of ≪ cosmology , which ≫ would oblige
r there might be a ≪ bottle which ≫ he had forgotten
ut all those other ≪ men in whose ≫ name no book has
rget the dark cup of ≪ tea which ≫ is beginning to
⋆ ⋆

pied piping:

[pos="N. ∗ | DD1. ∗ "] [pos!="YX | YS | YQ | YN | YF"] {0,1} [pos="II. ∗ | IF. ∗ | IW. ∗ "] [pos="DDQ | DDQG | PNQO | PNQS" & word!="[Ww]hat"];
⋆ ⋆
efore there were ≪ oaks under which ≫ I went alone to
' s trade . The ≪ woman , to whom ≫ Stefano had
iver a serious ≪ judgement on which ≫ a man ' s life
cal to us . The ≪ planet on which ≫ Professor Splitt
customs of the ≪ country in which ≫ he appeared .
⋆ ⋆

phrasal coordination:

[pos=”N. ∗ ”] [pos=”CC”] [pos=”N. ∗ ”];
plus [pos=”V. ∗ ”] [pos=”CC”] [pos=”V. ∗ ”];
plus [pos=”J. ∗ ”] [pos=”CC”] [pos=”J. ∗ ”];
plus [pos=”RL. ∗ | RP. ∗ | RR. ∗ | RT. ∗ ”]
pos=”CC”
[pos=”RL. ∗ | RP. ∗ | RR. ∗ | RT. ∗ ”];
⋆ ⋆
rpose . Dozens of ≪ men and women ≫ immediately
; once again I ≪ stopped and turned ≫ around . ” Oh
ecause writing ≪ insignificant or ugly ≫ things usual
ted . We talked ≪ quietly and happily ≫ throughout
⋆ ⋆

nominalizations:

[pos=”N. ∗ ” & word=”. ∗ ment”];
plus [pos=”N. ∗ ” & word=”. ∗ ness”];
plus [pos=”N. ∗ ” & word=”. ∗ ity”];
plus [pos=”N. ∗ ” & word=”. ∗ tion”];
⋆ ⋆
n . She saw this new ≪ development ≫ as a happy
ed the onlookers that ≪ happiness ≫ was close at hand
s into a new objective ≪ reality ≫ for himself , for
959 , when his first ≪ collection ≫ of short stories
⋆ ⋆

conjuncts:

[pos=”YX | YS | YQ | YN | YF”] [ ] {0,1} [pos=”R. ∗ | CC. ∗ ”];
⋆ ⋆
more lunar thann solar ≪ . And ≫ I am in danger of
rican as possible . Yes ≪ ? So ≫ you can speak
bacon in the Kachelofen ≪ . But ≫ it ’ s such a
d come to a quiet end ≪ . However ≫ , one early
preferred it that way ≪ . Then ≫ I could imagine her
⋆ ⋆

passives:

[pos=”VB. ∗ ”] [ ] {0,1} [pos=”VVN. ∗ ”];
⋆ ⋆
the feast ’ s flame ≪ was extinguished ≫ there was a
n , in which Kuhlmann ≪ is described ≫ as a quasi-
. Telephone numbers ≪ are exchanged ≫ . Intimate
the case which would ≪ be distributed ≫ throughout
aberration . ” Guyotat ≪ was born ≫ in 1940 in a
⋆ ⋆

by passives:

[pos=”VB. ∗ ”] [ ] {0,1} [pos=”VVN. ∗ ”] [ ] {0,5} [word=”by”];
⋆ ⋆
ere , it cannot have ≪ been caused by ≫ tobacco smoke
going to start , he ≪ was pacified by ≫ the chairman
s it is to see a man ≪ being killed by ≫ a mental
n . Rebecca ≪ was told all this and more by ≫ Nell
the corridor and I ≪ was distracted by ≫ the noise of
⋆ ⋆

past participle clauses:

[pos=”N. ∗ ”] [pos=”YC”] [pos=”VVN. ∗ ”] [pos!=”N. ∗ | Y. ∗ ”];
plus [pos=”YF | YN | YQ | YS | YX”] [pos=”VVN. ∗ ”] [pos!=”N. ∗ | Y. ∗ ”];
⋆ ⋆
eviously an inert ≪ formula , created and ≫ developed
ld Seth ’ s Cape ≪ Matapan , published by ≫ Garzanti
tiary at the City ≪ limits , reserved for ≫ prisoners
co , and Gregorio ≪ . Unaccustomed to ≫ being away
ecording equipment ≪ . Compared to ≫ me , the Corrier
⋆ ⋆

subordinators:

[pos=”CS”];
⋆ ⋆
ne , and he was amazed ≪ when ≫ he learned that the
aiting on the flat , ≪ because ≫ there are two of
luctant to walk on . ≪ Although ≫ they can ’ t see me
envy , or resentment , ≪ if ≫ you prefer a more
swer , I ’ m put out . ≪ When ≫ he arrived home the
⋆ ⋆

that clauses:

[pos=”CST”];
⋆ ⋆
Terrified , I realised ≪ that ≫ it was a preordained
the dishonour and shame ≪ that ≫ would cloud over her
use . But the truth was ≪ that ≫ she had become
ht I was only imagining ≪ that ≫ it was getting better
Then it dawned upon him ≪ that ≫ the money mentioned
⋆ ⋆

demonstratives:

[pos=”DD1 | DD2”];
⋆ ⋆
ss . They never abandon ≪ that ≫ night , that black
truck him . He had seen ≪ this ≫ officer before . He
nd times better . Kids ≪ these ≫ days are brighter
’ Did his mother see it ¡≪ that ≫ way ? ’ ’ Huh ! ’
oney , the money which ¡≪ those ≫ poor women had to
⋆ ⋆

final prepositions:

[pos=”N. ∗ | DD2. ∗ ”] [pos!=”YX | YS | YQ | YN | YF”] {0,1} [word!=”W. ∗ | [Ww]hat” & pos=”CST | DDQG | DDQ | PNQO | PNQS”] [pos!=”YX | YS | YQ | YN | YF | YC”] {2,8} [pos=”IF | II. ∗ | IO | IW. ∗ ”] [pos=”YX | YS | YQ | YN | YF | YC”];
⋆ ⋆
itary ≪ service which she had so longed for . ≫ And
a ≪ gadget that I ordinarily had no dealings with . ≫
≪ wound , which I ’ d temporarily forgotten about . ≫
ts of ≪ things that no one else knows about : ≫ They
the ≪ sergeant who he ’ d been speaking with . ≫ ’
⋆ ⋆

existential *there*:

[pos=”EX”];
⋆ ⋆
dness . 19 May SECRETS ≪ There ≫ are times when my
kept on its way . But ≪ there ≫ were only chicken
, we say . And isn ’ t ≪ there ≫ anybody else here ?
s of the day to see if ≪ there ≫ ’ s anything left to
ting to be saved . But ≪ there ≫ was no one there .
⋆ ⋆

7.4 Appendix 4: Experimental instructions

The experiment consists of two different parts. Both parts include a German-English translation, on the one hand, and a French-English translation, on the other. Altogether four different translation tasks have to be carried out (4 x 40 sentences).

Before the test starts, please say out loud your name, your home country, the kind of your academic education and your professional experience as an translator (in terms of years and languages).

Part 1 – verbal translation task:
Read out loud every sentence and verbalize its translation immediately (without thinking about it or writing it down). Translate the SL sentences in the given order (first the German sentences, then the French ones). Do not worry about errors or lacking words. The only important thing is that your thoughts are immediately verbalized.

Part 2 – written translation task:
Translate the German and French sentences in the given order and write them down – you are under no time pressure and you may use the provided dictionaries (a German-English, a French-English, a monolingual German and a monolingual French dictionary).

It is important that all translated sentences (in part 1 and 2) are as idiomatical as possible. The English audience should not recognize the translated text as a translation, but as a text originally produced in English.

Please speak as loud and clear as possible during the verbal translation task!
Please write as legibly as possible in the written translation task!

Thank you very much for taking part in the experiment!

7.5 Appendix 5: Experimental design

German-English verbal translation task:

A1: In der Dealer-Szene macht sich Unruhe breit, weil die Polizei mit Hilfe von Rauschgiftfahndern Spuren finden konnte.
A2: Der Personalrat feierte seinen Erfolg, nachdem die Verhandlungen mit der neuen Geschäftsführung beendet waren.
A3: Es wurden kurzerhand alle Beteiligten bestraft, da keiner der Anwesenden die Schuld auf sich nehmen wollte.
A4: Der Lehrer versucht seinem Schüler zu helfen, indem er ihm kostenlose Nachhilfe anbot.
A5: Die Mädchen nahmen das Mitfahr-Angebot an, damit sie früher zu Hause waren.
A6: Dem Ausflug steht nichts mehr im Wege, wenn das Wetter sich nicht verschlechtert.
A7: Ira sollte eine Wohnung finden, bevor sie sich an der Universität anmeldet.
A8: Hans sollte vorher seine Kollegen benachrichtigen, falls das wirklich seine Absicht ist.

B9: Nach der Trauung fährt die Hochzeitsgesellschaft ins Restaurant.
B10: Vor der Vorstellung sollten die Schauspieler noch etwas essen.
B11: Wegen Christas Unfähigkeit ist die Theatergruppe nun in dieser Lage.
B12: Seit dem Unfall ist Kurt ein anderer Mensch.
B13: Durch die Hilfe des Kindergartens konnte die Gemeinde das Fest erst organisieren.
B14: Mit Dieters Mutter gehe ich nicht zu dem Theaterstück.
B15: Anstatt der Vorwürfe sollte Michael ihr lieber Hilfe anbieten.
B16: Anläßlich des Jubiläums fährt der Sohn doch sehr gerne nach Hause.

C17: Dem Abriß der alten Schule folgte der Aufbau eines neuen Bürogebäudes.
C18: Inas mangelnde Sehstärke verursachte schon einige Autounfälle.
C19: Daß Peters Mutter ihrem Sohn nie verzeihen konnte, bedingt leider seine Abneigung.
C20: Petras Geschrei ergänzt genau das Verhalten, das sie bisher an den Tag legte.
C21: Die Aussage der Nachbarin bekräftigt, daß der Junge die Blumen gestohlen hat.
C22: Die Befragung der Polizei klärte den Befund am Tatort.
C23: Daß die Nachbarin den verheirateten Ehemann küßte, sorgte für allerhand Dorfgespräche.
C24: Ständiges Betteln verschaffte ihm den Job.

D25: Der Besuch beim Anwalt steht in kausalem Zusammenhang zu dem Schreiben des Anklägers.
D26: Marys Haarfrisur ist verantwortlich für den tollen Erfolg bei Männern.
D27: Die Erläuterungen des Vaters sind hilfreich für das Auffinden des Kindes.

D28: Das Schuldeingeständnis wirkt auf jeden Fall beschwichtigend für den Prozeß.
D29: Dieser Hinweis wird sich klärend auf den Diebstahl auswirken.
D30: Das leise Singen der Mutter wirkt beruhigend auf den schlechten Schlaf ihrer Tochter.
D31: Das Auftreten der Prinzen steht in temporalem Zusammenhang mit dem Schrei der Prinzessin.
D32: Das Verschwinden des Mädchens steht vermutlich in direktem Zusammenhang mit dem Einbruch in das Haus.

E33: Der Besuch Georges Mutter ist wohl die Ursache für sein Verhalten.
E34: Schwere Verletzungen waren die Folge seines Unfalls.
E35: Die Ergänzung zu dem neuen Vertrag beinhaltet die Festlegung der Arbeitszeit.
E36: Der Grund für Mikes Untreue war Christines ständige Nörgelei.
E37: Die lebenslängliche Verurteilung war die gerechte Strafe für seine Verbrechen.
E38: Pünktliches Erscheinen am Arbeitsplatz ist die Bedingung für die Vertragsverlängerung.
E39: Die Jubiläumsfeier des Professors war der Zweck des Besuchs.
E40: Die Verleihung des Ordens war der Anlaß für seine Rede.

French-English verbal translation task:

A1: Frank croit Anna parce qu'il l'aime.
A2: La femme attend son fils après qu'il a raté son examen.
A3: Il y a de l'espoir tant qu'il y a de la vie.
A4: La fille travaille autant qu'elle peut.
A5: Sandra travaille dur pour qu'on veuille l'écouter.
A6: Gustav réparera sa voiture s'il y a un problème avec le moteur.
A7: Les traits d'un caractère se forment avant qu'on en ait pris conscience.
A8: Lucie ne peut pas aider son ami puisque c'est trop tard.

B9: Après le voyage en France, Anita doit rendre visite à sa grand-mère.
B10: Avant la guerre, il a travaillé dans une usine.
B11: À cause de sa victoire, il a gagné beaucoup d'argent.
B12: Dès l'assistanat, l'enseignant est considéré comme membre du corps universitaire.
B13: Sans le petit déjeuner, la chambre coûte cinq cents francs par jour.
B14: Avec tous ces touristes, le village est bien agité.
B15: Au lieu des questions, nous ferions mieux de lui présenter des solutions.
B16: Lors de leur installation à Paris, Sylvie veut organiser une grande fête.

C17: L'appel du professeur suit la mauvaise performance des élèves.
C18: La mauvaise performance de la voiture cause le terrible accident.
C19: La réponse du professeur provoque la désobéissance de la classe.
C20: Le cri de la concierge appelle a la réaction du médecin.

C21: La réponse de la voisine assure que le jeune homme a volé les fleurs.
C22: La bise de Jean-Luc explique ses sentiments pour Claire.
C23: Les activités de l'inventeur produisent des idées extraordinaires.
C24: Les détails des derniers jours empêchent l'opération du grand-père.

D25: L'initiative de la mère avait une influence négative pour le développement du fils.
D26: La lettre du professeur est nécessaire pour les études de l'étudiant.
D27: L'interprétation du roman est essentielle pour la compréhension des lecteurs.
D28: La présence de la mère est illogique considérant l'age de la fille.
D29: Des révisions régulières sont indispensables pour le fonctionnement du moteur.
D30: Le rire des gens est très comique dans ce chapitre de la pièce du théâtre.
D31: Le succès du groupe est louable considérant les difficiles conditions d'entraînement.
D32: La beauté de cette forêt est caractéristique du bon traitement des jeunes peuplements.

E33: Son mauvais comportement est certainement la raison de son renvoi.
E34: Le déclin du chômage peut être un indicateur de la meilleure situation économique du pays.
E35: L'annulation de la participation est la conséquence des événements des derniers jours.
E36: Le retard du train est la cause de l'embouteillage.
E37: La mauvaise note est le résultat de son inattention.
E38: La réponse exacte est la condition pour l'intégration dans le groupe.
E39: La croissance de la population est un développement positif de l'action de l'État.
E40: La réparation du moteur était le sujet du cours des dernières semaines.

German-English written translation task:

A1: Die Studentin muß Ihre Eltern um Unterstützung bitten, weil ihre staatliche Förderung ausläuft.
A2: Der Junge wollte eine Limonade trinken, nachdem er sein Eis verspeist hatte.
A3: Die Mutter verzieh ihrem Sohn ein letztes Mal, da sie ihn sehr gut verstehen konnte.
A4: Die Katze versuchte die Maus zu fangen, indem sie sie in die Ecke trieb.
A5: Die junge Frau heiratete den fremden Mann, damit sie endlich ihre Aufenthaltsgenehmigung bekam.
A6: Alle Kellner würden mehr Trinkgeld verdienen, wenn sie freundlicher zu den Gästen wären.
A7: Paul muß unbedingt noch einen Brief schreiben, bevor er das Büro verläßt.
A8: Sei Paula bitte nicht böse, falls sie ein wenig zu spät zu der Verabredung kommen sollte.

B9: Nach der Vorlesung wird der Professor ziemlich müde sein.

B10: Vor der Prüfung war Elke sehr nervös.
B11: Wegen der Krankheit seiner Großmutter konnte Uli leider nicht früher kommen.
B12: Seit der letzten Familienfeier haben sich die beiden Schwestern nicht mehr getroffen.
B13: Durch das Engagement ihrer Schwester konnte Heike die Kinderklinik retten.
B14: Mit ihrem Chef muß Cora dringend noch einmal reden.
B15: Anstatt des Mitleids brauchte Caroline viel eher seine Hilfe.
B16: Anläßlich des Geburtstags seiner Freundin muß Alex auf jeden Fall noch ein Geschenk kaufen.

C17: Den turbulenten Ereignissen der letzten Tage folgte zum Glück etwas Ruhe.
C18: Das überschäumende Temperament des Mädchens verursachte schon immer viel Wirbel.
C19: Die fantastische Figur der Schauspielerin zog schon immer die Blicke der Männer an.
C20: Das Gejammer des Unfallopfers verschlimmerte die ganze Lage noch.
C21: Das Aussehen des Diebes bekräftigte den Verdacht.
C22: Die Aussagen der Zeugen konnten den Sachverhalt auch nicht klären.
C23: Die Antwort seiner Zukünftigen sorgte für ein Lächeln auf seinen Lippen.
C24: Das Gutachten des Lehrers verschaffte ihr die Lehrstelle.

D25: Das Versagen der Bremsen steht natürlich in kausalem Zusammenhang zu dem Alter des Autos.
D26: Der schöne Gesang des Mädchens ist verantwortlich für die vielen Job-Angebote.
D27: Die Erklärung des Studenten ist hilfreich für die Beantragung eines Urlaubssemesters.
D28: Das Gedicht der Tochter wirkt beschwichtigend auf die schlechte Laune der Mutter.
D29: Die Aussage der Nachbarin wird sich positiv auf das Geschehen des gestrigen Tages auswirken.
D30: Das Einwirken der Salbe wirkt beruhigend auf die entzündete Haut.
D31: Die schnelle Reaktion der Feuerwehr steht sicherlich in temporalem Zusammenhang mit dem anonymen Anruf.
D32: Die Entführung der Jugendlichen steht vermutlich in direktem Zusammenhang mit dem Banküberfall.

E33: Der Verlust der Brieftasche ist die Ursache für die Wut des jungen Mannes.
E34: Eine gewaltige Geldstrafe wird die Folge seiner Dummheiten sein.
E35: Der Vertrag des Mieters beinhaltet die Einhaltung der Hausordnung.
E36: Das Versprechen des Sohnes war der Grund für sein Schweigen bei der Beerdigung.
E37: Die Verhaftung durch die Polizei war die gerechte Strafe für das Verhalten der Bande.
E38: Regelmäßige Gespräche mit dem Therapeuten sind die Bedingung für die

Besserung des Seelenzustandes des Mädchens.
E39: Der Besuch des Museums war der Zweck der Reise.
E40: Die schlechte Laune ihres Vaters war der Anlaß für das Verschwinden der Tochter.

French-English written translation task:

A1: Une pierre tombe parce qu'elle est lourde.
A2: Le garon a attendu trois semaines après que cette fille ait apparu.
A3: Il soutenait ses ministres tant qu'ils avaient la majorité.
A4: Je te hais autant que je t'aime.
A5: Julie accepte l'offre de Paul pour qu'il soit tranquille.
A6: Sarah ira à la maison si elle ne peut plus travailler.
A7: Les élèves ne parlent pas avant que le professeur ait fini son exposé.
A8: Il y a certainement des vrais miracles puisque j'ai les vus.

B9: Après la naissance de Lucas, sa mère ne travaille plus.
B10: Avant le lever du soleil, il doit arriver à la maison.
B11: À cause des hommes politiques, on ne peut plus travailler dans un pays étranger.
B12: Depuis cette époque, Louise est une poète célébrée.
B13: Sans son avocat, l'inculpé a refusé d'être interrogé.
B14: Avec l'assistance de sa mère, Eric tente l'aventure.
B15: Au lieu de l'avion, nous prendrons le train.
B16: À l'occasion de son mariage, il a acheté un nouveau chapeau.

C17: L'initiative des parents suit l'appel du docteur.
C18: Les résultats extraordinaires de Philippe causent la réaction enthousiaste de ses parents.
C19: La froideur du mari provoque la haine de la femme.
C20: La lettre de Julie appelle d'autres filles à l'aide.
C21: La question du gendarme a assuré l'incertitude de l'inspecteur.
C22: L'enquête de la police a expliqué le comportement des voleurs.
C23: L'assassinat du jeune homme provoque beaucoup de rumeurs.
C24: Les victoires de son ennemi empêchent les exercices du champion.

D25: La visite de l'avocat peut avoir une influence positive sur la situation de l'accusé.
D26: La naissance d'un droit est nécessaire pour le développement d'un État.
D27: La formation de l'oxygène est essentielle pour la vie des plantes.
D28: La confession du mari est logique dans cette situation.
D29: La consommation des vitamines est indispensable pour la santé du bébé.
D30: Cet incident est tragique dans cette phase du développement.
D31: Les idées des enfants sont difficilement louables considérant le mauvais état de la directrice.
D32: L'intelligence de cette fille est caractéristique de l'héritage de cette famille.

E33: La visite de Carol est la raison pour les sentiments du voisin.
E34: Le bilan positif est un indicateur des achats des consommateurs.
E35: Des blessures internes sont la conséquence de l'accident.
E36: La mort de sa grand-mère est la cause de la tristesse de la jeune fille.
E37: Cette peine est le résultat de tes bêtises.
E38: Une meilleure motivation est la condition pour un renouvellement du contrat de travail.
E39: Le déclin du chômage est l'effet de l'abattement de l'aide de l'état.
E40: L'assassinat du directeur est le sujet de la discussion entre les travailleurs.

7.6 Appendix 6: Experimental results

Teacher – German-English verbal translation task (transcribed):

E39: The anniversary celebration of the professor was the purpose of the visit.
C19: The fact that Peter's mother could never forgive her son unfortunately conditioned his disaffection.
B10: The actors should eat something before the performance.
A4: The teacher tried to help his pupil by offering him free tuition after hours.
A1: There is unrest throughout the drug dealer scene because the police were able to find some traces by using drug store officers.
D28: The convation of guilt will definitely have a modifying effect on the trial.
B9: After the wedding ceremony the wedding guests went to a restaurant.
D27: The father's statements was assistance in finding the child.
C17: The pulling down of the old school was followed by the construction of the new office building.
D31: The appearance of the prince stands in a temporal connection with the princess's cry.
E33: The visit of George's mother is probably the cause of his behavior.
A3: All those involved were punished summarily as none of those present was prepared to take the blame.
E40: The award of the order was the occasion for his speech.
C18: Ina's poor vision has already caused a few car accidents.
A6: There is nothing preventing the excursion as long as the weather does not turn bad.
E38: Punctual attendance at the place of work is a condition for the extension of the contract.
E36: The reason for Mike's infidelity was Christine's constant badgering.
B16: The son likes going home for the anniversary.
B14: I am not going to the play with Dieter's mother.
A8: Hans should tell his colleagues beforehand if that really is his intention.
B12: Since the accident Kurt is a changed person.
A7: Ira should find a flat before enrolling at the university.
D32: The girl's disappearance probably stands in a direct connection with the break-in.
D25: The visit to the lawyer's stands in a causal connection to the letter of the prosecutor.
C20: Petra's screaming perfectly complemented to the behavior that she had previously displayed.
E34: His accident resulted in severe injuries.
D29: This information will help to solve the theft.
B11: It's Christa's incompetence that has got the theatre group to such a situation.
C23: The neighbor's kissing the married man like that caused all sorts of gossip in the village.
E35: The appendix to the new contract also regulates the time that has to be worked.

B13: The help of the kindergarten made it possible for the community to organize the celebration properly.
D26: Mary's coiffure is responsible for the huge success she has with men.
A5: The girls accepted the offer of the lift in order to get home quicker.
C24: Constant begging got him the job.
B15: Michael should offer to help her instead of making all these accusations.
A2: The staff council celebrated its success after negotiations with the new management were concluded.
D30: The mother's quiet singing has a calming effect on her daughter's trouble to sleep.
E37: The sentence of life imprisonment was the just punishment for his crime.
C22: The police questioning solved the evidence found at the scene of crime.
C21: The neighbor's statement strengthens the case for the boy having stolen the flowers.

Teacher – French-English verbal translation task (transcribed):

A7: Character traits are formed before one is aware of them.
B14: What with all these tourists the village is really abused.
E38: A correct answer is the condition for membership in the group.
B9: After her trip to France Anita has to visit her grandmother.
E40: Motor repairs were the subject of the course of the last few weeks.
A1: Frank believes Anna because he loves her.
B10: Before the war he worked in a factory.
A5: Sandra works hard so the people want to listen to her.
D27: The interpretation of the novel is essential for the reader's understanding.
A6: Gustav will repair his car if there is a problem with the motor.
C22: Jean-Luc's kiss explains his feelings for Claire.
D30: The people's laughter is very funny in this chapter of the play.
E35: The cancellation of participation is the outcome of the events of the last few days.
C20: The genitor's shout appealed the reaction of the doctor.
C24: The details of the last days are preventing the grandfather's operation.
E33: His bad behavior is certainly the reason for his being sent away.
D32: The beauty of this forest is characteristic of the good treatment of the young populations.
C21: The neighbor's answer makes it certain that the young man stole the flowers.
B12: Since his time as an assistant professor the teacher is considered as a member of the university body.
C17: The teacher's appeal followed the bad performance of the pupils.
A2: The woman is waiting for her son after he failed the exam.
B15: Instead of questions we had been better off offering him solutions.
A3: Where there is life there is hope.
B11: His victory has made it possible for him to earn lots of money.
D29: Regular checks are indispensable if the motor is to function properly.
A4: The girl works as much as she can.

D31: The group's success is worthy of praise considering the difficult training conditions.
C19: The teacher's answer provokes the class's disobedience.
D25: The mother's initiative had a negative influence on the son's development.
C18: The car's bad performance causes the terrible accident.
D26: The teacher's letter is necessary for the student's studies.
B16: When she sets up house in Paris Silvia wants to have a big party.
E37: The bad grade is the result of his lack of attention.
B13: The room costs 500 francs a day without breakfast.
E36: The traffic jam is the result of the delayed train.
E34: The decline in unemployment can be an indicator of a better economic situation in the country.
D28: The mother's presence is illogical considering the daughter's age.
E39: The growth of the population is a positive development of the action of the state.
A8: Lucie can't help her friend because it is too late.
C23: The activities of the inventor produce extraordinary ideas.

Teacher – German-English written translation task:

D29: The neighbor's statement will impact positively on yesterday's events.
A6: All the waiters would earn more tips if they were nicer to the guests.
B14: Cora urgently needs to talk to her boss again.
C21: The thief's appearance strengthened our suspicion.
A5: The young woman married the foreigner in order to get her residency permit at last.
A4: The cat tried to catch the mouse by chasing him into the corner.
C17: Thankfully the turbulent events of the last few days were followed by some peace and quiet.
C18: The girl's effervescent temperament had always caused a stir.
E37: The police's putting them in prison was the just punishment of the band's behavior.
D30: Application of the ointment has a soothing effect on inflamed skin.
B11: Unfortunately Uli couldn't come any sooner due to his grandmother's illness.
E40: Her father's bad mood occasioned the daughter's disappearance.
D28: The daughter's poem has a calming effect on the mother's bad mood.
A8: Please don't be angry with Paula if she's a bit late for her appointment.
D32: The kidnapping of the teenagers is probably directly related to the bank robbery.
E33: The loss of the wallet is the cause of the young man's rage.
E34: The silly things he did will result in a hefty fine.
A1: The student has to ask her parents for help as her stipend is about to run out.
D27: The student's declaration assists in applying for a vacation semester.
B10: Elke was very nervous before the test.
B9: The professor will be quite tired after the lecture.

A2: The boy wanted to have a soft drink after eating his ice-cream.
D26: It's the girl's beautiful singing that is responsible for all the job offers.
B15: Caroline needed his help far more than his sympathy.
B13: Her sister's commitment enabled Heike to save the clinic.
E35: The tenant's lease obliges him to respect the regulations regarding the use of the building.
E39: The visit to the museum was the purpose of the trip.
D25: The failure of the brakes is of course causally related to the vehicle's age.
A7: Paul before leaving the office just has to write one more letter.
C23: His bride-to-be's answer brought a smile to his lips.
D31: The swift reaction of the fire brigade is definitely temporally related to the anonymous phone call.
B12: The two sisters haven't seen each other since the last family get-together.
C24: The teacher's report got her the position.
E36: The son's promise was the reason for his silence at the funeral.
E38: Regular consultations with the therapist are essential for an improvement in the girl's mental condition.
B16: Alex definitely has to buy a present for his girl-friend's birthday.
C22: The witnesses' statements didn't help solve the case either.
C19: The actress's fantastic figure had always attracted male glances.
A3: The mother forgave her son one last time as she really understood him.
C20: The victim's moaning made the whole situation even worse.

Teacher – French-English written translation task:

B12: Ever since that time Louise has been a famous poet.
B16: He bought a new hat when he was getting married.
D25: The lawyer's visit may have a positive influence on the situation of the accused.
B13: Without his lawyer the suspect refused to answer any questions.
C21: The officer's question overcame the inspector's uncertainty.
D32: That girl's intelligence is characteristic of that family's heritage .
E38: Better motivation is the condition for an extension of the employment contract.
C17: The parents' initiative followed the doctor's call.
D28: The husband's confession is logical in this situation.
B10: He has to reach home before daybreak.
C24: His enemy's victories prevent the champion's exercises.
E36: It's the death of her grandmother that's causing the girl's sadness.
E33: It's Carol's visit that's engendering the neighbor's feelings.
B15: Instead of flying we'll go by car.
B11: Due to the politician it's no longer possible to work in a foreign country.
A3: He supported his ministers as long as they had a majority.
A1: A rock falls because it is heavy.
C19: A husband's coldness provokes a wife's hate.
A6: Sarah will go home if she can't work any longer.
E40: The manager's assassination is the subject of the workers' discussion.

E34: The positive balance is an indicator of consumer purchasing.
A4: I hate you as much as I love you.
D29: Consumption of vitamins is indispensable for the baby's health.
E35: The accident has resulted in internal injuries.
D30: This incident is tragic in this phase of development.
C23: The young man's murder gave rise to many rumors.
D26: The inception of a system of law is a necessary condition for the evolution of a state.
C22: The police enquiry accounted for the thieves' behavior.
D27: Formation of oxygen is essential for plant life.
C20: Julie's letter calls on other girls to help.
A8: Certainly there are genuine miracles; I've seen some.
A7: The pupils refrain from talking until the teacher has finished his explanation.
A2: The boy waited three weeks after that girl disappeared .
B9: After Lucas was born his mother stopped working.
C18: It's Philippe's outstanding results that are responsible for his parents' enthusiastic reaction.
E37: This punishment is the outcome of all the stupid things you've done.
B14: With his mother's help Eric dared to have a go at it.
A5: Julie accepts Paul's offer to get him to shut up.
E39: The decline in unemployment is the result of slashing state aid.
D31: It's difficult to give the praise to the children's ideas in view of the poor state of the headmistress.

Student – German-English verbal translation task (transcribed):

E39: The professor's birthday party was the reason for the visit.
C19: That Peter's mother could never forgive him unfortunately it meant his denial.
B10: Before the performance the actors should eat something.
A4: The teacher is attempting to help his pupil by giving him free after-school help.
A1: As regards to the drug dealer scene there is quite a lot of problems because the police are using Rauschgiftfahnderspuren to solve crime.
D28: The understanding of guilt works in every case for a court case.
B9: After the wedding ceremony the whole party moved into the restaurant.
D27: What the father said was helpful to the discovery of the child.
C17: The destruction of the old school was followed by the building of a new office block.
D31: At the same time the prince entered there was a cry by the princess.
E33: The reason for George's behavior was due to the fact that his mother visited.
A3: In short all participants were punished because none of them wanted to take responsibility.
E40: The giving out of the order was the reason for his speech.
C18: Ina caused many car accidents because of her failing eye sight.
A6: We will not be able to go out on an excursion if the weather becomes worse.

E38: In order to have a longer contract you must always arrive on time at work.
E36: The reason Mike was always unfaithful was due to Christine's constant complaining.
B16: Without taking the party into consideration the son liked to go home.
B14: I am not going to the theatre with Dieter's mother.
A8: Hans should tell his colleagues beforehand if he really wants to do what he wants.
B12: Since his accident Kurt has become a completely different person.
A7: Ira should find an apartment before she registers at the university.
D32: The girl's disappearance is probably due to the break-in at the house.
D25: The lawyer's visit comes directly thanks to the plaintiff's writing.
C20: Petra's shouting clears up her behavior that she showed on that day.
E34: As a result of the accident the person had serious injuries.
D29: These instructions will affect quite clearly the theft.
B11: Due to Christa's Unfähigkeit the theatre group now finds itself in this position.
C23: As a result of the lady neighbor kissing the married man there was much gossip in the village.
E35: They cleared up the new contract by stating more clearly work times.
B13: Thanks to the help of the kindergarten crash the community could only now arrange the festival.
D26: The reason Mary is so successful with men is due to her lovely hair.
A5: The girls took advantage of sharing a ride with someone so that they can get home earlier.
C24: Constant begging enabled him to get the job.
B15: Instead of complaining Michael should offer his help.
A2: The personal department celebrated its success after the negotiations with the management were finished.
D30: The sweet singing of her mother helped her daughter to sleep better.
E37: That he had to spend a long time in prison was due to the fact of his crime.
C22: Through questioning the police cleared up what they found at the crime scene.
C21: Through the neighbor's evidence it was found out that the boy stole the flowers.

Student – French-English verbal translation task (transcribed):

A7: Character traits form before we even know about them.
B14: With all these tourists the village is alive.
E38: The precise response is a condition for integration into the group.
B9: After traveling to France Anita must visit her grandmother.
E40: The engine's repairs were the subject of the lesson these last few weeks.
A1: Frank believed Anna because he loves her.
B10: Before the war he used to work in a factory.
A5: Sandra works hard so that people can hear her.
D27: The interpretation of the novel is essential for the reader's understanding.
A6: If there is a problem with the engine Gustav will repair his car.

C22: Jean-Luc's kiss explains his feelings for Claire.
D30: The people's laughing is very comical in this part of the play.
E35: The cancellation of participation is the consequence of events of the last few days.
C20: The shouts of the care-taker called for a reaction by the doctor.
C24: The details of the last few days prevented the grandfather's operation.
E33: His bad behavior is of course the reason for his being fired.
D32: The beauty of this forest is characteristic of the good treatment by young people.
C21: The neighbor's response backed up that the young man stole the flowers.
B12: Since being taken on the teacher is now considered as a member of the university body.
C17: The professor's call followed bad performance by pupils.
A2: After failing his exam the woman waited for her son.
B15: Instead of asking questions it would be better if we gave him our solutions.
A3: Where there is life there is hope.
B11: He won much money due to his victory.
D29: Regular check-ups are very much needed for the good running of an engine.
A4: The girl works as much as she can.
D31: The group's success is commendable considering the difficult conditions of their training.
C19: The teacher's response provokes the class's disobedience.
D25: The mother's initiative had a bad influence on her son's development.
C18: The car's bad performance causes the terrible accident.
D26: The teacher's letter is necessary for the person continuing his studies.
B16: Having arrived in Paris Silvia wants to organize a big party.
E37: The reason for his bad marks is his inattention.
B13: Excluding breakfast the room 500 francs per day.
E36: The train's delay was caused by traffic jams.
E34: The decline in unemployment could be an indicator for a better economic situation in the country.
D28: Her mother's presence is quite illogic considering the age of the daughter.
E39: The population growth is a positive development thanks to action by the state.
A8: Lucie cannot help her friend because it is too late.
C23: The inventor's activities produce extraordinary ideas.

Student – German-English written translation task:

D29: The neighbor's declaration will affect yesterday's events quite positively.
A6: Waiters would earn more tips if they were friendlier to guests.
B14: Without haste Cora must speak once again with her boss.
C21: The crime's appearance increased suspicion.
A5: The young woman married the stranger so that she could finally obtain residency.
A4: The cat was trying to catch the mouse by cornering it.
C17: Thankfully after the turbulent events of the last few days there was calm.

C18: The girl's hot-headed temperament always caused quite a lot of turmoil.
E37: The police had just cause in arresting the gang.
D30: The ointment soothes inflamed skin.
B11: Unfortunately as his grandmother was ill Uli could not come any earlier.
E40: The girl's father was in a bad mood and this caused her to vanish.
D28: The daughter's poem had a soothing effect on her mother's bad mood.
A8: Please do not be angry with Paula if she arrives a little late for her appointment.
D32: It is quite possible that the boy's kidnapping and the bank robbery are directly related.
E33: By losing his briefcase the man became angry.
E34: The reason for the huge fine was his stupidity.
A1: The student had to ask her parents for help since her grant had run out.
D27: The student's declaration will help his application for a semester off.
B10: Before the exam Elke was very nervous.
B9: After his lecture the professor will be quite tired.
A2: After eating his ice-cream the young boy wanted to drink some lemonade.
D26: Her sweet singing is the reason why the girl has so many job offers.
B15: Instead of sympathy Caroline required his help a lot earlier.
B13: Through her sister's involvement Heike managed to save the children's clinic.
E35: In the contract it is stated that the renter is responsible for ensuring order.
E39: The reason for the trip was a visit to the museum.
D25: The brakes failed and this was of course all to do with the car's age.
A7: Paul really has to write another letter before he leaves the office.
C23: His future wife's answer led a smile to appear on his lips.
D31: The quick reaction of the fire service and the anonymous tip are no doubt connected.
B12: Since the last family get-together the sisters have not seen each other.
C24: Her teacher's report got her the apprenticeship.
E36: The son's promise was the reason for his silence at the funeral.
E38: To improve the girl's mental state regular sessions with a therapist are castled for.
B16: As it is his girl friend's birthday Alex must of course buy her a present.
C22: The witnesses' statement could not clear up the facts of the case.
C19: For a long time men have been looking admiringly at the actress's fantastic figure.
A3: His mother forgave him again as she very much understood why he did it.
C20: The victim of the accident was moaning and this made the whole situation much worse.

Student – French-English written translation task:

B12: Since then Louise has been a famous poet.
B16: As he was getting married he bought himself a new hat.
D25: The lawyer's visit could have a positive influence on the accused's predicament.

B13: As his lawyer was not present the accused refused to be questioned.
C21: The gendarme's (French policeman's) question put doubt into the inspector's mind.
D32: This girl's intelligence is quite characteristic of the family's heritage.
E38: To obtain a contract renewal the employee must be better motivated.
C17: What the parents did was a result of the doctor's call.
D28: The husband's confession is quite a logical thing to do in this situation.
B10: Before sunrise he must be at the house.
C24: His enemy's victories prevented the champion from training.
E36: As her grandmother has died the young girl is quite sad.
E33: The neighbor's feelings all stem from Carol's visit.
B15: Instead of flying we shall take the train.
B11: Thanks to politicians we can no longer work abroad.
A3: As long as his ministers held the majority he supported them.
A1: A stone falls because it is heavy.
C19: Her husband's coldness led her to hate him.
A6: Sarah will go home if she cannot work anymore.
E40: The employees have been gossiping about their director's assassination.
E34: As a result of consumer consumption the company's results have been positive.
A4: I hate you as much as I love you.
D29: Consuming vitamins is vital to the baby's health.
E35: As a result of the accident the person had internal injuries.
D30: It is terrible that the incident occurred in this phase of development.
C23: The man's assassination has caused much gossip.
D26: The creation of laws is necessary for a country's development.
C22: Through police inquiries it was worked out why the robbers behaved in such a way.
D27: Oxygen formation is essential to the life of plants.
C20: Julie's letter called for other girls to help her.
A8: Of course miracles exist; I have seen them happen.
A7: The pupils did not speak until the teacher finished his presentation.
A2: The boy had to wait three weeks before the girl appeared.
B9: Since Lucas's birth his mother no longer works.
C18: Philippe's wonderful results have made his parents immeasurably happy.
E37: This punishment is a result of your clowning around.
B14: With his mother's help Eric is going to go off on an adventure.
A5: In order to quiet him down Julie accepted Paul's offer.
E39: The decrease in unemployment has come about due to the reduction in state subsidies.
D31: The children's ideas are not really commendable considering the poor condition of the headmistress.