‘Translating Science’: an empirical investigation of grammatical metaphor as a source of difficulty for a group of translation trainees in English-Italian translation

by
Giuseppe Palumbo

Submitted for the Degree of Doctor of Philosophy

Department of Languages and Translation Studies
Faculty of Arts and Human Sciences
University of Surrey

February 2008

© Giuseppe Palumbo 2008

Abstract

The present work reports on an empirical study aimed at investigating the difficulties emerging in the translation of scientific texts from English into Italian for a group of student translators at an advanced level of training. The study is based on the assumption that scientific discourse presents a set of lexico-grammatical features (a "syndrome", in M.A.K. Halliday's words) that contribute to making it opaque for non-specialists. This syndrome is largely the result of what Halliday calls "grammatical metaphor", which is manifested principally as a tendency towards heavy nominalisation. In particular, it was hypothesised that as the density of grammatical metaphor, and the heavy nominalisation this leads to, increases, so does the difficulty experienced by translators, especially as regards text comprehension. The hypothesis is investigated by analysing a range of datasets related to the translation process and the translations produced by a small group of subjects (N=5) for three English source texts, chosen with a view to presenting increasing degrees of nominalisation. Specifically, the analysis looks at the editing performed by the translators on their own drafts, the degree of inter-translator variation manifest in the target texts and the errors identified in the target texts. Relevant signals of processing effort are sought in each set of data, and these signals are taken to point to the difficulty experienced by the student translators. After triangulating findings form each dataset, it was concluded that the hypothesis underlying the study could only partially be supported: the two more nominalised texts were associated with a higher number of difficulty indicators but no further differentiation emerged between them. From a methodological viewpoint, the combination of data sources and methods employed for analysing the data (Choice Network Analysis and Error Analysis) proved to be reliable for identifying a cluster of ST segments observed to be difficult for translators on account of their high nominal density.

Table of Contents

List of tables and figures vii

List of abbreviations x

Acknowledgments xi

Chapter 1 INTRODUCTION

1.1 Aims and scope of the investigation 1

1.2 Modelling the translation process 2

1.3 Difficulty, competence, evaluation 4

1.4 Research Question 6

1.5 A preliminary note on methodology 7

1.6 Overview of the present work 10

Chapter 2 LITERATURE REVIEW: A FRAMEWORK FOR STUDYING SCIENTIFIC DISCOURSE IN TRANSLATION

2.0 Introduction 12

2.1 Empirical Studies of Translation 13

2.2 Investigating translation as process 14

2.2.1 Process-oriented studies adopting verbal report procedures 15

2.2.2 Process-oriented studies based on triangulation 17

2.2.2.1 Logging target-text production – *Translog* 18

2.3 Choice Network Analysis 22

2.3.1 Exploring text difficulty through CNA: the role of translation accuracy 25

2.4 Focus on the Source Text 26

2.5 The Language of Science 29

2.5.1 Scientific language as an LSP 29

2.5.2 The 'rhetoric of science' 34

2.5.3 D. P. Hayes' "inaccessibility of science" 36

2.5.4 M.A.K. Halliday's description of scientific English 38

2.6 Translation error analysis – a brief overview of past research 42

2.6.1 Traditional error categories in translation evaluation 42

2.6.2 Functionalist approaches to error analysis 43
2.6.3 Translation as selection – Pym's categorisation of errors 44
2.7 Problems and difficulties in translation: one and the same notion? 47

Chapter 3 METHODOLOGY

3.1 General methodological considerations 58
3.1.2 Types of empirical study 60
3.2 Design issues 61
3.3 Materials: Selection of the Source Texts 63
3.3.1 Relevance of Halliday's model of register 64
3.3.2 Grammatical metaphor 65
3.3.3 Degree of metaphoricity as a selection criterion 68
3.4 Subjects 72
3.5 Data collection 74
3.5.1 Organising the translation tasks 75
3.5.2 Real-time keystroke logging 76
3.5.2 Translation errors 77
3.6 Overall analytical procedure 79

Chapter 4 AN ANALYSIS OF EDITING AS AN INDICATOR OF EFFORT ORIENTATION

4.1 Introduction 81
4.2 Studying target-text editing by translators 81
4.2.1 Editing by student translators working into L2: a summary of Campbell's (1998) study 83
4.2.2 Implications of Campbell's findings for the present work 84
4.3 Studying editing in keystroke logs 86
4.3.1 Identifying types of edit as indicators of effort orientation 88
4.3.2 Counting edits 92
4.4 Frequency and distribution of editing in the three tasks 95
4.4.1 Aggregate editing in the three tasks 95
4.5 Mapping editing effort onto the STs 100

Chapter 5 ANALYSING TARGET TEXTS FOR CLUES ON SOURCE-TEXT DIFFICULTY

5.1 Introduction 106

5.2 Analysing choice networks 107

5.2.1 Identifying units for analysis 109

5.2.2 Differentiating between TL renditions 110

5.2.2.1 Variation across TTs at the structural level 111

5.2.2.2 Variation across TTs at the lexical level 115

5.2.2.3 Examples of variation identified on a case-by-case basis 116

5.2.3 Summing up 121

5.3 An overall picture of variation in the three translation tasks 123

5.3.1 Measuring variation in TT renditions 124

5.4 Using CNA to identify textual 'hot spots' 128

5.4.1 A closer look at ST 'structural' hot spots 132

Chapter 6 ERRORS AS SIGNALS OF DIFFICULTY

6.1 Translation errors 134

6.2 Criteria for Error Analysis of the TTs 134

6.2.1 Non-binary errors (N=79): some examples 136

6.2.2 Binary errors (N=120) 139

6.2.2.1 Inaccurate lexical equivalents in the TT (n=75) 141

6.2.2.2 Inaccurate transfer of ST content (n=45) 143

6.3 Distribution of errors within and across the three tasks 150

6.4 The interaction between lexico-grammar and domain knowledge 152

Chapter 7 TRIANGULATING FINDINGS

7.1 Introduction 156

7.2 Integrating findings from the previous analytical steps 158

7.2.1 Intra-subject variation hot spots and TT errors compared 160

7.2.2 Inter-subject variation hot spots and TT errors compared 161

7.2.3 Collating results 162

7.3 Back to the Research Question 167

7.4 Some reflections on methodology 169

7.5 Implications for training, evaluation and research 172
7.6 Concluding remarks 173

Appendix 1: Source Texts used for the study 175
Appendix 2: Segmented Source Texts and explanatory notes 177
Appendix 3: Binary and non-binary errors in the target texts 186
Appendix 4: Editing process as reconstructed from keystroke logs 195
Appendix 5: Comparison tables used for Choice Network Analysis on the TTs 207

References 213

List of tables and figures

Tables

2.1 Range of lexical difficulty in selected text categories (based on Hayes 1992: 739) 37

2.2 Mapping of notions of translation problem/difficulty onto Toury's analysis of the concept 56

3.1 The relationship between metafunctions and register variables 64

3.2 Measures of metaphoricity in the STs 71

3.3 Information on the participants to the study 74

4.1 Dimensions of editing, from Campbell (1998: 142) 84

4.2 Frequency of edits in Task 1, per category and translator 96

4.3 Frequency of edits in Task 2, per category and translator 96

4.4 Frequency of edits in Task 3, per category and translator 96

4.5 ST segments on which editing effort was concentrated 101

5.1 ST segment count according to type/token ratio of TL renditions at structural level 125

5.2 ST segment count according to type/token ratio of TL renditions at lexical level 125

5.3 ST segments resulting in strong variation across TT renditions 129

5.4 Segments containing lexical items repeated from previous segments 130

5.5 ST hot spots in terms of TT variation 131

5.6 Hot spots as a percentage of total segments in the STs, per category 131

6.1 Non-binary errors in the three tasks (N=79) 137

6.2 Binary errors in the three tasks (N=120) 140

6.3 Distribution of binary errors and non-binary errors in the three tasks (per translator) 151

6.4 Distribution of binary errors (N=120) in the tasks (per subcategory and translator) 152

7.1 Nominalised ST segments leading to increase in processing effort or TT content error against total of nominalised ST segments (threshold = 2 nouns per segment) 166

7.2 Nominalised ST segments leading to increase in processing effort or TT content error against subtotal of nominalised ST segments (threshold = 3 nouns per segment) 166

Figures

2.1 Choice Network for the translation of *Anglicare* (English-Arabic; source: Hale and Campbell 2002: 20) 23

2.2 Choice Network for the translation of *Anglicare* (English-Spanish; source: Hale and Campbell 2002: 20) 23

3.1 Place of the three STs along a cline of increasing grammatical metaphoricity 72

4.1 Counting edits in successive TL renditions typed by VM for segment 4, ST 2 93

4.2 Couting edits in successive TL renditions typed by CB for segment 35, ST 1 94

4.3 Percentage of each category of edits in the three translation tasks 97

4.4 Cross-task variation in editing, per category, against variation in ST metaphoricity 99

4.5 Textual segments in the STs on which editing was concentrated in each task (with indication of editing orientation) 103

5.1 Translations of segments 17-22 in ST 1 112

5.2 Translations of segment 29 in ST 2 113

5.3 Translations of segment 31 in ST 1 114

5.4 Translations of segments 24-25 in ST 114

5.5 Translations of segments 62 to 65 in ST 3 (segment 63 omitted) 115

5.6 Translations of segment 4 in ST 2 117

5.7 Translations of segment 12 in ST 2 119

5.8 Translations of segments 13 to 19 in Task 2 120

5.9 "Comparison table" used to differentiate between renditions 124

5.10 Relationship between type/token ratio of 4/5 and number of translators providing distinct renditions 128

5.11 Hot spots in ST 1 (structural level) 132

5.12 Hot spots in ST 2 (structural level) 132

5.13 Hot spots in ST 3 (structural level) 133

6.1 Translations of segments 62 to 65 in ST 3 138

6.2 Translations of *rate* in segment 55, ST 2 (the segment is only partially reproduced) 142

6.3 Inaccurate TT renditions of segment 7, ST 2 144

6.4 Translations of segments 65-66, ST 1 145

6.5 Translations of segment 4 , ST 2 146

6.6 ST segments with syntax as contributing factor to erroneous renditions 148

6.7 Translations of segment 54, ST 1 148

6.8 ST segments misinterpreted on no apparent syntactic grounds 149

7.1 Triangulation of findings 157

7.2 Segments emerging as ST-oriented editing hot spots (intra-subject variation in editing based on log data) 160

7.3 Segments resulting in ST-oriented content errors (Error Analysis data) 160

7.4 ST-oriented editing hot spots matching segments leading to content errors in the TTs 161

7.5 Segments associated with high structural variation in TT solutions (inter-subject variation data based on TT solutions) 161

7.6 Inter-subject structural hot spots matching segments leading to content errors in the TT 162

7.7 Structurally difficult ST segments 162

7.8 ST segments with high error-engendering potential (Error Analysis data) 163

7.9 Final consolidated cluster of structurally difficult ST segments based on the 3 datasets 163

7.10 Hypothesised association between increase in ST grammatical metaphoricity and the three dimensions observed in the study 168

List of abbreviations

CED	Collins English Dictionary
CNA	Choice Network Analysis
LSP	Language for Special Purposes
MT	machine translation
SL	source language
ST	source text
TAPs	think-aloud protocols
TL	target language
TT	target text

Acknowledgments

There are many people I should give special thanks to for the enormous help they gave me over the long journey that led to the present work. Prof. Margaret Rogers, who supervised the thesis, was a constant source of invaluable support and a continuous source of inspiration – I hope she realises that these are not just conventional formulas. Prof. Khurshid Ahmad, who was the initial supervisor for the thesis, taught me a lot, but first of all that research is all about evidence. Prof. Federica Scarpa, who acted as collaborative supervisor for the thesis, is the person who initiated me to translation research and never failed to provide me with moral and material support all along the journey. I am also immensely grateful to Prof. Maria Teresa Musacchio, a friend and colleague whose passion about translation and research never failed to fuel my own. Prof. Andrew Chesterman provided penetrating feedback at an early stage of the project. Prof. Marina Bondi, Prof. Marc Silver and other colleagues at the University of Modena and Reggio Emilia were enormously patient with me as I struggled to combine work on this project with my duties as a teacher and researcher there.

Special thanks go to the teaching and researching staff of the SSLMIT of the University of Trieste, particularly Prof. Christopher Taylor, Prof. Lorenza Rega and Prof. Marella Magris, for the suggestions and the support they gave me as the research progressed. I am also particularly grateful to the group of students of the SSLMIT who took part in this project.

I am also very grateful to Prof. Arnt Lykke Jakobsen of the Copenhagen Business School, who provided me with a copy of the *Translog* software package, and to Prof. Luisa Barbieri and Cristina Siligardi of the Department of Materials and Environmental Engineering at the University of Modena and Reggio Emilia, who kindly accepted my invitation to participate in the project as the domain experts asked to evaluate the translated texts.

Many other colleagues I met at various conferences and seminars also gave fundamental inputs to the research presented here – my thanks go to all of them.

Needless to say, any fault found in the present work is entirely my own.

Chapter 1
INTRODUCTION

1.1 Aims and scope of the investigation

The present work reports on a study aimed at investigating the difficulties emerging in the translation of scientific texts from English into Italian. In particular, the aim of the study is to shed light on the difficulty associated with syntactic aspects of the chosen source texts, an issue that has received scant attention on the part of researchers investigating scientific translation or, more generally, the translation of Language for Special Purposes (LSP) texts. Lexical items, or more specifically terminology, are often taken to be the main obstacle for translators tackling a scientific text. The relatively recent tradition of studies that can be grouped under the label 'rhetoric of science' has shown, however, that scientific texts are highly rhetorical artefacts and that the style of scientific texts (including their syntax) can be an independent or at least specific source of difficulty for the readers of those texts – including those who are trained scientists but specialise in different fields from that which a given text belongs to. In parallel with the 'rhetoric of science' tradition, M. A. K. Halliday has elaborated a description of scientific discourse based on systemic-functional grammar, showing how the construction of such discourse is a result of the joint effect of lexis and syntax. Terminology, in this perspective, is not a linguistic resource developing on its own, irrespective of the meaning-making mechanisms of non-terminological resources. It is, rather, a long-term product of the very same mechanisms that scientific discourse resorts to with higher frequency than other discourses.

Based on the Hallidayan characterisation of scientific discourse, the study presented here focuses on one specific lexico-grammatical feature, i.e. what Halliday calls "grammatical metaphor", in order to observe what role it plays as far as the difficulty experienced by a group of trainee translators faced with scientific texts is concerned.

The investigation is placed at the interface between different areas of translation research. One is the investigation of the cognitive processes involved in

translating. This is an area where a substantial number of studies have appeared in the last 20 to 25 years; very few of them, however, have investigated the connection between specific syntactic aspects of the source texts and the mental processes of translators. Another relevant area is that of translation competence. In particular, in the context of translator training, source-text classification schemes are rarely related to the degree of difficulty of the texts, an aspect which is of fundamental importance if the criteria for text selection are to be established according to some kind of pedagogical progression (notice, however, that this study is not concerned with the development of translator competence). Finally, and in close relation with translation competence, a third relevant area is that of translation evaluation. In particular, in contexts where translation competence must be assessed (e.g. for training or accreditation purposes) the need to categorise source texts according to degrees of difficulty is evident, but no universally accepted parameters have been established.

In the rest of the present chapter, a brief illustration will be given (1.2) of the model of the translation process on which the study is based; the notion of difficulty will then be introduced, and its relation with the already mentioned areas of translation competence and translation evaluation will be elaborated on (1.3); after that, the research question the study is intended to investigate will be presented (1.4), followed by some introductory remarks on the methodological approach adopted by the investigation (1.5); an overview of the study (1.6) will finally close the chapter.

1.2 Modelling the translation process

For the purposes of the present investigation, the *translation process* is to be understood (Hansen 2003: 26) as "everything that happens from the moment the translator starts working on the source text until he finishes the target text"; in particular, it comprises everything "from every pencil movement and keystroke, to dictionary use, the use of the internet and the entire thought process that is involved in solving a problem or making a correction".

More specifically, from the perspective of the "thought process" of the translator, the investigation accepts the operational model proposed by Gile (1992), which sees the translation process as comprising two stages: a comprehension stage

and a reformulation stage. At the comprehension stage the translator puts forward a hypothesis regarding the meaning of a given source text (ST) unit. This hypothesis is based on both the translator's linguistic and extra-linguistic knowledge and the results of his/her research in reference materials. The hypothesis is then tested for plausibility. If this test gives a positive result, the translator proceeds to reformulate the unit in the target language (TL) (or to analyse the following translation unit). Reformulation is based, again, on the translator's linguistic and extra-linguistic knowledge and on the result of research into reference materials and leads to a first TL rendition. This TL rendition is tested for both fidelity to the ST and acceptability in the TL (at the grammatical, lexical, terminological and stylistic levels). Comprehension errors, one of the focal points of the process for the present investigation, result from a malfunctioning of the plausibility test, although their origin lies primarily in a wrong hypothesis made about ST meaning. This wrong hypothesis may be due to a lack of knowledge on the part of the translator or to inadequate research strategies.

As a corollary to this basic model, the notion of translator *effort* will also be considered. Effort is to be understood, in cognitive terms, as the lack of automaticity in providing a translation solution and therefore as the mental processing carried out by the translator in connection with the element to be translated. Two basic effort orientations will be identified, on the basis of the two process stages of comprehension and reformulation as proposed by Gile. 'ST-oriented' effort is directed at assessing the meaning of a given ST element (from an individual morpheme or word to a whole sentence or even a paragraph). 'TT [target text]-oriented' effort has to do with the processing directed at evaluating the translation solution in terms of TL correctness or appropriateness (at the lexical, syntactic, textual or pragmatic levels). The notion of effort is particularly relevant in relation to some of the datasets employed for the investigation (e.g. the editing performed by translators on their own drafts). Where necessary, further distinctions will be introduced as far as the orientation of the effort is concerned, e.g. between the editing effort oriented towards ST terminology and the other instances of ST-oriented editing.

1.3 Difficulty, competence, evaluation

As already suggested, the issue of difficulty is relevant in different areas of translation research, particularly in connection with translation competence and translation evaluation. As far as translation competence is concerned, the question of the difficulty attached to the source text as part of a translation task is of immediate interest to either trainers or other researchers investigating the development/acquisition of translation competence. As stressed by Kelly (2005: 117), in a translator training context "sequencing and hence selection of materials and texts are complex issues dependent on a variety of factors". The degree of difficulty of the selected texts is certainly one of these factors, but "as yet there do not seem to be clear criteria to categorise texts for translation according to degrees of difficulty" (Orozco Jutorán 2003: 224). In a research context, difficulty turns out to be an equally elusive notion:

> it would appear that researchers who wish to study the translation process must face the problem of their choice of texts to translate. The problem [...] comes not from the "definable" parameters (source, translation brief, language pair, length, main subject, written original in the source language, etc.) but from the parameters for which there is no generally accepted definition and/or classification (i.e. level of difficulty) (Orozco Jutorán 2003: 233).

Such parameters, it can be added, have to do with a variety of factors, which can be schematically listed as follows:

- the translator's linguistic and stylistic competence;
- the individual translator's own awareness of an incorrect choice of equivalent;
- the translator's knowledge of the subject matter of the ST (which for trainees can be assumed to be scarcer than for experienced professional translators specialised in a given domain);
- the inherent ambiguity or lack of clarity of the ST (note, however, that ambiguity is in itself a relative notion, as a text can be highly ambiguous for certain readers and perfectly clear for other readers who are more familiar with the subject matter or the genre);
- the translatability of the text into a certain language at different levels: semantic, lexical, syntactic or pragmatic.

A consideration of all of these factors goes well beyond the scope of the present investigation. The study focuses on one aspect only (the lexico-grammatical features that may give rise to the ambiguity or lack of clarity of source texts), but, in discussing text difficulty in translation, the interaction between the different factors listed above (and possibly others) should always be taken into account. A 'place to start' has nevertheless to be chosen, and for the present investigation a decision was taken based on the considerable amount of attention that the structural aspects of scientific texts have received (at both macro- and micro-level) over the last few years (cf. the review of studies in Chapter 2). One particular structural aspect, i.e. the strong tendency towards nominalisation, was chosen around which a consensus has emerged among scholars concerning its 'problematic' nature. This was deemed to be a sufficiently specific basis to start investigating 'difficulty' in relation to scientific texts in translation. Seen from a methodological perspective, this decision appears to be in line with the recent tendency of translation scholars to formulate "more and more small-scale, and hence more testable hypotheses" (Toury 1995: 239).

As far as translation evaluation is concerned, text difficulty can be seen to establish an obvious link with one particular aspect of evaluation, i.e. translation error. With the emergence of functionalist approaches to the study of translation, attempts at defining the nature of errors have lent prominence to factors that were external to the translation process (with process here intended as the actual transfer carried out by the translator; see 1.2). In the words of House (1997: 16), propagators of functionalist approaches "believe that the original is a quantité négligeable and emphasize the translation's total dependency on its purpose and its recipients"; House (1997: 159) believes "this recent shift of focus in translational studies to be fundamentally misguided". Her reservations concerning functionalist approaches may be considered overstated, but a reconsideration of the role played by the ST is perhaps not gratuitous.

The relatively recent tradition of empirical studies in translation seems to have implicitly accepted House's plea for a consideration of the 'internal' aspects of the translation process. Research on the nature and causes of errors in translation can be said to have received a new impetus not only from studies of the mental processes of translators but also from research on the acquisition of translation competence and on translation evaluation (both in a teaching and a professional setting). Such studies,

although centred on very different aspects of translation, are contributing to the creation of an interrelated set of findings related to errors (and to methods of identifying and observing them). These findings could in turn lead to a typology of errors exempt from the inconsistencies and the rigidity characterising traditional error categories. Most of the work carried out on translation errors, however, still concerns general-language translation. Reference to specialised translation relating to texts which presuppose domain knowledge usually arising from explicit training is only made in passing and no systematic approach to studying why and how translators 'get it wrong' in specialised texts has been proposed.

1.4 Research Question

The basic assumptions of the present investigation, discussed in detail in Chapter 2 (see in particular 2.5), can be summarised as follows: 1) scientific discourse is as persuasive as it is informative; 2) the style of scientific texts is based on rhetorical devices as much as that of texts in other domains; 3) scientific discourse is becoming increasingly opaque for non-specialist readers; 4) a "syndrome of features" (Halliday 1993a) contributes to this opacity at the lexico-grammatical level. In particular, according to Halliday, in English scientific discourse this syndrome is largely the result of what he calls "grammatical metaphor", which in turn is manifested principally as a tendency towards heavy nominalisation.

The investigation starts from the assumption that translators tackling a scientific text are often faced with a highly opaque text, such opacity being engendered as much by the subject matter as by the particular, heavily nominalised style used by writers. In particular, the question that the investigation will try to give an answer to is the following:

> *What is the relationship between grammatical metaphor, specifically the heavy nominalisation it leads to in English scientific texts, and the difficulty experienced by Italian student translators facing such texts?*

This question will be investigated by analysing different sets of data associated with the translations produced by a group of subjects for three different source texts,

which were chosen so as to present different degrees of nominalisation (in particular, each text was increasingly more nominalised than the others; see 3.3 for details). Relevant signals of processing effort will be sought in each set of data, and such signals will be taken to point to the difficulty experienced by translators. In particular, the analysis will look at: 1) the editing performed by the translators on their own drafts; 2) the degree of variation between the target texts; 3) the errors identified in the target texts. As a general hypothesis, it is expected that, as nominalisation increases in the source texts, so does the processing effort made by the translators in relation to the structural (as opposed) to lexical aspects of the STs. Three further, smaller-scale hypotheses can be proposed, in relation to each of the three sets of data mentioned above. These hypotheses can be formulated as follows:

a) as nominalisation increases in the source texts, so does the number of edits relating to ST structural (as opposed to lexical) aspects;

b) as nominalisation increases in the source texts, so does the degree of variation in TT final renditions, especially as regards the structural (as opposed to lexical) profile of the renditions;

c) as nominalisation increases in the source texts, an increasing number of TT errors associated with ST structural (as opposed to lexical) aspects is observed in the TTs.

The analyses carried out on the three datasets will focus on the way these relate to the lexico-grammatical aspects of the STs. The proposed distinction between the structural and lexical aspects of the texts will be considered as it emerges in the analysis of each individual dataset. Higher textual levels (e.g. aspects having to do with theme/rheme distribution) and pragmatic elements will be ignored.

1.5 A preliminary note on methodology

The approach adopted in the present study closely follows that employed in other studies of text difficulty, most of which, however, have been carried out outside the field of translation studies. The aim here is to conceptualise 'difficulty' in translation as a function of certain stylistic features of texts. In translation research, most studies

examining difficulty deal with aspects other than comprehension, focusing on the reformulation/re-expression phase. The source text is often taken as a given and, as regards ST comprehension in particular, translators are often implicitly compared to ideal native speakers of the source and target languages whose problem lies mainly in deciding on the best way to transfer meaning in accordance with the translation brief. The stylistic features of source texts have been considered mainly in the field of Machine Translation, where they are seen as one of the determining factors of the translatability of a given text (see 2.4).

Outside translation, research on text complexity has often been triggered by concerns over the increasing opacity of the texts written by specialists (cf. Hayes; 1992; Halliday and Martin 1993). While the emphasis was traditionally placed by researchers on the role played by terminology, more and more studies have shifted the focus of research in the direction of syntax. The attempt in these studies is to see whether certain stylistic features (as determined by syntactic as opposed to lexical phenomena) contribute to the relative accessibility of a given text. Lassen (2003) has analysed the accessibility of technical manuals, starting from the assumption that this can be adversely influenced by high lexical density *and* a high density of grammatical metaphor.

The present work explores whether difficulties in the translation of scientific texts can be described and, possibly, explained in terms of the role of grammatical metaphor and the grammatical problems it gives rise to, as described in the Hallidayan model of scientific language. The explanation does not seek to identify an automatic 'cause-and-effect' mechanism triggering difficulty, such that whenever a source text presents the translator with a grammatical metaphor this invariably determines either difficulty in the process or an outright error in the target text. Rather, the explanation assumes a probabilistic nature and can be presented in a more tentative form such as "the more grammatically metaphorical a source text is, the more difficult it is likely to be for translators".

The presence of grammatical metaphor in a text is not *per se* a criterion for low text accessibility (cf. Lassen 2003). In order to explain why some texts are more accessible than others, we would have to invoke a larger set of principles likely to have influenced the readers' reactions and not just grammatical metaphor alone: "[t]ext accessibility may be presumed to be a function of both text internal and text

external criteria" (Lassen 2003: 127). In this respect, two other areas of relevance are information structure (i.e. the succession of given and new information in each sentence as the text unfolds) and the readers' socio-cultural situation. Much as in Lassen's own study, however, the present work focuses on one aspect only of the texts (style as affected by the density of grammatical metaphor) and starts from the hypothesis that as the degree of grammatical 'metaphoricity' (and the nominalisation it leads to) of a text increases, so does the processing effort required of translators, which contributes to making the text more difficult. This, however, should by no means be understood as excluding other sources of difficulty.

In order to test this hypothesis, the study employs empirical research methods and, in particular, relies on a research design directed at isolating specific features of source texts in order to observe what effects they have on the performance of a small group of student translators (N=5). In Chesterman's (2001) terminology of 'context' and 'profile' variables, the present study aims to observe the effect produced by a *context* variable – here, different degrees of grammatical metaphoricity in STs – on translation difficulty, as reflected in ST processing/editing, variation in translation solutions between translators, and errors in the TTs.

Other datasets were chosen apart from the target texts since attempting to observe the translated texts alone risks giving a limited picture of the factors engendering difficulty, for the very simple reason that the difficulty experienced by a translator during the performance of the translation task does not necessarily leave textual traces. In other words, a particular aspect of the task may actually prove difficult but then be somehow solved by the translator, so that no traces of that difficulty are left in the final product. In order to obtain a more comprehensive picture of the ways in which the style of scientific texts turns out to be difficult for translators, a broader set of data is therefore needed than that emerging from an analysis of the target texts alone. It follows that a suitable methodological framework must be developed for a description of difficulty that is capable of accounting, simultaneously, for both process and product. In this respect, the triangulation approach adopted by the present study (discussed in Chapter 3) seems to profitably serve the double purpose of ensuring a broader scope for the analysis and providing firmer ground for the interpretative claims made on the role of style in contributing to difficulty. In particular, three different datasets were employed. One is the target

texts themselves, which were submitted to a particular method of analysis (i.e Choice Network Analysis; see 2.3) aimed at measuring the degree of inter-translator variation, considered, in turn, as an indicator of processing effort. The same method of analysis was also applied to the interim translation solutions observed in the drafting process carried out by translators, which was reconstructed by looking at the the computer logs of the keyboard activity performed by the translators. The third dataset was provided by an Error Analysis of the target texts.

From a methodological viewpoint, the necessary preliminary steps for the analysis proposed in the investigation concern the following aspects:

1) a justification for the particular methods of data collection used in the study;
2) a clear specification of the features of the ST that are assumed to be relevant as far as difficulty in translation is concerned;
3) the ways in which this specification has been used in selecting the source texts submitted to the translators who took part in the study.

These three aspects are treated in detail in Chapter 3, which illustrates the study design and concludes by giving a brief overview of the analytical steps that have been followed in examining the data obtained for the investigation.

1.6 Overview of the present work

The rest of the present work is organised as follows. Chapter 2 presents a review of the existing literature, focussing at first on empirical process-oriented translation studies, then moving on to the relevant descriptions of scientific language and finally returning to translation in order to review the existing research on translation errors and discuss the mutually related notions of 'difficulty' and 'problem'.

Chapter 3 describes the methodological approach and illustrates the design the investigation is based on. In particular, it considers in some depth the notion of "grammatical metaphor" as discussed by Halliday, given that the degree of metaphoricity was taken as the basis for selecting the three source texts employed in the study.

Chapters 4, 5 and 6 each relate to one of the three datasets collected in the study and therefore share a common structure: each chapter starts by describing in

detail the procedure for data analysis and then goes on to discuss results. In particular, Chapter 4 first describes the way keystroke-based data were used to analyse the editing performed by translators on their own drafts and then illustrates how such editing is distributed in each task. Chapter 5 reports on the analysis of the target texts, aimed at identifying the extent to which they varied from each other – variation being considered a measure of the processing effort made by translators. Chapter 6 presents the Error Analysis carried out on the target texts with a view to identifying the errors that were more clearly oriented towards the STs.

Finally, Chapter 7 proposes a review and integration of the findings arising from the three previous chapters; upon this basis, the chapter evaluates whether the hypothesis underlying the investigation can be supported and then proposes some final methodological considerations. The conclusion reached in the chapter is that, while the analytical methods employed for the investigation proved to be reliable, the results of the analysis provide only partial support for the hypothesis linking an increasing density of grammatical metaphor with an increase in the difficulty experienced by translators.

The materials included in the appendices comprise the three source texts, the target texts and some of the other data employed for the analysis. Of particular importance is Appendix 2, where a note illustrates the criteria used for the constituent analysis carried out on the source texts at the lexico-grammatical level. This was the step that led to the identification of the ST segments that provided the units for analysis in relation to all the datasets employed in the investigation.

Chapter 2
LITERATURE REVIEW:
A FRAMEWEWORK FOR STUDYING SCIENTIFIC DISCOURSE IN TRANSLATION

2.0 Introduction

This work aims at studying how, in the translation of scientific texts, some linguistic features of the source text influence the performance of translators, which is evaluated by looking primarily at the target texts, i.e. the product of translation. The study, however, also takes into account data relating to the process of translation. The main focus of interest will be in the linguistic features of the source texts and in the possibility to characterise these features as specific sources of difficulty for translators. Methodologically, the study follows a 'triangulation' approach whereby hypothesis-testing is conducted on different sets of data so as to provide firmer empirical grounds to the claims made in discussing and interpreting the results of the study.

This chapter is organised as follows: section 2.1 deals with empirical studies of translation; section 2.2 focuses on studies investigating the translation process, briefly reviewing studies adopting verbal report procedures (2.2.1) and studies based on triangulation (2.2.3); Choice Network Analysis, a complementary research method to those normally employed in process-oriented studies, is then discussed (2.3); section 2.4 introduces the perspective of ST difficulty; section 2.5 presents an overview of studies of scientific language; in particular, scientific language is first presented as a variety of LSP (2.5.1) and then observed as a rhetorical artefact (2.5.2); this section also reports on a study measuring the 'opacity' of today's scientific texts (2.5.3) and then presents (2.5.4) M.A.K. Halliday's description of scientific discourse and of the primary role played in it by "grammatical metaphor"; section 2.6 goes back to translation in order to present a framework for analysing translation errors which is suitable for the needs of the present study; finally, section 2.7 discusses the related, but as yet not clearly distinguished, notions of 'translation problem' and 'translation difficulty'.

2.1 Empirical Studies of Translation

With the emergence of approaches to translation that are not based exclusively on linguistic models and methods, translation has increasingly been seen as an act of communication rather than just a linguistic product. As such, it has been observed to reflect two different categories of processes: on the one hand, *cognitive* and *psycholinguistic* processes, emerging when the attention is focussed on the mental activities of one or more translators carrying out a given translation task; on the other, *sociological* processes, scrutinized when the inquiry is into the various aspects of the social context in which a translation task takes place. Not merely a linguistic product nor an essentially linguistic operation any more, translation has thus been reconfigured as the result of an interpretative process that aims at reformulating a text in another language and operates in a given socio-cultural context. The translated text, and an analysis of it in relation to the source text and to a systemic comparison between the two languages involved, are no longer felt by researchers to be a sufficient basis to elucidate all the aspects of translation.

The attention of many researchers has therefore turned to translation (or translating) as an activity. Studies within this area generally adopt one of two very general perspectives. An 'internal' perspective is that adopted by the inquiries into the cognitive or psycholinguistic processes involved in translation. An 'external' perspective characterises the studies devoted to the relational and social aspects of translation and the social role(s) played by translators. Studies of this kind address questions such as the different meanings given to the notion of translation by different communities (e.g. the general public vs. practitioners), the self-image of professional translators or the "norms" adhered to by a given community of translators (see, for instance, Toury 1995; Simeoni 1997). The two perspectives, however, must not be seen as mutually exclusive. In areas of inquiry such as translation competence the methodology of research can be seen to alternate or integrate 'external' and 'internal' approaches: whether the focus is on the acquisition or the components of translation competence, research in this particular area is confronted with the elucidation of a diverse set of mental and relational processes and skills (cf. Orozco and Hurtado Albir 2002).

When translation is understood and studied mainly as an activity, two specific aspects acquire particular importance: on the one hand, the collection of relevant data; on the other, the set of methodological tools to be applied in the interpretation of such data. This is the area where Translation Studies has seen one of its most significant developments of the last two decades or so. Snell-Hornby (2006) and others have even come to speak of an "empirical turn", a crucial development which has come into being "from within the field" (Snell-Hornby 2006: 115) although in parallel to the interdisciplinary borrowing of theories and methods that characterised Translation Studies in the 1990s.

2.2 Investigating translation as process

In the last two decades, two general lines of research have emerged in studies of translation based on empirical methods. The first is research into translation as product, i.e. into translated texts; this is generally based on corpora of translations (either parallel or comparable) and aims either at establishing whether translated language exhibits features that set it apart from non-translated language (the so-called "translation universals") or at identifying regular linguistic patterns that can help to shed light on the strategies and techniques employed by translators in given language pairs, genres or text types. The other line of research is the one including studies of translation as process, with "process" understood, rather loosely, as what goes on in the translator's mind as s/he is engaged in the translation task.

This identification of two separate strands of research should not be seen as pointing to two completely different and separate areas of research each having its specific object. The distinction between product and process should not "ignore the fact that the one is the result of the other, and that the nature of the product cannot be understood without a comprehension of the nature of the process" (Holmes 1978: 81). In fact, it would probably be better to use the terms "product-oriented" and "process-oriented research" (rather than "research into" either the process or the product). The distinction between these two dimensions (product and process) has, in some cases, more to do with the methodological choices made by researchers than with the object of their research. So, for instance, a study such as Hale and Campbell (2002), although based on "products" (more specifically, translations produced by

trainees), is aimed at shedding some light on some cognitive aspects of translation, i.e. on "process" (see 2.3). By the same token, corpus-based studies of translation may lead to an identification of the (processual) strategies employed by translators. Corpora, for instance, have been used to investigate the claim that explicitation is an inherent feature of translated text (Laviosa-Braithwaite 1996; Pápai 2004), a feature which some researchers see as deriving from an underlying cognitive process (Oloahn and Baker 2000).

It is customary to mark the beginning of the process-oriented research tradition in Translation Studies with the publication, in 1986, of H. P. Krings' pioneering work *Was in den Köpfen von Übersetzern vorgeht*, a study investigating the use of time and reference books on the part of translators (Krings' subject were really language learners) and analyzing the nature of both the problems translators encountered and the problem-solving strategies they employed. Krings' study was based on data gathered through verbal reporting, a method originally developed in cognitive science and human information processing. Most of the subsequent process-based research in translation has made use of this method, adopting one (or sometimes more) of its possible variants. More recent studies have complemented verbal reports with other data elicitation methods in an attempt to redress what some researchers thought were the limitations of verbal reporting. The next two sections review some relevant process-oriented studies based on verbal reports (2.2.1) and studies in which verbal reports have been complemented with other data elicitation methods (2.2.2).

2.2.1 Process-oriented studies adopting verbal report procedures

Verbal reporting as a method of data collection is based on the assumption that human cognition is information processing and that information is stored in memories with different capacities (cf. Ericsson and Simon 1984). Short-Term Memory (STM) stores information from the surrounding world, i.e. the information that is heeded or attended to. The amount of this particular kind of information stored in STM is limited. Long-Term Memory (LTM) is "an enormous collection of nodes" (Ericsson and Simon 1984: 13) which can be accessed either by recognition or by way of links associating nodes to others that have been already accessed. Both these

processes bring information into STM. Information stored in STM remains accessible for further processing and for producing verbal reports. These can be of three types: introspective, retrospective and concurrent. In introspective reports, the subject of an experiment carries out a self-analysis of his/her thought processes. Retrospective reports are reports on thought processes that are given after the performance of a task. Concurrent reports take place at the same time as the task: here the subject is not asked to verbalise specific information but to think aloud.

In translation research, verbal reports are used to obtain an account of the translator's thoughts either while s/he is translating (i.e. concurrently) or at some temporal distance from the task itself (i.e. retrospectively). Verbalisation is recorded and subsequently transcribed into protocols, which may contain additional information provided by the researcher (e.g. about the participants, the materials used in the experiment, non-verbal behaviour and so on). More specifically, in the "think aloud" method of data collection a translator is asked to translate a text while concurrently verbalising as much as s/he can of his her thoughts. The verbalisation is audio- or video-recorded and then transcribed: the transcript is referred to as the "Think-Aloud Protocol" and constitutes the object of study on the part of the researchers, with or without reference to the actual recording.

Opinions on the respective validity of the various verbal report procedures differ, although TAPs are usually seen as the method which has proved "the most popular" with translation researchers (Jääskeläinen 1998: 266). However, Hatim (2001), following Fraser (1996), notes how "IRs [i.e. immediate retrospections] are thought to be superior in many respects", as "the accounts tend to be more structured, inferential and thus much more revealing than the instantaneous TAPs" (Hatim 2001: 157). Bernardini (2001), on the other hand, stresses that only concurrent verbalisation (as opposed to *post hoc* verbalisation) reflects the mental states of subjects. She also points out that, in order for the concurrent reports to be reliable, they must be strictly monological (cf. also Jääskeläinen 2000).

Researchers have used TAPs to investigate specific questions such as problem-solving strategies (Krings 1986; Lörscher 1991), creativity in translation (Kussmaul 1991, 2005), comprehension processes (Kussmaul 1995b), the use of reference materials (Livbjerg and Mees 2003) and the differences between novices and professionals (Tirkkonen- Condit 1990; Jääskeläinen and Tirkkonen-Condit

1991). After a first, enthusiastic phase (which, in chronological terms, goes from the publication of Kring's study in 1986 till the second half of the 1990s), the use of verbal reporting in process-oriented studies of translation has come under closer scrutiny, as it has been felt that the validity of the method had previously been "assumed rather than proved" (Bernardini 2001: 242). Some reservations have been put forward as regards particularly the use of TAPs with professional translators: as some of their skills have been automatised, they may by-pass STM and may therefore not be available for verbalisation. Such a risk is mentioned (although not in connection to translation) by Ericsson and Simon (1984: 15) themselves, who warn that automation means that "intermediate steps are carried out without being interpreted, and without their inputs and outputs using STM". Other concerns voiced by researchers have had to do with the potential effect of verbalisation on the process being investigated (Jääskeläinen 2000), the lack of a clear definition of the object of study in many TAP-based investigations (Tirkkonen-Condit 2000) and the lack of a clear definition of the notion of "problem" in those studies which had this notion as their object, either implicitly or explicitly (Bernardini 1999) – this last being the question which is of more immediate interest for the present study (cf. section 2.7 below).

2.2.2 Process-oriented studies based on triangulation

Most recent overviews and reviews of process-oriented studies of translation (cf. Hansen 2003; Jakobsen 2003; Tirkkonen-Condit 2006) agree on the need to combine different data collection methods so as to test hypotheses on firmer empirical grounds. In particular, in order to overcome the limitations of TAPs, other methods have been proposed that allow the researcher to gather data deriving from the observation of the translation process – quantitative data felt to be complementary to the qualitative data supplied by data collection methods based on verbal reports. Some researchers have adopted a "triangulation" approach, which is based on the convergence of different methodologies used to collect, elicit and interpret data (Jakobsen 1999; Alves 2003). The metaphor inspiring this approach assumes that "navigating through uncharted waters requires several location points to establish one's position" and that therefore several instruments of data gathering and analysis

can be simultaneously used to throw light on the nature of the translation process (Alves 2003: vii). The aim is therefore that of reducing the risk that observations of a given phenomenon are "mere methodological artefacts" (Jakobsen 1999: 19) and of ensuring that observational data are successfully validated.

2.2.2.1 Logging target-text production – *Translog*

One particular method used in most of the studies belonging to this recent "triangulation" tradition is the computer logging of the keyboard activity performed by the translator during a written translation task. A software tool such as *Translog* (Jakobsen 1999; Jakobsen and Schou 1999) can be used to elicit this kind of data. In particular, *Translog* registers all the keyboard activity performed by the translator and has a replay function that can be used to observe the typing process either in real time or at a different speed. The raw data provided by *Translog* consist of the final printout of the translation and a detailed register of the typing process. This register gives information about the timing of each keystroke, pauses, revisions, deletions, corrections and so on. When gathered for a group of subjects, such information alone can, for instance, give indications as to the process features that correlate with the quality of the target texts (assuming a suitable benchmark can be found). In other words, the researcher can verify whether the best translations are produced through a smooth text generation process (i.e. one with few revisions and deletions) or whether they are arrived at via a more uneven process characterised by frequent revisions, deletions and corrections. The replay function of *Translog* can also be used, once the text generation phase is over, as a prompt in eliciting a retrospective report on the task. While looking at the replay of his/her text generation process, the translator can comment on his/her own work, focussing, for instance, on why certain revisions were made or why a certain structure for a TT sentence was immediately discarded.

Tirkkonen-Condit (2006: 684) identifies two advantages of *Translog* over think-aloud methods: it is less obtrusive for the translators and less labour-intensive from the researcher's viewpoint. Tirkkonen-Condit also notes (2006: 684) how, if combined with other data such as eye-gaze patterns, information obtained using *Translog* could enable researchers to test on a larger scale hypotheses that have already been put forward as a result of small-scale studies based on think aloud and

other labour-intensive methods. As far as the present study is concerned, Translog has been used as one method for gathering observational data about the translation process; these data are integrated with information elicited through other methods and with an evaluation of products. The present investigation is thus following, from a methodological point of view, the recent tradition of studies based on triangulation. Whereas, however, most of the studies in this recent tradition are still concerned with finding out "the shared patterns that single out those translation processes that result in high quality translations" (Tirkkonen-Condit 2006: 684), the present investigation attempts to apply the same integrated methodology to an investigation focussing on a different aspect of the process, i.e. the linguistic features of the ST and their role as a source of difficulty for translators. Before moving on to reviewing other methodological approaches which are relevant in light of this change of focus, some more space will be devoted to a discussion of studies following a triangulation approach, some of which do share the focus on ST linguistic features which characterises the present investigation.

The studies collected in Hansen (2002) are for the most part based on data collected through more than one method of elicitation (typically, verbal protocols plus data from keyboard logs obtained using Translog). Livbjerg and Mees (2002), for instance, investigate the problem-solving strategies adopted by a group of semi-professional translators tackling a non-specialist text, with and without the use of dictionaries. Of more direct interest for the present work is the study by Halskov Jensen (2002) in the same volume (reporting on the results obtained by Halskov Jansen in her PhD thesis). Halskov Jensen's hypothesis is as follows: for translators of legal texts from Spanish into Danish, both the syntactic structure at clause and sentence level and the text structure are likely to be highly problematic; if a text were simplified in these respects (and without affecting technical terms and phrases), it would lead to a reduction in the number of problems experienced by translators and, consequently, to target texts of higher quality. In order to test this hypothesis, Halskov Jensen designed an experiment in which she submitted four different Spanish texts to a group of six Danish translators. The four source texts included two legal texts (both of which were judgements: one non-simplified; the other simplified) and two non-legal text (newspaper editorials: again one non-simplified, the other simplified). The features that characterise the two non-simplified texts compared to

the simplified texts are: presence of particularly long sentences (30 words or more), greater number of subordinate clauses, presence of more pre- and post-modified nouns, highly nominalised style, greater degree of ambiguity in references, greater ambiguity in the semantic relations within the text. For the four different texts, then, variation can be characterised along two axes: simplification and genre. By analysing data sets gathered through different methods, Halskov Jensen concludes that it is variation in genre that affects the quality of translations the most, while the simplification of syntactic and textual features only plays a marginal role. Translators, in other words, seem to be better equipped to tackle texts that they are more familiar with (here, newspapers editorials), whether or not these texts exhibit syntactic features that in principle can be assumed to be particularly difficult for them.

Apart from the conclusions it reaches, however, Halskov Jensen's study is to be considered relevant for the present purposes mainly from a methodological point of view. More specifically, it can be seen as a good example of an empirical investigation that (successfully) integrates various data collection methods. Besides TAPs, the process investigation is based on retrospective reports and keyboard logs: together, these sets of data provide a more complete picture of the process. In turn, they are combined with questionnaires submitted to translators and with evaluations of the target texts provided by experts. Overall, the data taken into consideration in Halkov Jensen's study include:

- reading time of the source texts;
- think-aloud protocols (based on a preliminary sight translation of the STs);
- retrospective reports;
- keyboard logs for the written translation task;
- questionnaires submitted to translators (concerning the difficulties encountered in the translations tasks);
- evaluations of the target texts (provided by professional translators and translation teachers).

Another methodologically relevant aspect of Halskov Jensen's study is her operationalisation of the notion of translation problem – a pervasive concept in translation research but one that has seldom received systematic treatment (cf. section 2.7 below). In particular, Halskov Jensen analyses TAPs following the

"interpretive reconstruction" approach (cf. Lörscher 1991: 56), i.e. one taking into consideration the foci of attention actually expressed by the verbalising subject (as opposed to the *assumed* problematic aspects of the translation task). Such a reconstruction leads Halskov Jensen (2002: 132) to identify four different levels of difficulty in the TAPs:

- "attention" – at this level, the subject is merely indicating where his/her attention is directed; no doubts are expressed on the validity of a translation solution;
- "potential problem" – this is where the subject expresses doubts but then proceeds to propose a translation solution;
- "difficult problem" – at this level, the subject express doubts and leaves further traces of processing (e.g. prolonged pauses, reading aloud of the ST segment in question, reference to extralinguistic knowledge) before verbalising a final solution;
- "unsolved problem" – this is the level at which subjects either do not verbalize or reject a translation solution, thus proposing no final rendering for a given ST element.

The study by Halskov Jensen is probably unique in what can perhaps now be called the 'mainstream' tradition of empirical process-oriented research on translation. Its uniqueness, however, lies mainly in the object of the investigation. Methodologically, the study is, as already noted above, a particularly illustrative example of how different methods of data collection are currently being combined by researchers so as to overcome the limitations of each method when used individually. Verbal reports (either concurrent or retrospective), logs of keyboard activity, questionnaires and evaluations of the target texts have established themselves as the most popular methods for collecting information about the process of translation and they are today used (in different combinations) for testing a variety of hypotheses. There are, however, other investigations of an empirical nature which share the same interest in text difficulty as the present study but fall outside the 'mainstream' tradition described above. These studies, based on Choice Network Analysis, will be reviewed in the next section.

2.3 Choice Network Analysis

Choice Network Analysis (CNA) is a research method proposed by Stuart Campbell of the University of Western Sydney. It is intended to investigate the mental processes of translators by looking at the target texts they have produced. It can be seen as a complement or an alternative to research methods based on verbal reports, such as think-aloud protocols and word-based experimental techniques. CNA and the results of studies based on it have been presented by Campbell in various papers (some of them co-authored with Sandra Hale: Campbell 1999; Campbell & Hale 1999; Hale & Campbell 2002); a general illustration of CNA and examples of possible applications are given in Campbell (2000), on which the following is largely based.

CNA examines the translations of a source text produced by a number of translators working into the same target language. The principles CNA is based on can be summarised as follows (cf. Campbell 2000: 32): target texts are a source of evidence of mental processing in translation; the texts produced by a sample of subjects translating the same texts will reveal a range of differences and similarities in the behaviours of the subjects; the larger the sample, the more likely it is that the complete range of behaviours is approached; a model of the mental processing underlying the translation of a given text in a given language combination can be inferred by comparing and classifying the behaviours of the subjects; the analyses of specific texts and language combinations can help in extracting general principles about mental processing, and these principles can be used as hypotheses for investigating other texts and language combinations.

As an example of how CNA is carried out, let us look at how Hale and Campbell (2002) interpret the data from an experiment they conducted with Arabic and Spanish translators with the aim of investigating source text difficulty. The participants were asked to translate two passages from English. In these passages, Hale and Campbell identify the lexical items and syntactic structures for which the translators provided alternative renditions; from these they then proceed to infer choice networks. In one of the two pools of 20 translations, for instance, they found different renditions for the official term *Anglicare* (the name of an Australian church welfare agency). Based on each list of different renditions, Hale and Campbell infer the two choice networks reproduced below (Figures 2.1 and 2.2), from which it

emerges that the difficulty posed by the term can be described in terms of selection from a set of strategic options.

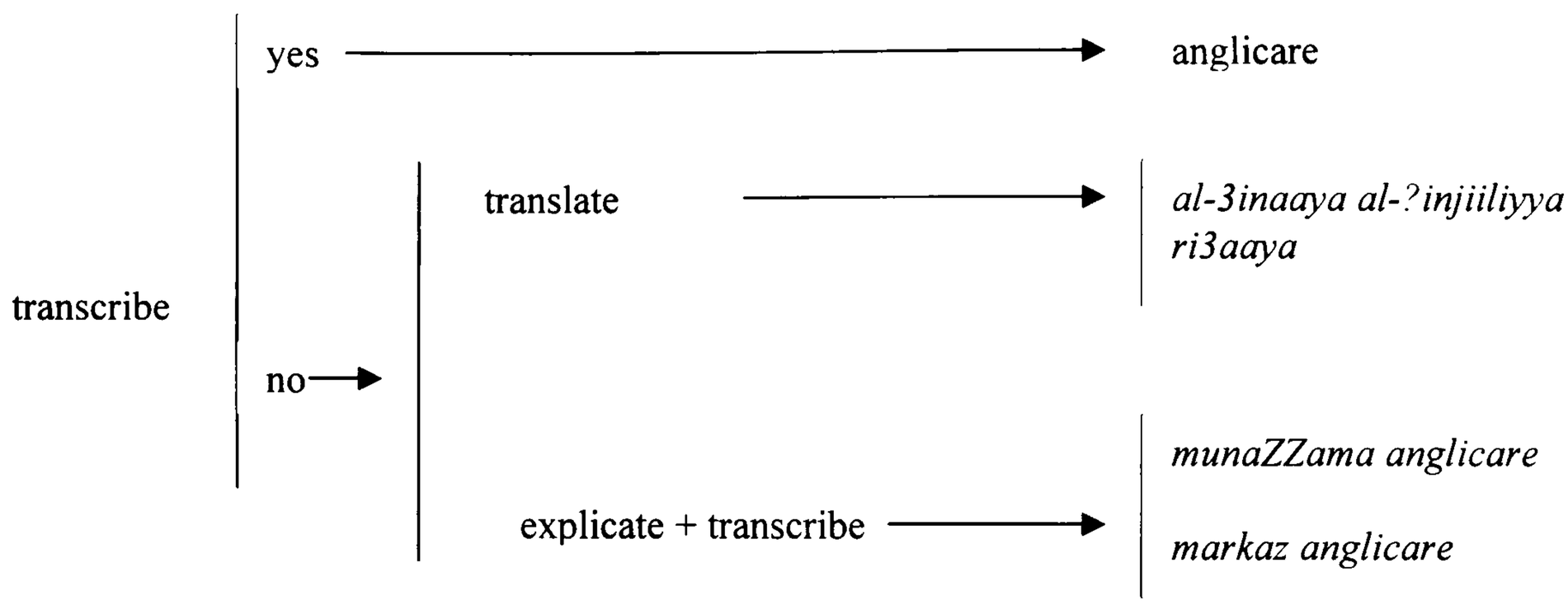

Figure 2.1 – Choice Network for the translation of *Anglicare* (English-Arabic; source: Hale and Campbell 2002: 20; spelling of Arabic translations as in source)

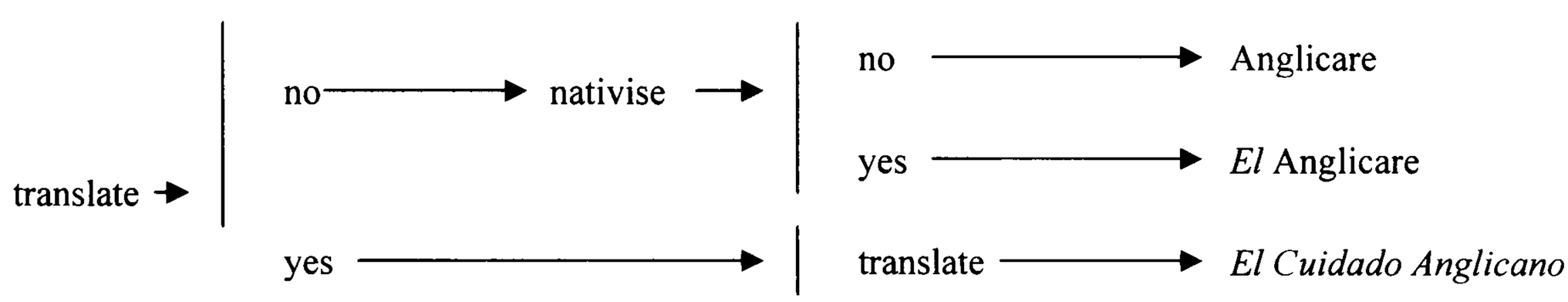

Figure 2.2 – Choice Network for the translation of *Anglicare* (English-Spanish; source: Hale and Campbell 2002: 20)

CNA is presented by Campbell as being theory-free: "[w]hat it generates are models, i.e. representations of complex systems" (2000: 38; see also 5.2). The researcher is free to make these models more or less complex and to incorporate theory in a network. When a theory is incorporated, CNA can indeed act as a test bench for it. An example would be the translation of complex noun phrases: analysing the network of choices proposed by a group of Spanish and Arabic translators for the phrase *methadone treatment*, Campbell (2000: 37) observes how the network provides confirmation for de Groot's (1997) hypothesis of transcoding, i.e. the replacement of an SL item with a TL item via the lexeme route rather than the conceptual route (in this particular case, all the Spanish translations had the cognate *tratamiento* for *treatment*, while the Arabic translators had chosen different TL items, suggesting recourse to the conceptual level).

As regards the possibility of using CNA to test theoretical assumptions, Campbell (2000: 39) is also careful to stress that the building of a choice network should respect three fundamental principles: 1) it must account for every piece of data in the sample "that is relevant to the domain of the theoretical framework of the investigator", which means incorporating in the analysis even those translation solutions that are not immediately recognizable as equivalents for a given ST (the network will force the researcher to put forward a hypothesis about their presence); 2) the choice network must be "linguistically plausible", i.e. it should not defy the notion of grammatical constituency; 3) it must be "parsimonious", which means that it should contain the minimum number of nodes and branches but remain plausible at the same time.

The main application of CNA on the part of Campbell has been the investigation of source text difficulty, presented in a series of papers in which the notion of difficulty is increasingly refined (see section 2.7 below for a discussion of the notion of difficulty in translation). Other possible applications suggested in Campbell (2000) are the investigation of expert-novice translator behaviour and the research on the translatability of specific language pairs (in itself an area that could be of some relevance to studies of text difficulty). Of direct relevance for the present study is also Campbell's (2000) hint at the possibility to cross-test hypotheses emerging from CNA with other empirical methods such as think-aloud protocols.

Cross-testing is indeed the approach that will be taken in the present study, where it is intended to compensate for possible limitations deriving from the fact that only a small sample of translators was observed. In particular, the reduced size of the sample could weaken the validity of the inferential steps taken in using product data (i.e. TL renditions) to discuss a processual aspect (i.e. mental processing by translators). In order to overcome this limitation, the analysis of choices made by translators will be performed not only on the TTs but also on the editing process that led to each TT (reconstructed on the basis of keystroke logs). Editing is, of course, primarily a reflection of processing aimed at reformulation in the TL, which means that an element of inference will still be present. It is felt, however, that editing can reveal with a reliable approximation whether translators really considered more than one choice as a translation for a given ST element.

2.3.1 Exploring text difficulty through CNA: the role of translation accuracy

CNA was initially developed by Campbell (1998) to study translation competence. The method was then used, with increasing fine-tuning, in a series of papers (Campbell 1999; Campbell & Hale 1999; Hale & Campbell 2002) that investigated "text difficulty" or, more explicitly, "what makes a text difficult to translate". The perspective from which difficulty is investigated in these papers is broader in scope than that adopted here. In particular, difficulty is linked to more general questions of translatability between pairs of languages, and the aim is that of establishing whether a given text is easier or harder to translate into different languages. Campbell's "difficulty" therefore has no clearly predefined orientation: it can emerge either from ST features or from factors that have to do with the TL. My own concern is with difficulty as a function of variation in style in the ST (understood as the preference for a more or less grammatically metaphorical expression form). The investigation of difficulty proposed here thus leans more clearly towards the SL. These differences notwithstanding, a few methodological considerations emerging in Campbell and Hale's papers are worth reporting in that they are of direct interest for the present study.

Originally, the relative difficulty of parts of the same texts was considered by Campbell (e.g. 1999) as a function of variation in the TT renditions of those parts by a group of translators. In other words, the more TL alternatives could be observed for the same ST element, the more difficult this element was taken to be. Based on these different alternatives, inferences were then made as to what led to one choice instead of the others.

In Hale and Campbell (2002) a new notion was incorporated that the authors felt had been neglected in the previous studies, i.e. the relationship between difficulty and accuracy. In particular, it was found that it was possible for different subjects to translate an element identically but inaccurately or, conversely, that subjects could translate an item differently but accurately. That of establishing the nature of the relationship between the number of alternative renditions and the accuracy of these renditions was thus felt to be a crucial issue. More specifically, Hale and Campbell asked whether difficulty should still equate only to a higher number of TL

realisations, regardless of accuracy, or to a higher number of *inaccurate* realisations only.[1]

To this question Hale and Campbell provide no clear-cut answer, since accuracy itself is by no means an easily defined concept in translation.[2] In some cases the criterion for accuracy is more easily identified. In others, accuracy has to be judged on the basis of "a complex of differently weighted criteria on the part of each individual who evaluates the translation, and of the readership" (Hale and Campbell 2002: 21). As an example of the former, the authors study the translation of complex noun phrases. These are elements where TL equivalents are constrained by grammar and semantics, so that "source text heads generally need to be translated as target text heads", making translations "very much amenable to judgements of accuracy" (Hale and Campbell 2002: 21), i.e. an evaluation in terms of 'right or wrong'. Not all scenarios, however, easily lend themselves to such a binary evaluative approach (cf., in this respect, Pym's (1992) distinction between binary and non-binary translation errors, of which more will be said in 2.6.3), which complicates matters at the time of judging how difficult a given ST element is. Hale and Campbell's (2002: 26) tentative conclusion is that when there is a high number of mostly accurate renditions, then these "only weakly" correspond to the level of difficulty; a high number of inaccurate renditions can more safely be associated with a high degree of translation difficulty.

2.4 Focus on the Source Text

Much has been written about the role of the ST in translation in general terms. In recent years, however, scholars have agreed (to a greater or lesser extent) on a broadly functionalist perspective. As a result, in descriptive translation studies the ST has almost been eclipsed from view. The attention of scholars and researchers has moved to the cognitive aspects of translating, the shifts and strategies occurring in

---

[1] Hale and Campbell (2002: 17) also noted that, in considering difficulty, a number of other, different factors should be taken into account: the individual's own awareness of an incorrect choice; the individual linguistic and stylistic competence; the individual's knowledge of the subject matter; the inherent ambiguity or lack of clarity of the text; the translatability of the ST at different levels (lexical, syntactic, semantic, pragmatic, discoursal).

[2] Sager's (1983: 21) remark that there are "no absolute standards in translation quality but only more or less appropriate translations for the purpose for which they are intended" can now be seen as commonplace in translation research.

translation (as product and/or as process), the universal features of translations and, more recently, the sociocultural aspects of translating (see, for instance, Pym *et al.* 2006). The majority of studies on difficulty in translation (however it is defined; see section 2.7 below) focus on what Toury (1995: 34) calls "transfer operations", i.e. the set of operations whereby a translation comes into being.

There is a substantial body of research on text difficulty *outside* translation. One strand of research is devoted to readability and, more generally, complexity (see, for instance, the various studies in Davison and Green 1988). Other studies have considered text difficulty in more narrowly defined areas, such as LSP or the language of science (Hayes 1992; see 2.5.3). Studies concerned with issues of language simplification (see, for instance, the research on controlled language or the Plain English Movement) also indirectly address the notion of text difficulty, as do efforts in areas such as technical writing (Lassen 2003). In general, the research on text complexity or difficulty, especially when applied to scientific discourse, is triggered by concerns over the increasing opacity of the texts written by specialists (cf. Halliday and Martin 1993).

In relation to translation, text difficulty can be characterised, in very broad terms, as "that which makes a text difficult to translate", a general formulation which of course says nothing about the *scope* of the concept (are we considering the text in relation to one TL, to a family of TLs or to any TL? Are we considering a specific translator or any possible translator?) nor its *orientation* (is the difficulty attached to comprehension or to reformulation?). A source text can be considered difficult from all these points of view. The present work aims at studying a particular feature of texts in a particular domain of communication for a particular language pair and assessing this feature as a specific source of difficulty for particular group of translators.

In empirical studies of translation, the ST is often taken as a given – a factor of the translation task whose influence on the way translators perform transcends the interest manifested by researchers, committed as they are to peering into the "black box" or to describing with increasing precision the procedural aspects of the translation task. When it is considered, references to it are only made in passing. Apart from a few studies associated with the work of Stuart Campbell discussed above (see Campbell 1999; Campbell & Hale 1999; Hale & Campbell 2002), no

systematic study exists of how the source text influences the way translators translate. Discussions of the source text and of the difficulties encountered by translators (especially trainees) are usually considered to fall within the remit of language teaching and language acquisition. Translators are often implicitly compared to ideal speakers of the two languages in question whose problem lies mainly in deciding on the best way to transfer meaning in accordance with the translation brief. Although much attention has been paid by researchers to a wide array of factors influencing the translation task, the interaction between text difficulty and translation accuracy has rarely been scrutinized (except, as noted above, by Campbell and associates; cf. also the study by Halskov Jensen discussed in 2.2.2.1).

In Tirkkonen-Condit (2005) mention is made of the fact that experts "are able to take the upper hand" in relation to the ST (cf. also Séguinot 1989: 25-30); indeed, such an approach "in non-fictional translation is often necessary, as source texts may be generated with little attention to the form of expression" (Tirkkonen-Condit (2005: 407) – a rare reference to ST formal features in terms of quality and to the way they can affect the translator's performance or decision-making.

It is interesting to note that: a) the few studies (apart from the ones by Campbell and colleagues mentioned above) which *have* focussed on the features of the source texts as one of the factors determining the translatability of a given text have all been carried out in relation to machine translation (MT; cf. Bernth and Gdanic 2001; Reuther 2003; Rojas and Aikawa 2006); and b) that these studies have frequently adopted an explicitly Hallidayan perspective in discussing ST features and the difficulties they pose for MT systems. In one such study (Møller 2003) grammatical metaphor is even taken as an 'umbrella' concept under which a whole host of lexico-grammatical phenomena is grouped which have proven to be sources of difficulty for MT systems.

The ST can be presented as one of the factors that constrain decision-making by the translators. In particular, by looking at what is not accurate in a TT (the product), and by studying data on the process in order to find what might have led to such inaccuracy, the hypothesis can be put forward that it is some specific features of the ST that constitute a source of difficulty (or a problem; cf. section 2.7) for the translator. This can be seen as part of "an attempt to gradually reconstruct both

translation decisions and the constraints under which they were made" (Toury 1995: 88) – the focus being on what negatively constrains these decisions and in particular on certain aspects of the ST.

The framework for the present study treats the ST much like an independent variable, i.e. one that – all other things being equal – is likely to affect the performance of translators (this last being the dependent variable). Methodologically, the study needs to 1) define, with reference to the language of science, the features of the ST that are assumed to be relevant as far as difficulty in translation is concerned, 2) define the notion of difficulty in translation, highlighting the relevant aspects for a study of the role of STs in determining difficulty, and 3) adopt a framework for the analysis of translation errors capable of shedding further light on the 'loci' of difficulty to be identified in an ST. The following sections review existing studies in relation to each of these three aspects, concentrating on the issues that are more directly relevant for the purposes of the investigation.

2.5 The Language of Science

Descriptions of the language of science usually fall within either of two main strands of research. On the one hand, scientific language is considered as a specific variety in the broad category of Language for Special (or Specific) Purposes (LSP). Research on LSP is essentially aimed at contrasting it to Language for General Purposes and identifying its characteristic features, both at a general level and in relation to each individual field of knowledge or activity. The other research tradition is the one concerned with the rhetorical aspects of the language of science and draws on research on genre and studies in the philosophy of science. For the present purposes, it seems useful to dwell on some general themes in both traditions and then move on to describe in more detail some studies that are more relevant in relation to the theoretical framework adopted here.

2.5.1 Scientific language as an LSP

Research on LSP in English is a vast area, with studies reflecting an extremely diverse range of interests determined both from the many different perspectives adopted by researchers and the specificity of each particular field whose language is

investigated. Science, with its individual sub-domains, is only one such field along with others as diverse as business, gastronomy or cinematography. At a very general level, LSP studies focus on the degree of overlap between LSP and LGP, the elements that characterise (each) LSP and the different levels of LSP communication. The range of methods and applications found in such studies is, again, enormous. Recent trends in methodology include corpus-based research, discourse-analytical methods and combinations of these two (see, for instance, Partington *et al.* 2004). Applications of LSP research include language teaching, translator training and, increasingly, natural language processing (e.g. terminology extraction and automatic summarisation).[3]

A classic reference work in LSP research in English is Sager, Dungworth and McDonald (1980), a descriptive overview of the English special languages of science and technology based on an essentially functionalist perspective. Sager *et al.* stress the difficulty of defining LSP satisfactorily in purely linguistic terms and claim that LSPs "are more readily recognised as pragmatic and extra-linguistic subdivisions of a language" (1980: 2).[4] The first part of their work is therefore devoted to an in-depth analysis of the social rules governing the use of LSP and the demands placed on LSP by the communicative needs of users. Two other sections of the study, however, appear to be more relevant for the present purposes. One is the authors' theory of reference (1980: 70-86); the other is their description of LSPs at the syntagmatic and lexical levels (1980: chap. 8 and 9).

Reference in LSP is seen by Sager *et al.* as "the mapping of a discrete, numerable repertory of symbols, the lexicon, onto a continuous knowledge space" (1980: 73); as a consequence, each lexical item has its image in a region of knowledge space whose bounds vary between individuals but "are restricted by the social norm" (1980: 73). The social norm is a "tacit agreement" (1980: 74) on the bounds of the region in knowledge space that a lexical item refers to. The individual is, however, free to interpret the social norm in view of the total structure of his/her

---

[3] The list of works to be cited for both the methodological and applicative aspects of LSP research would be endless. Here are a few representative works: Dudley-Evans and St John (1998) and Douglas (1999) for the general aspects of LSP in English; Pearson (1998) and Ahmad and Rogers (2001) for term extraction; Bowker and Pearson (2002) for using LSP corpora in translation work.

[4] Elsewhere in their study, Sager *et al.* (1980: 69) define special languages as "semi-autonomous complex semiotic systems based on and derived from general language; their use presupposes special education and is restricted to communication among specialists in the same or closely related fields"

own knowledge. This implies that a certain flexibility exists in the determination of bounds. More importantly for the present purposes,

> [t]his flexibility in the face of individual variation ha important consequences for communication since the participants involved may each use the same lexical item, fixed in form by the social norm, but with reference to subtly different regions of their own knowledge space, each of which individually represents only partially the reference sanctioned by the social norm (Sager et al. 1980: 74).

The particular position of the translator in the communicative process initiated by an LSP text immediately comes to mind here: not being a specialist (at least, not typically), his/her regions of the "knowledge space" represented by the text will frequently diverge from those of the author and the intended readers of the text. An analysis of the role played by the particular expression form (the lexico-grammar, in Hallidayan terms – see 2.5.4) of a text in determining this possible divergence is one of the aims of the present study.

As to the syntagmatic description of the LSP of science and technology, Sager *et al.* note how specialist writers "show a distinct preference for only a limited number of structures" (1980: 185). The set of preferred structures includes nominal groups, non-finite verb forms and finite verbs in the passive. Nominal groups are presented by Sager *et al.* as the most prominent among these structures: "the most important components of the vast majority of SE [Special English] sentences are conceptual units expressed in nominal groups" (1980: 219). Nominal groups are seen by Sager et al. (1980: 219) to fulfil a variety of functions: they are used 1) to describe machines or processes, 2) to expose an idea or theory in a "logical" fashion, 3) to present a "reasoned explanation" of natural phenomena and 4) to give an "objective evaluation" of experimental data (scholars from the 'rhetoric of science' research tradition would probably not fail to notice the somewhat positivist overtones of the authors' vocabulary here). Of these functions, those in (2) and (3) appear to be of particular interest here: M.A.K. Halliday's characterisation of "grammatical metaphor" (read: nominalisation) as the basic principle of scientific discourse (see 2.5.4 below) is an attempt at identifying one fundamental mechanism whereby scientists unfold their arguments and, in the long run, contribute to the creation of a vast taxonomy of 'nominalised' concepts, i.e. terms.

From the structural point of view, Sager *et al.* see nominal groups acting as the building blocks of sentences, mainly by virtue of their ability to be expanded through the insertion of different types of modifiers. It is not the presence of nominal groups *per se* that distinguish special languages from general English, "but the amount of modification normally employed by the authors of scientific and technical texts (Sager *et al.* 1980: 219). On this respect, a significant remark made by Sager *et al.* (1980: 219) is that

> [t]he possibilities of combining, extending and sequencing nominal groups available to the specialist writer are limited only by the ability of the intended recipient easily to comprehend the resultant text.

An observation that once again immediately brings to mind the particular position of the translator as a reader of a scientific text and, at a more general level, points to the question of the interaction between text difficulty and the (presupposed) knowledge of readers.

The unfamiliarity of readers with the forms used in an LSP text is, as Sager *et al.* (1980: 230) do not fail to notice, even more apparent at the lexical level: many of the lexical items in an LSP text "are likely to be either completely new to [a non-specialist reader, GP] or at least used with a meaning different from that with which he has previously been familiar". Sager *et al.* also note (1980: 230) how "the lexicon of special languages is their most obvious distinguishing characteristic" and "the most fully documented area of special language studies" – a remark which probably still holds true even after the twenty years that have passed since their study was published. Of the various aspects connected to the lexical level of LSPs treated by Sager *et al.* (1980: chap. 9), three in particular appear to be of relevance for the present purposes: the question of lexicalisation, the nature of what they refer to as the "content" and "expression forms" and the modes of designation, among which compounding plays a major role.

The question of lexicalisation is seen as "fundamental to the description of special languages " (Sager *et al.* 1980: 231). Although "terms cannot [...] be conceived of as a single fixed set of designations", the assignment of a particular grouping of lexical items to the category of free collocation or that of compounds has a very significant effect in special communication, as it can "considerably unbalance the information load of an utterance" (Sager *et al.* 1980: 231). Closely connected to

the question of lexicalisation is the identification of extended terminological units, which Sager *et al.* (1980: 233) see as causing particular difficulties in English as a consequence of the following phenomena: the tendency of English not to write compounds in one word and its inconsistent use of the hyphen; the fact that extended term units are often distinguished in opposition to other term units of different origin or nature; the fact that terminologisation makes a limited use of the linkages required in normal free collocation and "relies more directly on the users' knowing the appropriate concepts and thus being able to identify them" (Sager *et al.* 1980: 230). These difficulties are discussed by Sager *et al.* mainly from a lexicographical point of view. It would be easy to imagine, however, that similar difficulties in recognising extended terminological units might be experienced by another category of special readers of LSP texts, i.e. translators.

As regards the forms of special languages, Sager *et al.* follow a Wüsterian approach and distinguish between two planes, "content" and "expression". The content plane concerns the relationship between lexical items and concepts; at this level, the lexicon of special languages is seen by Sager et al. (1980: 242) as consisting of three major groups of words: 1) general language words used in all disciplines or in one discipline in particular; 2) general language words used with some restriction or modification of meaning; 3) terms specific to a discipline – these are newly-coined words or terminologised general language words. On the expression plane, the one having to do with the way concepts are designated, Sager *et al.* (1980: 243) notice how special languages "strive to systematise principles of designation" in order to achieve the general aims of transparency and consistency. As far as English is concerned, Sager *et al.* (1980: 244) stress its considerable potential for conversion, defined as "the ability to use one word category for another, to the extent that it is difficult to decide which category a word belongs to without reference to etymology".

One more difficulty for non-specialists in reading LSP texts in English (although Sager *et al.* themselves do not present it explicitly as such) is their heavy reliance on compounding as a mode of designation of specialised concepts. A compound is a "combination of two or more words into a new syntagmatic unit with a new meaning independent of the constituent parts"; it is formed on the principle that "one, usually the first, element determines the other" (Sager *et al.* 1980: 265).

Compounds make use of adjectives (as in *heat-resistant* and *flame proof*), verbs (formed by backformation or conversion; e.g. *outline* and *overcharge*). particles (e.g. *off-press*, *on-line*) and, most frequently, nouns (such as *speed ratio*, *heat transfer* and *temperature control*).[5] Compounding can be cumulative and combinable with affixation: the number of constituents of multiple compounds can increase to 5 or 6 elements, as in *self-aligning radial ball bearing*. In discussing compounds, Sager *et al.* seem to adopt an ambivalent attitude as to their potential for causing difficulties for readers, be they specialists or non-specialists. On the one hand, they attribute to compounds "a unique capacity to contribute to the building of terminological systems" (1980: 268) and generally minimize the potential ambiguities generated by compounding by stressing the role of textual reference.[6] On the other, they acknowledge that their own in-depth analysis of this mode of designation is meant to "assist understanding of the meaning of compounds" (Sager *et al.* 1980: 267).

2.5.2 The 'rhetoric of science'

Traditionally, studies of the language of science conducted from the point of view of linguistics were aimed at identifying features that were characteristic, and tended to see scientific language as an autonomous system. Scientific language was thus variously characterised as a sublanguage (Kittredge and Lehrberger 1982), a special register (Crystal and Davy 1969) or a language variety distinguished by particular lexical items (Savory 1965; Hogben 1969) or by syntactic forms (Huddleston 1971; Gopnick 1972). The specific features of scientific language were mainly discussed in lexical and morpho-syntactic terms: with reference to English, the study by Sager *et al.* (1980) reviewed in the previous section is a good example of a lucid and rigorous description falling within this line of investigation, although the authors themselves are repeatedly careful to stress that theirs is an analysis carried out from an essentially functionalist point of view and taking into great consideration the

[5] All examples of compounding are taken from Sager *et al.* (1980: 267-275).

[6] They also note how "the interpretation of compounds is assisted by the relatively high frequency of neoclassical or Romance words with clear word category indicators" (Sager *et al.* 1980: 267). It would be interesting to see whether this is still true today.

communicative needs of the users of special languages.[7] More recently, studies of the language of science have tried to look at it adopting the perspective of genre studies, the sociology and the philosophy of science or a combination of these.[8] The language scientists produce has been seen not only as a means of transmitting knowledge, but also as an instrument used for purposes that, at first sight, would seem to be distant from the basic aim of scientific writing, i.e. that of presenting experimental data in a purely objective, expository manner. Argumentation, in particular, has been identified as a key rhetorical feature of scientific writing across different disciplines and genres. Analysing both textual forms (in their various generic realizations) and linguistic structures at sentential and phrasal level, an attempt has been made at explaining how scientists use writing to convert "brute facts" into "meaningful statements", of whose truth readers must be "persuaded" (Gross 1990: 4).

A pioneering study among those devoted to the rhetoric of science is Bazerman (1988). Defining rhetoric as "the study of how people use language and other symbols to realize human goals and carry out human activities" (1988: 6), Bazerman sets out to investigate "how language realizes the work of science as a naturally situated, social semiotic system" (1988: 291). More specifically, Bazerman focuses his inquiry on the experimental article, in an attempt to show how "the experience is shaped by the arguments just as the arguments exploit the experience" (1988: 155). The experimental research article emerges form Bazerman's analysis as a dynamic and highly rhetorical product influenced by a number of heterogeneous forces: historical, empirical, social and cognitive. One section of the study (Bazerman 1988: chapt. 6) is devoted to experimental reports in physics, where some of Bazerman's findings appear to be of direct interest for the present work. A diachronic analysis of articles concerned with spectroscopic research published in the *Physical Review* from 1893 until 1980 shows that sentence length and syntactic complexity in these reports remained stable, which suggests that "neither changes in article length nor perceived changes in the 'difficulty' of reading can be attributed to

---

[7] Sager *et al.* (1980: 2) are explicit in rejecting a view of special languages as a separate system from general language.

[8] Bazerman (1988: 7) notes how the sociology of science can contribute to an understanding of how communication is organised in academic contexts and how texts fit in with the larger systems of disciplinary activity; the philosophy of science, on the other hand, is relevant for the study of scientific language "not for the ultimate questions of epistemology, but for the more modest ones of how people conceive [...] of disciplinary activity".

changes in sentence patterns or sentence style" (Bazerman 1988: 168). What did change in these articles was the increasing abstraction of sentence subjects and the increasing density and theoretical import of multiword noun phrases:

> [s]uch phrases are to be distinguished from ordinary adjective-noun clusters in that they modify not just by adding information, but by placing the object, event, or concept within a more specific framework of knowledge (Bazerman 1988: 171).[9]

This "ontological" role played by noun phrases in scientific language as described by Bazerman is extended by Gross (1990: 18) to syntax in general. Gross notes how in the traditional view of scientific language nouns stood for natural kinds and predicates for natural processes, thus making syntax a mere reflection of reality, of the structure of nature (1990: 18). According to Gross, however, such aspects of language as syntax, metaphor and analogy are central to scientific enterprise, and particularly to the way scientists persuade readers of the truth of their statements (cf. 1990: 18). Hence the extension to syntax in general of the ontological task Bazerman identified for noun phrases only. Before discussing in more detail another description of scientific language (M.A.K. Halliday's) stressing the joint effect of lexis and syntax, one more study will be reviewed which lends support to Bazerman's thesis regarding the increase in density of scientific English over the years.

2.5.3 D. P. Hayes' 'inaccessibility of science'

The increasing density of scientific texts also emerges in a paper published in *Nature* by Donald P. Hayes (1992). Trying to go beyond the "anecdotal evidence" regarding the opacity of scientific literature, Hayes proposes a method for measuring what he calls "text difficulty". This method he then applies to a corpus of texts comprising samples taken from science journals published in the 145 years prior to his study (the other genre represented in the corpus is that of science textbooks for college students).

Hayes' method measures text difficulty in terms of lexical choice, with reference to an estimated 600,000 word types assumed to represent the full English

[9] A remark reminiscent both of the delicate issue of lexicalisation in special languages as treated by Sager *et al.* (1980: 231; see 3.1) and of what M.A.K. Halliday calls the "systemic" (i.e. long-term) effect of nominalisation in scientific discourse (see 2.5.4).

lexicon. The measure is relative: the language of newspapers is taken as the standard for comparison and Hayes' method calculates the discrepancy between a specific text's pattern of word choice and the pattern observed in newspapers. The higher the score obtained for a text, the more difficult it is considered. Table 2.1 reproduces some of the results reported by Hayes.

Table 2.1 - Range of lexical difficulty in selected text categories (based on Hayes 1992: 739)

Nature (specific article from 1960 issue)	55.5
Science (abstracts, 1990)	44.8
Cell (articles, 1990)	38.0
Nature (research articles, 1990)	31.6
Science (research articles, 1990)	28.0
Physics Today (articles, 1990)	13.3
New Scientist (articles, 1986)	4.0
International English-language newspapers	0.0
Discover (popular science magazine, 1990)	-4.7
Fiction books, American	-19.3
Comic books, American and British	-26.8
Casual conversation	-41.1

Hayes' analysis is carried out both diachronically and synchronically. Diachronically, it shows how in the period 1950-1990 each science publication became increasingly difficult for non-specialists according to this measure. This increase is particularly evident in journals publishing reports of original research, such as *Nature*, *Science* and a series of basic science journals in fields such as biology, chemistry and geology. An upward trend, however, is observable even in journals such as *Scientific American*, whose readers were not traditionally characterised as trained scientists. Up to 1970, *Scientific American* was, in terms of difficulty, at or slightly below the level of a newspaper. After 1970, its language became more similar to that of academic science journals, as reflected in the increase in lexical difficulty measured by Hayes' method (Hayes notes how readers were quick to react to this increased opacity: when the difficulty of the average article approached a score of 15, the number of subscribers to the journal dropped by 125,000).

As acknowledged by Hayes himself, there are several contributory factors to this observed increase in the difficulty of science texts. One is the greater level of

detail with which texts are written. Another has to do with the dynamics of publishing: as publications compete for authors, readers and, at times, advertisers, they look for new "lexical niches" (Hayes 1992: 740), which results in an overall upward trend as far as the difficulty of texts is concerned. At the same time, "from the abruptness with which the changes in text difficulty occurred [in the professional science journals, GP] it seems that editorial policy may have something to do with it" (Hayes 1992: 740). In other words, choices regarding the selection of major articles and short reports, the intended readers of articles and (in general science publications) the disciplinary fields to be featured have resulted in increasingly inaccessible texts, both for non-specialists and for scientists from other disciplines.

Hayes' study is here quoted at length in that it supports the hypothesis forming the basis of the present study, i.e. that the style of scientific texts can be an independent or at least specific source of difficulty for readers (even, as we have seen, when they are trained scientists but come from a different field from that which a given text belongs to). Hayes' analysis, however, is centred on lexis alone and does not go into detail as regards other specific traits of the language of science which readers might find difficult or unfamiliar. A description of such traits is provided in a series of papers written by M.A.K. Halliday, which will be treated in the next section.

2.5.4 M. A. K. Halliday's description of scientific English

Most descriptions of scientific discourse centre on the role of lexis and treat terminology as a distinctive feature of the discourse. As was seen in the previous sections, the grammar of scientific discourse, on the other hand, was traditionally studied in terms of either a restricted selection of grammatical categories or a preference for some categories with respect to others – both perspectives frequently agreeing in considering nominalisation as the grammatical feature that distinguishes scientific texts from general language texts (see, in particular, the discussion of Sager *et al.* 1980 above). Other features include a more extensive use of the passive voice and the absence of certain verb tenses and moods. Few approaches, however, have tried to consider lexis and grammar as part of the same phenomenon. One such

approach is followed by Halliday in his analysis of the lexico-grammar of scientific discourse (see Halliday 1992, 1993a, 1993b).

According to Halliday, in the language of science information is organized following one basic principle which is particularly suited to the kind of argumentation that, historically, came to be accepted as scientific. This principle is "far from being an arbitrary or ritualistic feature" and is "an essential resource for constructing scientific discourse" (Halliday 1993a: 61). In elaborating on the principle, Halliday uses the term "grammatical metaphor", defined as the phenomenon whereby "the grammar shifts from a predominantly CLAUSAL to a predominantly NOMINAL mode of construal" (Halliday 1992: 30). Consider the following phrase (cfr. Halliday 1992: 31):

the orbital motion of an electron,

which is an implicit rewording of the verbal clause:

an electron moves in an orbit (sic).

Actual examples taken from scientific papers – also quoted in Halliday (1992: 25) – are:

the net effect of electron emission is the conversion of a neutron to a proton;
gene recombination results from some sort of physical exchange, or crossover, between chromatids.

Both sentences are an example of the syntactic pattern resulting from the shift from a clausal structure[10] to a highly nominalised structure. The underlying verbs (*to emit* and *to convert* in the first sentence, *to recombine* and *to exchange* in the second) are converted into complex nominal groups, thus realizing the pattern on which the "prototypical syndrome of features" identified by Halliday (1993a: 54) in scientific English typically converges:

nominal group + verbal group + nominal group.

[10] The underlying clausal structure of the two sentences could read as follows:
neutrons are converted into protons because the neutron emits an electron;
genes are recombined because chromatids are exchanged in some way.

There are two main reasons why this pattern, where nominalisation is largely the result of what Halliday calls "grammatical metaphor", is found particularly in scientific discourse (cf. Halliday 1992: 30-31):

1) it helps in unfolding an argument and contributes to the ongoing reasoning in the text; Halliday calls this the "instantial effect" of nominalisation: the nominalised form is here being employed as a "temporary construct set up to meet the needs of the discourse" (Halliday and Martin 1993: 14);
2) it contributes to the framework of the technical concepts of a given discipline; these are related to one another taxonomically and condense in themselves large amounts of accumulated knowledge – the effect here is "systemic".

Halliday (1992: 31) is careful to stress that the effect brought about by nominalisation is really of the same nature whether it is instantial or systemic; what varies, in other words, is its range: immediate in (1), long-term in (2). Illustrating these effects properly would of course require an entire text or preferably a text sequence reflecting the development of a scientific theory.

Halliday's study of nominalisation goes beyond the establishment of its significance in quantitative terms and takes into account its functional/pragmatic significance both intra- and intertextually. Nominalisation is here linked to the emergence and establishment of a particular kind of argumentation, for which it seems to be a natural choice. Halliday, in other words, shows the reasons why nominalised clauses are particularly suited to the needs of scientific discourse. Also, he stresses that in considering scientific discourse lexis and grammar cannot be considered separately: grammar has as much meaning potential as lexis. He therefore prefers to use the term lexicogrammar, as technical vocabulary and grammar are "different aspects of a single semiotic process" (Halliday and Martin 1993: 8).

Grammatical metaphor is one of the seven features of scientific English identified by Halliday (1993b: 69-85), the other six being: interlocking definitions, technical taxonomies, special expressions, lexical density,[11] syntactic ambiguity and semantic discontinuity. Of these, technical taxonomies, lexical density and syntactic ambiguity are very closely related to grammatical metaphor, of which they can be seen as by-products (cf. Halliday 1993b: 79): technical taxonomies are made up of

[11] In Halliday's definition, lexical density is "the number of lexical words per clause" (Halliday & Martin 1993: 76). Other definitions of lexical density are based on different criteria.

terms, and many terms are really the result of a process of grammatical metaphorisation; lexical density (i.e. the number of lexical words per clause) is higher where there is a great number of nouns; finally, syntactic ambiguity often arises in interpreting both the verbal group joining two nominal groups and the nominal groups themselves.

Halliday applies this conceptual framework to the analysis of English texts (extracts from Newton's *Opticks*, from Darwin's *Origin of the Species* and from a modern-day scientific journal). He repeatedly makes reference to other languages, remarking that their pattern of evolution might be very similar to that undergone by English. As regards the difficulty in *reading* scientific texts, Halliday (1993: 71) is careful to stress that this difficulty "lies more with the grammar than with the vocabulary", that "it is the total effect of the wording – words and structures – that the reader is responding to", and that "technical terms are part of this overall effect".

In Halliday's and in the other descriptions briefly reviewed above, nominalisation emerges as one key feature of the language of science in all its different registers and text types. As Halliday clearly shows, this heavy reliance on nominalised groups is by no means an arbitrary feature but an essential resource for the construction of scientific discourse. While "to write science is commonly thought not to write at all, just simply to record the natural facts" (Bazerman 1988: 14), the inquiries into scientific language briefly illustrated here show how scientific communication is the result of "massive linguistic work" (Bazerman 1988: 14) at all levels: phrasal, sentential and textual.

The basic assumption of the present thesis is indeed that the style of scientific writing can be a specific source of difficulty for translators, with style defined, in Hallidayan terms, as the preference for certain lexico-grammatical constructs (mostly those resulting from nominalisation) over others.[12] At the level of syntax alone, scientific language may probably be seen to have undergone a process of simplification, with writers showing a distinct preference for a limited number of constructs, the most typical of which is the "noun group + verb group + noun group"

[12] The view that specialist texts in general sometimes display an inherent complexity which is, at least partly, the result of (poor) style is also assumed by Göpferich (1998:1-2), who sees "the increasing specialisation in all disciplines" as causing "obstructions in the flow of information between specialists and non-specialists". In an intralingual perspective, Göpferich acknowledges the important role of "mediators" such as technical communicators in overcoming these obstructions.

pattern discussed above. This syntactic simplification, however, seems to have been counterbalanced by the considerable strain placed on nominal groups. If today "much of the vocabulary of the various subject fields is incomprehensible, not only to the layman, but also to specialists in other fields" (Sager *et al.* 1980: 230), the reason probably lies not only in the extremely high degree of specialisation reached in each field, but also in the highly compacted, highly nominalised forms this vocabulary employs.

2.6 Translation error analysis – a brief overview of past research[13]

The concept of "translation error" lies in a much larger field of translation research, which is that of translation evaluation, and is directly linked to at least two other broad areas of research: translation competence (which can be seen as the object of evaluation) and translation training. Within translation evaluation, the discussion of errors is part of the investigation of evaluation criteria and evaluation procedures and instruments. A clear definition of translation error (and of translation problem) seems to be crucial for the establishment of evaluation criteria. At the same time, the particular context in which the evaluation takes place has to be taken into account. Two broader contexts may be identified: translation teaching and professional translation. The definition of "error" will change according to the particular context where error analysis is carried out, although elements of the definition are bound to overlap to a greater or lesser extent.

2.6.1 Traditional error categories in translation evaluation

Traditionally, typologies of translation errors were based on categories such as wrong sense, opposite sense and overtranslation, i.e. errors related to *transfer*, whilst grammar, spelling and other errors related to *re-expression* in the TL. As acknowledged by Waddington (1999: 37) these categories are today still used by many translation teachers, but they began to be criticised by researchers as far back as the 1970s. Their continuing success is perhaps due to three main factors: 1) the

[13] Extensive critical overviews of existing studies on translation errors are presented in Waddington (1999) and Hurtado Albir (2001), two works published in Spain (Spain, Canada and Germany are the countries where most of the research on translation errors has been carried out to date).

sheer force of habit, 2) the fact that they are more or less overtly based on the principles of the *stylistique comparée* (cf. Vinay and Darbelnet 1958), an approach which was very influential in translation research starting from the 1950s, and 3) their simplicity of use. Criticisms moved to these traditional categories of translation evaluation are mainly concerned with their rigidity and their insistence on translation as having an essentially linguistic, rather than communicative, dimension (cf. Gouadec 1989: 36; Waddington 1999: 64-65). The rigidity derives from their tendency to establish universally valid parameters which do not take into account the uniqueness of each ST, if not of any act of transfer into the TL. On the other hand, their lack of a communicative perspective makes these categories unable to evaluate a given TT element in terms of its appropriateness to the textual function or any other pragmatic consideration entering the transfer process or explicitly linked to the translation brief.

2.6.2 Functionalist approaches to error analysis

The first explicit attempt to account for a functionalist perspective in the evaluation of translation errors is to be found in House (1977), where a distinction is proposed between *covert errors* (i.e. those solutions which fail to establish a functional equivalence between the ST and TT) and *overt errors* (i.e. mismatches of the denotative meanings of ST and TT elements). This identification of two broad categories of errors, with one category usually accounting for the translation errors "proper" (i.e. errors due to a lack of translation competence) and the other concerned with errors resulting from a lack of linguistic competence, is to be found in most of the subsequent studies dedicated to either translation errors or translation evaluation. Even when a functional perspective is not explicitly mentioned or adhered to, such as in Gouadec (1989), the distinction nevertheless remains.[14]

According to Kupsch-Losereit (1985: 170), the definition of translation error cannot be derived from a structural description of the linguistic system but should result from the consideration of the specific equivalent text function which the ST

[14] Gouadec (1989) follows a previous paper on translation errors (Gouadec 1981), which identified as many as 675 types of errors (see Waddington (1999: 62-63) for an evaluation of this study). In the later paper, Gouadec makes a distinction between "absolute" and "relative" errors. "Absolute" errors result from a violation of the cultural or linguistic norms or from a violation of usage rules, while "relative" errors are solutions that do not conform to the requirements of a given translation project.

and the TT realise in a given communicative situation. Along the same lines, Nord (1996) operates a clear distinction between errors of re-expression and translation errors, the identification of the latter being based on pragmatic factors. Translation errors are thus discussed in relation to the specific functional aspects emerging from the translation brief. As a consequence, such a concept as the "quality" of a translation becomes a relative one; its measurement is not based on absolute standards but is carried out in terms of "appropriateness" with respect to the purpose of a translation (cf. also Pym 1992;[15] Hatim and Mason 1997). By the same token, the impact of errors is seen as varying according to the relative importance of the erroneous element within the text as a whole (cf. Sager 1983: 127; Kussmaul 1995a: 139-141).

As suggested by this brief overview, research on translation errors has covered much new ground since House's explicit introduction of a functional perspective in the 1970s. However, while there is no aspect of translation (be it teaching, research or professional practice) which does not refer, "implicitly or explicitly" (Gouadec 1989: 35), to the notion of error, the number of studies dedicated to translation errors remains small and in the past few years other authors (cf., for instance, Hurtado Albir 2001; Martínez Melis and Hurtado Albir 2001) have lamented the lack of empirical research on a typology of errors, on the seriousness and impact of errors and on the causes of errors. In particular, almost no study focuses on the features of the ST as an explanation for some translation errors. In other words, the problem of difficulty in translation is always discussed in terms of poor performance on the part of the translator and issues regarding the features which make some texts more difficult to translate than others have rarely been tackled – with the usual exception of the work carried out by Stuart Campbell (1999; see also Campbell and Hale 1999; Hale and Campbell 2002).

2.6.3 Translation as selection – Pym's categorisation of errors

As noted above, the study of errors is closely connected to the research on translation competence. Translators seem to be able to act as such thanks to a specific kind of competence. While there is little consensus on how such competence can be defined,

[15] See below (4.3) for a detailed discussion of Pym's scheme for translation error analysis.

almost all researchers agree that there is something that distinguishes translators from other language users who deal with more than one language, e.g. bilingual individuals.

Following Pym (2002), in the models of translation competence proposed so far four tendencies can be identified:

1) competence as summation of linguistic competencies (SL competence + TL competence)

2) competence as 'no competence' – translation deals with the actual use of languages, therefore it would be paradoxical, in Chomskyan terms, to describe as 'competence' something which is really 'performance' (the term is thus replaced with 'proficiency', 'strategies', 'expertise', etc.)

3) competence as multicomponential, i.e. made up of several skills (but lists of skills can be expandend *ad infinitum*)

4) competence as 'supercompetence', i.e. something which defines translating and nothing but translating.

Pym himself opts for a "minimalist" notion of competence, seen as "the ability to generate a series of more than one viable target text (TT_1, TT_2 ... TT_n) for a pertinent source text (ST)" plus "the ability to select only one viable TT from this series, quickly and with justified confidence" (2002). The question, in other words, is the following: can the acquisition of competence essentially be defined as the ability to break free from a "word-for-word" approach and start thinking in selectional terms?[16] Or to put it in Pym's terms, is the acquisition of competence a revelation that "translation is not simply a matter of literalist fidelity" (1992: 287)?

Closely linked to this definition of competence is Pym's scheme for the analysis of translation errors. If we accept that being a competent translator means being able to select from a series of alternatives, we are also assuming that a translation error must, by definition, be *non-binary*. Whereas a binary error opposes a wrong answer to the right answer (thus being identifiable as a language error), non-binarism "requires that the target text actually selected be opposed to at least one further target $text_2$ which could also have been selected, and then to possible wrong

[16] This, of course, does not mean that the 'word-for-word' approach will *never* give perfectly acceptable results.

answers" (Pym 1992: 282). In other words, non-binary errors are those that provoke in the translation teacher reactions such as "It's correct, but...".

If we accept this definition, then the question arises as to how trainee translators can best acquire the skills necessary to select the most appropriate alternative in the TL. And whereas in general-language texts specialist knowledge rarely enters the picture, in LSP texts specialisation is a *datum* and plays a primary role. In other words, specialist knowledge is one of the factors, sometimes the predominant one, which help the translator establish what the preferred TT alternative for a given ST element must be.[17] Moreover, where specialist knowledge (and specialist terminology) enters the picture, it is possible that binarism comes to the fore again. There often is one and one equivalent only for a SL term, and the student who does not select exactly that equivalent in the TL obviously incurs in a binary error (such a statement, of course, is an oversimplification, as it assumes that the ST and the TT always share features such as text function and text addressee).

From the definition of binary and non-binary errors given above, it would seem that the former should come within the remit of language teaching, leaving non-binary errors only to translation teaching. However, Pym himself sees such a distinction as questionable. More specifically, in the translation of LSP texts, the (re)emergence of binary errors may be attributed to failure in comprehending the ST or a lack of terminological/phraseological accuracy in the TT. In turn, the failure in comprehending the ST may derive from a lack of subject-matter knowledge or from an inability to extract meaning from a lexico-grammatically complex segment of text. If we go back to Chesterman's question ("So what do we mean by *better* and why are some options better?"), we can't help noting that in the case of LSP translation the answer also depends on the degree of subject-matter knowledge displayed by students (and teachers), so that the more familiar they are with the domain and its terminology, the better they understand the source text and the better equipped they are to provide a satisfactory TL equivalent.[18] As we shall see later, this should not rule out the possibility that some (purely) linguistic or textual features

[17] Cf. Chesterman (2000: 85): "Different options are not just alternatives, but usually some seem better than others. So what do we mean by *better* and why are some options better?"

[18] Add to this the difficulty deriving from the fact that "[s]pecialist terms are used slightly differently in different genres of specialist texts" (Ahmad 1999: 148).

may influence the translator's performance. Think, for instance, of aspects such as cohesion or style and the different ways these are handled across languages. Again, however, students may be faced with textual features that are typical of a given specialist domain, i.e. these features may be part of the background knowledge that a domain expert puts into the text more or less consciously.

Pym's distinction between two broad classes of errors will be taken as one of the bases for analysing the translations produced by the participants in the present study. More specifically, errors will be taken as one of the indicators of difficulty (or as problems that the translator was not able to solve) and will be analysed so as to identify those that are more clearly oriented towards the ST.

The two notions of 'problem' and 'difficulty' are compared and contrasted in the next section.

2.7 Problems and difficulties in translation: one and the same notion?

Empirical studies of translations, both product-and process-oriented, are essentially centred around two sets of questions. On the one hand, they try to describe what goes on in the translator's mind as he or she is performing the translation task. In doing so, these studies make hypotheses on those elements pertaining to either the text or the context which lead translators to go beyond automatic solutions or "unmarked processing" in Jääskeläinen's [1990: 173; 1993: 102] terms ("direct mappings" for Olohan [2000: 597]) and necessitate a problem-solving or decision-making approach. Other, perhaps more traditional studies, are centred on translation as product and essentially aim at either a critique of the target-language text or a description of the regularities observed in translated texts (what has come to be known as the "norms" governing the activity of translation; cf. Toury 1995). Both approaches are concerned with the identification of the strategies deployed by translators, either to achieve certain goals or in response to what they, or the researchers, perceive to be the problems found in the source texts or related to any other aspect of a translation task.

Although a strategy can be seen as something more than a problem-solving mechanism (Jääskeläinen's [1993: 116], for instance, adopts a broader notion, considering it as a plan to achieve something), the notion remains useful as a link between errors and problems themselves. Errors can first of all be seen as problems

which the translator was not able to solve (which of course still leaves open the question of how to define "problem"). Indeed, problem-solving abilities are taken to be at the heart of translation competence by all of the various research groups currently studying it in different parts of the world.[19] The notion of problem, in other words, emerges as a central issue in practically any kind of study concerning translation and based on data obtained from either the process or the product.

Problems (and the way they are solved) are also one of the crucial aspects taken into consideration when evaluating translations. Given its centrality, it is perhaps not surprising to find that many different definitions have been proposed for the term "translation problem", although not always explicitly and often in studies devoted to other aspects of translation. The lack of an uncontroversial definition and of a convincing categorization of problems has been noted by more than one author (e.g. Wilss 1996: 47; Hurtado Albir 2001: 280). Others (see for example Lörscher 1991: 92; Toury 2002) have stressed how the speculative approach favoured by many researchers in the past prevented a rigorous treatment of the notion. The uncertainty of researchers as to what translation problems are and how they should be categorized is reflected in the terminology: some researchers seem to use the two terms "problems" and "difficulties" in much the same sense; others distinguish between the two, as we shall see below. The following paragraphs offer a brief review of the available definitions, focussing on those which are more directly linked to the purposes of the present study.

Traditionally, problems have been seen as linguistic (lexical, syntactic or stylistic) discrepancies between the SL and the TL texts (see Vinay and Darbelnet 1958; Mounin 1963; Catford 1965). As happened with other, related notions, however (e.g. translation errors; see 2.6), such a view of problems, based as it was on the principles of the *stylistique comparée* (see especially Vinay and Darbelnet 1958), failed to take into account the communicative dimension of translation and the need to evaluate a problematic element in terms of the textual function and the other pragmatic considerations entering the transfer process or explicitly linked to the translation brief.

---

[19] Groups such as PACTE (Process in the Acquisition of Translation Competence) in Spain, CORPRAT (Corpus on Process for the Analysis of Translation) in Brazil, PRONIT (Process and Product in Translation Investigation) again in Brazil, TRAP (Translation Process) in Denmark and

Among less recent studies, Wilss (1982: 158-175) devotes a whole chapter to what he refers to as "Translation Difficulties" (TDs). Within an essentially didactic perspective, TDs are considered by Wilss as one of the four "focal points" of applied translation research, the other three being translation teaching, error analysis and translation criticism. In particular, Wilss assigns TDs, together with translation teaching, to the "prospective" pole of what he calls the "science of translation", the one focussing on *transfer*. Error analysis and translation criticism, on the other hand, are said to belong to the "retrospective" pole and their emphasis is on the transfer *result* (what researchers would today call translation as "product"). All four domains are subsumed "under a global applicative concept whose aim it is to improve the prospective and retrospective transfer competence of the translator-student and thus to raise the level of efficiency of practical translation work" (Wilss 1982: 159) – an explicit declaration of Wilss' ultimately didactic perspective.

Translation Difficulties are defined by Wilss (1982: 164) as occurring "whenever a lexical or syntactic one-to-one correspondence between SLT [source-language text] and TLT [target-language text] cannot be practised, because literal translation would inevitably entail a negative transfer". By giving prominence to transfer, the definition explicitly rejects a consideration of difficulties more closely related to the ST, whether as a result of problematic reception of the text or as a consequence of its intrinsic semantic or stylistic complexity. To be sure, Wilss is fully aware of what he calls "text-receptive difficulties" but prefers not to include them in its discussion in light of what he views as their high degree of subjectivity: insufficient SL text understanding, he says, is "in principle subject-bound (translator-bound)" and any attempt at systematizing receptive text difficulties will result in failure (Wilss 1982: 164). He prefers, therefore, to concentrate on the "results of transfer-preceding SLT analysis" following a "person-unrelated" approach (Wilss 1982: 164). To counter Wilss' scepticism on the role of the ST and ST reception in determining translation difficulty, it could be argued that data-collection methods established after his study appear to have put researchers in a better position to observe how translators react to source texts. Think-aloud protocols or keystroke logs, for instance, even with all their limitations, seem to enable researchers to make

---

EXPERTISE (Expert Probing through Empirical Research on Translation Processes), an international network, are all investigating translation competence focussing on problem solving.

more than educated guesses at what a translator finds difficult in a source text and why. With an appropriate design, studies based on such data can 1) aim at providing a convincing and more complete picture of what makes a text difficult to translate and 2) allow insights into differences of ST interpretation.

Wilss' definition of translation difficulty, based as it is on the transfer aspect, and more specifically on the non-isomorphic nature of languages and texts and the subsequent necessity to concentrate on the cases in which translating means essentially to provide an alternative for a literal translation felt as unsatisfactory, has 'set the scene' for many subsequent definitions of translation problem.

In studies of translation as process, strategies are a series of ordered behaviours (cf. Krings 1986: 175; Dancette 1997: 89), and problems are seen as those aspects of the SL texts that translators tackle by adopting a certain strategy. Lörscher (1991: 79-81) sees a problem as occurring "when a subject realizes that, at a given point in time, s/he is unable to transfer or to transfer adequately a source-language text segment into the target language" (1991: 80) – a definition given from the perspective of the subject rather than that of the analyst. For Lörscher, who draws on studies in cognitive psychology (e.g. Dörner 1976) and on a previous study by Krings (1986), the notion of problem has a psychological validity and can be related to the various aspects and components of the translation process. An interesting observation about problems made by Lörscher (1991: 85) is that they do not necessarily surface in the products of translation: "problems in the reception of the source-language text need not necessarily cause problems in the production of the target-language text". By the same token, "subjects could have latent problems in SL text reception which would become manifest if the task was to understand the SL text, but which remain latent in translation and do not cause problems in TL text production either" (Lörscher 1991: 85). It seems reasonable to assume that at least some of these latent problems are linked to aspects of the source text (possibly, as regards scientific texts, those discussed in section 2.5), and that their identification requires an investigation of both the TT (the product) and other kinds of data pertaining to the translation task (i.e. data from the process).

In a small-scale study recently carried out by Kussmaul (2005), translation problems are not defined explicitly but indirectly: "unproblematic cases" are those where "direct equivalents exist and where routine processes are involved" (2005:

382). This would seem to imply a dynamic notion of "problem", as problematic elements probably differ according to the level of professionalism. In other words. certain aspects of the translation task that constitute problems for novices would probably not be problematic for professionals. Kussmaul (1995: 378) himself suggests that "[p]roblem solving may even, when practised frequently, in the end become a routine". As noted by Kussmaul, this still awaits investigation. For the purposes of the present study, it would be interesting to see whether this 'routinization' of the task involves formal aspects of the source texts – in other words, whether certain formal properties of the source texts remain problematic for any translator or can be seen to be gradually integrated in the procedural knowledge, or more generally in the competence, of translators. Data from TAPs and loggings of the translator's writing process can shed some light on both latent problems and on aspects likely to become routine as the level of professionalism increases. Even if they do not emerge in TL text production, their consideration and the identification of their exact nature could nevertheless help in establishing different types of source texts and in grading each type for difficulty (something which may be needed for pedagogical or professional accreditation purposes).

In studies of translation as product, problems have often been seen as the other side of the coin of a much more studied aspect, i.e. error. There is, furthermore, a tendency in some authors to further distinguish between "problems" and "difficulties". From an essentially pedagogical viewpoint, Nord (1991; 1997), for instance, sees problems as objectively identified phenomena of a textual, pragmatic, cultural or linguistic nature. Difficulties, on the other hand, are a subjective phenomenon depending on the individual translator (or trainee) and arising because of "deficient linguistic, cultural or translational competence" or because of a lack of "appropriate documentation" (Nord 1997: 64). Nord (1991: 152-155) distinguishes between four types of difficulties: 1) those depending on the degree of comprehensibility of the source text; 2) those depending on the translator, 3) those related to the nature of the translation task; 4) those related to the specificity of the subject matter.

Campbell and Hale (1999; see also Campbell 1999) only talk about "difficulties", which they define from a cognitive standpoint, i.e. as items in a text "that require more processing effort to resolve target language choices". The aim in

Campbell and Hale's research is that of exploring "how universal translation difficulties might be" (universality here an echo of Nord's "objectivity". which however was assigned to "problems"). Comprehension and production are two different "loci of difficulty", the former being "fairly universal", the latter giving rise to different levels of difficulty depending on the lexis and grammar of the target language (Campbell and Hale 1999). Campbell and Hale concentrate on production difficulties, which they describe and analyse via their "Choice Network" model, i.e. "a psycholinguistic model of the options facing the translator, based on the evidence of the target texts of groups of subjects translating the same texts" (Campbell and Hale 1999). The areas of difficulty thus identified are as follows: (1) words low in propositional content, (2) complex noun phrases, (3) abstractness, (4) official terms, (5) passive verbs. While the focus is on the production phase, it is interesting to note that most of these areas are closely related to lexicogrammatical characteristics of the ST. Campbell and Hale themselves discuss these features in essentially Hallidayan terms: words low in prepositional content, for instance, are seen as an example of "grammatical metaphor", while the difficulty in rendering official terms is seen as a consequence of the dense semantic packing of such terms. Therefore, although their focus is explicitly on the production phase of the translation process, Campbell and Hale are seen to fall back on a description of ST features in order to account for the difficulties emerged in the process.

The distinction between problem and difficulty is far from being universally acknowledged, and many authors can be seen to use the two terms interchangeably.[20] Upon closer scrutiny, the distinction as is proposed by Nord may even look questionable. It is true that in some cases an element of the SL seems to be objectively problematic, i.e. likely to be a source of difficulty for translators in whatever type of text the element appears and for whatever reason the text is translated. So-called *culture-bound* terms or phrases seem to be a case in point.[21] However, in many other cases the distinction fails to take into account such crucial

[20] Cf. also Lörscher (1991: 92), who notes that the notion of "translation problem" had hardly been paid any attention by translation theory because of the (then) speculative and non-empirical orientation of most studies, and that available taxonomies of "translation difficulties" were only based on unsystematic observation.

[21] Cf. however Pym (2004) for the discussion of culture-bound elements which acquire a relative importance in the overall economy of a text and can be solved exerting more or less effort. Are these elements problematic or difficult? (See below for a discussion of Pym 2004.)

aspects as the level of competence of translators. Translator-specific "difficulties" are taken by Nord (1991: 153) to exist even for experienced translators, in which case it is hard to see why they should be any different from "objective" problems – if not precisely for the level of competence of the translators the researcher is observing (which in turn is not easy at all to define with absolute objectivity). Hurtado Albir (2001: 287) agrees with Nord and sees the problem/difficulty distinction as worth taking into account when studying problems. In criticising Bell (1998: 188), who identifies problems with cases of non-automatic transfer, Hurtado Albir (2001: 287; my translation) states that "an automatic solution for a[n ST] element cannot be taken as tantamount to the non-existence of a translation problem" and that if that element "was not a problem for a certain translator, it could nonetheless be seen as such for another translator", i.e. it could be translator-specific or, in Nord's terms, it could be seen as a "difficulty". A few lines below, Hurtado Albir herself acknowledges that the borderline between problems and difficulties is not clear-cut, that empirical studies are needed to establish a more convincing differentiation between the two and that in defining both the notion of problem and the way it differs from that of difficulty, the level of competence must be taken into account: "different [kinds of] translation problem can be seen to emerge at each stage of the acquisition of translation competence" (Hurtado Albir 2001: 287; cf. also Presas 1997).

An attempt to go beyond a definition of translation problem based solely or mainly on linguistic grounds is to be found in Pym (2004). Pym's approach is based on the notion of *risk*, defined as "the possibility of not fulfilling the translation's purpose". Translation problems, then, are not only SL elements for which a selection between two ore more TL options is possible, a definition proposed in a previous paper (Pym 2002). Rather, a translation problem can be described in terms of the risk attached to it; more specifically, the risk that the solution proposed for it does not fulfil the purpose the translation is intended to have. This entails that "a linguistic rich point is not necessarily a high-risk problem" and that in order to efficiently solve certain problems the translator has "to go in search of information beyond the text, and certainly beyond comparative linguistics" (Pym 2004).[22] By the same token, it

[22] Notice, in passing, how close Pym's revised characterization of the notion of problem is to what Nord (1988/1991: 151) calls pragmatic "difficulties" (see above), i.e. those related with the nature of the translation task.

follows that there is no such thing as the definitive solution to a translation problem – there are different solutions with different degrees of risk attached to them. Also, some solutions involve more effort than others; "the goodness or badness of a solution is then a question of the way risk correlates with effort" (Pym 2004), which in turn has to do with the type of strategy the translator resorts to in order to solve a problem. This redefinition of the notion or problem seems to be particularly interesting in light of the particular type of texts the present study is based on. In discussing how errors could be analysed (2.6.3.), it has already been noted how, with specialized texts, a simple distinction between binary and non-binary errors is not always capable of accounting for the role of specialized knowledge. Pym (2004) himself seems to hint at this when warning that a failure in solving certain "binary" problems, e.g. dosage quantities on a medical label, can be very high-risk (and therefore translationally important in terms of his revised definition), whereas failure on "non-binary" solutions (i.e. the properly translational solutions according to the old definition) could have low-risk consequences.

Another attempt at discussing the concept of problem in the "internal terms" of Translation Studies is to be found in Toury (2002). Acknowledging the centrality of the concept, Toury also notes how the term "problem" is really used in Translation Studies with three different senses, serving three different kinds of expert discourse on translation. The first sense (which Toury refers to as "PROBLEM$_1$") is located in discussions of the source text and involves issues of translatability rather than actual translation. PROBLEM$_1$, in other words, is a prospective notion that refers to an SL element (e.g. a metaphor) for which we investigate the possibility of establishing optimal correspondence with an appropriate TL element. This investigation is not linked to any actual translation act; rather, the nature of the translation act is only speculated on in idealized terms and a translator "persona" is postulated who "is often ascribed almost mythical qualities: full mastery of the languages and cultures involved in the act, unlimited resources, unlimited memory, an ideal capacity to analyze and interpret texts, and the like" (Toury 2002: 62). The second sense identified by Toury ("PROBLEM$_2$") refers to actual instances of translation: it is associated with individual translation acts situated in a particular time and space. This meaning features in discourses on translation which are retrospective and see translated texts as a reservoir of realized solutions. PROBLEMS$_2$ are therefore

"reconstructed entities" (Toury 2002: 64) arrived at through an examination of coupled pairs of replacing + replaced segments established during the analysis of the SL and TL texts. There *is* a translation act behind such problems but it is hardly accessible in retrospect; it can only be speculated on. The third sense ("PROBLEM$_3$") is also associated with a single translation event but, unlike the second, it is not retrospective in nature; rather, it considers the event as it is unfolding and may therefore be regarded as processual. Observation is here possible "only inasmuch as ACT$_3$ *has* left traces comprising more than just the end-product; most notably, temporary, **interim replacements**, on the one hand, and **reflections** on the other" (Toury 2002: 65; emphasis in the original). It is in relation to this third sense that data collection methods such as TAPs and keyboard logs come to the fore, as the possible multiplicity of solutions proposed during the translation event could not be unearthed by an analysis of the end-product alone.

Toury's analysis of the various meanings attached to the term *problem* in Translation Studies provides a basis for reassessing some earlier proposals for both the categorization of problems and the distinction between problems and difficulties put forward by some scholars. In particular, Nord's problem/difficulty distinction can be mapped onto the three meanings identified by Toury: Nord's objectively identified "problems" will thus appear to map onto Toury's first meaning, i.e the one focussing on issues of translatability. Nord's "difficulties" on the other hand would seem to map (whether they are overcome by translators or not) onto Toury's second and possibly third meaning (provided an account of process data is given). By the same token, Campbell's "difficulty" (as defined in Campbell and Hale 1999) seems to be synonymous with PROBLEM$_2$. Table 2.2 shoes how the various notions of problem presented so far can be mapped onto Toury's treatment of the concept. Notice how some notions can be seen to fall into more than one category. Pym's notion of problem as a function of risk, for instance, can be considered retrospective or process-oriented depending on whether it is used to reconstruct a given choice or to decide, in the course of what he calls the translation act, what the most appropriate translation solution is. By the same token, Campbell and Hale's "difficulties" can be observed in relation to either the ST-comprehension phase or the TT-production phase. In the former case, they could here be placed under PROBLEM$_1$ but are eventually assigned to the PROBLEM$_2$ category as Campbell and Hale's study only

concentrates on the production phase, making their discussion of difficult elements essentially a retrospective one (although still largely referring to characteristics of the *source* texts).

Table 2.2 - Mapping of notions of translation problem/difficulty onto Toury's analysis of the concept

PROBLEM$_1$ - source-oriented - prospective	**PROBLEM$_2$** - target-oriented - retrospective	**PROBLEM$_3$** - process-oriented - processual
Wills' "translation difficulties" Nord's "translation problems"	Campbell & Hale's "difficulties" Pym's risk-correlated "problems" Nord's "translation difficulties"	Lörscher's "translation problem" Krings' "translation problem" Nord's "translation difficulties" Pym's risk-correlated "problems"

Of the two notions, then, that of problem seems to be the broadest in meaning, the overarching concept, as it were, as attested by the fact that Toury chooses the word "problem" for an overview of what other researchers have also called "difficulties". As table 2.2 shows, reference to one sense or the other of the notion is not reflected in the choice of the label used for it by the various researchers. On the other hand, preference for one term over the other could also be seen as simply the result of collocational preference, so that when the source text is seen as problematic (whatever the reason) the term "difficulty" suggests itself , as the phrase "text difficulty" is definitely more common than "text problematicity".

As far as the present study is concerned, the assumption is that certain linguistic features of scientific texts are likely to be difficult for most translators. Whether this means that they should be considered objectively problematic remains to be seen, as the group of translators involved in the present study is only really representative of a particular stage in the acquisition of translation competence (that of 'semi-professionalism'; see 3.4 for more information on the participants to the experiments). It could be argued that a particularly expert translator (i.e. one with great experience *and expertise* in the translation of scientific discourse) can develop efficient decoding mechanisms for the features that here taken to be difficult precisely as part of his/her expertise, just as in other fields, e.g. advertising or law, translators may be seen to develop a particular ability to recognize (and transfer) the

distinguishing features of the texts they are faced with. Perhaps the most interesting aspect to emerge from this overview of the notions of problem and difficulty is precisely the fact that, however they are defined, they should be considered as dynamic concepts whose definition changes according to the level professionalism of translators. In other words, different kinds of problems or sources of difficulty can be seen to emerge at each stage of the acquisition of translation competence. Ideally, a final stage could be reached where the translator coincided with Toury's idealised "persona": someone with full mastery of the languages and cultures involved, a perfect capacity to interpret and analyze texts and, consequently, the ability to solve any translation problem or difficulty s/he might encounter. Before that final stage is approached, however, a model of what is difficult for translators might turn out to be useful, be it for training, evaluative or accreditation purposes. The present work is meant as a contribution in the development of such a model.

Chapter 3
METHODOLOGY

3.1 General methodological considerations

In the years after James Holmes first proposed[1] his much-quoted map of Translation Studies, studies falling within the descriptive branch (Descriptive Translation Studies, or DTS) have increased in number and have seen a shift of focus which has brought them from an almost exclusive interest in the target text to an in-depth investigation of the process of translation and the ways in which this influences the final, textual product. The division of DTS into three branches (*product*-oriented, *process*-oriented and *function*-oriented studies) does not mean, however, that investigating translation from any of these three perspectives can or should thrust aside the other two. As noted by Toury (1995: 11),

> individual studies of whatever denomination emerge as a twofold enterprise: each one is a local activity, pertinent to a certain corpus, problem, historical period, or the like [...], as well as part of an overall endeavour, an attempt to account for ways in which function, process and product can and do determine each other.

Historically, this need to take into account the interrelations between product, process and function has emerged in parallel with the introduction, in translation research, of more genuinely experimental methods. But Toury (1995: 12), again, notes how experimental work cannot lead to significant results unless its design includes all the relevant parameters and takes into account the way they interact with each other. So, for instance, in corpus-based research, when the aim is that of describing TL patterns for the translation of a given SL element, reference will often have to be made to 1) the function of that particular element, and, more generally, of the texts it appears in, in the target culture, and 2) the particular environment in which the texts were translated (the "context of situation", to put it in linguistic, Firthian-Hallidayan terms).

---

[1] See Holmes (1972 and 1988). Holmes' paper illustrating the general framework for the discipline of Translation Studies first appeared in 1972 ; the paper, however, only had limited circulation until well into the 1980s. Toury (1995) played a key role in making Holmes' ideas known to a larger public.

Given the now recognised need for descriptive and experimental studies, particularly in process-oriented studies, the lack of the 'raw materials' for research has been noted by many. From the theoretical and methodological viewpoint, over the last two decades the debate among researchers has centred on the following issues (cf. Hurtado Albir 2001: 172):

- the need for empirical data;
- the identification of suitable data collection methods for the research question at hand;
- the ways in which descriptive studies may influence theoretical discussions and these, in turn, can lead to new descriptive approaches;
- the opportunity of inserting each individual study within the framework of the discipline, identifying differences and similarities with questions investigated in other studies.

In particular, it has become clear that data regarding translation as a cognitive and psychological process were, traditionally, scarcely available. Most process-oriented studies which have appeared in the last two decades have, therefore, as a priority been concerned with the establishment of suitable data collection methods, and the validity and reliability of different data elicitation methods has, in itself, been presented as the object of systematic investigation (cf. Jääskeläinen 2000).

Increasingly, then, translation research has felt the need to conform to investigative methods capable of achieving the two essential aims of empirical science, i.e. description and explanation. On the descriptive side, research has been given the role of providing more rigorous and systematic accounts of translational phenomena than were available in the past, when most studies would not go beyond impressionistic, albeit plausible, accounts, especially as far as the translation process was concerned. On the explanatory side, attempts have been made at accounting for translational phenomena from different angles: causal ("why is X translated so?"),[2] procedural ("how is the translation of X arrived at?") and finalistic ("what is the translation of X aimed at?"). The descriptive and explanatory stage of research can then be followed by a predictive effort which, in a discipline such as translation, is

[2] As remarked immediately below, the format of an explanation in Translation Studies is bound to be probabilistic rather than deterministic (cf. Toury 2004). As regards causal explanations in particular, they should also be distinguished from cases of correlation between variables.

bound to be of a probabilistic nature. As regards the present investigation, the modest size of the sample of subjects severely limits the predictability potential of the analysis, which remains of an essentially qualitative nature. The decision to adopt a triangulation approach was a means of partially overcoming this limitation, but of course a larger representative sample (perhaps combined with a statistical interpretation of the findings) could enhance the predictive significance of the results.

3.1.2 Types of empirical study

Following Gile (1998; cf. also Williams and Chesterman 2002) empirical studies can be seen to fall into two different categories, "observational" and "experimental" studies. The former observe a given phenomenon as it manifests itself in the real world, with no intervention of the researcher. An observational study can aim, for instance, at describing how one or more professional translators face translation tasks, focussing on particular aspects of the task such as revision, use of reference materials or the differences between translation from and into the second language. Studies in this category may be of an exploratory nature or start from a specific hypothesis that the researcher wants to test. Observation is based on data collected with different methods (sometimes used in combination): audio or video recordings, questionnaires, keyboard-activity logs, etc. Experimental studies are those in which the researcher deliberately intervenes in the conditions of the task s/he wants to study. The aim is that of isolating one particular feature (the one to be studied) while keeping the others unchanged or constant. As in observational studies, different data collection methods are employed.

As far as process-oriented studies (both observational and experimental) are concerned, the data collected for description and analysis are usually classified according to when and how they are collected in relation to the process (Krings 2005: 348; cf. also Englund Dimitrova 2005: 66):

- 'on-line' data are those collected from the ongoing process: computer logging of the translators' keyboard activity, eye-gaze patterns, brain imaging and concurrent verbal protocols all fall within this category;

- 'off-line' data are those collected after the process of translation has ended: the translated text, the revisions and annotations made by the translators, retrospective verbalisations and questionnaires are all types of data belonging to this category.

Early process-oriented studies based on empirical data of either category have frequently been criticised for giving an incomplete picture of the translation process. In trying to overcome this limitation, an increasing number of studies have resorted to triangulation methods (see 2.2.2), which essentially means resorting to an analysis of experimental data collected through different methods. A number of studies which have appeared in the last decade follow this approach and combine on- and off-line data in order to test hypotheses on firmer empirical grounds (cf. Halskov Jensen 2002; Livbjerg and Mees 2002; Alves 2003; Englund Dimitrova 2005).

The present study also adopts a triangulation approach and uses both on-line (computer logging) and off-line (the TTs and errors) data. The rest of this chapter will illustrate in more detail the design the study is based on, describe the methods used in collecting the data for analysis and, finally, give an overview of the analytical model that was applied in investigating the data.

3.2 Design issues

The investigation the present work reports on is an empirical study conducted using a small cohort of translation students from the SSLMIT[3] of the University of Trieste in Italy. At the time they carried out the translation tasks for the investigation, these students were attending courses for the two-year specialised degree in Translation, roughly equivalent to a Master's diploma (see 3.4 below for detailed information on the students). The group of subjects is a cross-sectional sample of student translators at an advanced stage of training. Their selection as subjects was motivated by essentially opportunistic reasons.

The selection of students as subjects in process-oriented studies is often controversial. Their behaviour in a translation task is often assumed to be different from that of expert translators. However, as in the present study the research question

[3] Scuola Superiore di Lingue Moderne per Interpreti e Traduttori (Advanced School of Modern Languages for Interpreters and Translators).

had to do with a text-internal variable, i.e. ST style, the choice of a homogeneous group of subjects was considered of paramount importance, on condition that these subjects displayed a reasonable degree of second-language competence.

The hypothesis guiding the study is that both the style and the subject matter of a text appear to be a source of difficulty for translators. In particular, the investigation aims at isolating a particular feature of style (i.e. the nominalisation resulting from grammatical metaphor) as a source of difficulty. In a truly experimental design based on a sufficiently large sample, a null hypothesis can be refuted on the basis of an analysis of results using inferential statistics, thereby ruling out chance (up to a given point) as a factor influencing the outcomes. In such a design, it is possible to establish correlational or even cause-effect relations between an independent variable which has been manipulated and another variable which is potentially dependent on it. In the present study, the focus is on a possible association between aspects of the ST style and the performance of a small group of trainee translators which relies to a large extent on a series of qualitative judgements and interpretations. Where statistical profiles are provided of ST features or of aspects of translator performance, these are descriptive. Hence the hypothesis is not stated in its null form.

The same group of participants was used throughout the investigation. This was done with the aim of eliminating any confounding variables, such as individual differences in ability at the task. One potential problem associated with this decision was the possible emergence of order effects, i.e. effects brought about by the fact that participants in the study performed one task after the other. Order effects can either lead to an improvement in the participant's performance (due to practice or increased familiarity with the task or the task demands) or to a deterioration in performance, due to fatigue or boredom (cf. Harris 2002: 129). In other words, participants can become either more practised (thus performing better regardless of what varies in the task) or fatigued or bored (thus performing worse, again regardless of what varies in the task). As a way of controlling for such effects the order in which participants performed their tasks was partially randomised (for details, see section 3.5.1).

The next section discusses the theoretical background to the notion of grammatical metaphoricity and illustrates the criterion that led to the choice of the three different source texts submitted to the translators who participated in the study.

3.3 Materials: Selection of the Source Texts

The initial assumption in the present study is that the particular style adopted by writers in scientific texts acts as a source of difficulty for translators, on top of their possible lack of subject-matter knowledge. This hypothesis closely resembles that underlying Lassen (2003), a study investigating the *accessibility*[4] of technical manuals and, in particular, "whether one style rather than the other would facilitate comprehension" (Lassen 2003: 161). As in Lassen's study, style is initially characterised, very broadly, as a "cluster of signs" (cf. also Hodge and Kress 1988) or a "cluster of textual traits" (Killingsworth and Gilbertson 1992: 138). These textual traits are here described in terms of Systemic Functional Grammar, again in line with the approach adopted by Lassen (2003; cf. also the studies on the translatability of texts in MT quoted in 2.4).

Systemic Functional Grammar (SFG) has been chosen as a theory of language that is capable of providing a detailed description of both language variation and the rhetorical purposes served by texts. More specifically, the SFG notion of grammatical metaphor is treated here as the basis for providing the following: 1) a qualitative description of the source texts, aimed at establishing traits that act as possible sources of difficulty for translators; and 2) a quantitative analysis of these same source texts, where such element as nouns and 'nominal-type units' are counted (see 3.3.3 below) with a view to establishing different levels of (assumed) intrinsic difficulty to be tested empirically. Following Lassen (2003: 17), I am well aware of the possible objections to using this kind of analysis as a method for assessing text difficulty. Lassen's own study demonstrates how an unequivocal conclusion as to the intrinsic accessibility (or difficulty) of a text is not easily arrived at, as factors having to do with readers, rather than just texts, should also be taken into account. The quantitative description of the source texts, however, is intended mainly as an initial step and a yardstick against which further, qualitative analyses are conducted (based on the 'on-line' and 'off-line' data mentioned above).

This section will give an overview of register theory within the framework of Systemic Fuctional Grammar, describe the role and operation of grammatical

metaphor and finally illustrate how the notion of metaphoricity has been used as a basis for selecting the three different source texts employed in the present investigation.

3.3.1 Relevance of Halliday's model of register

The Hallidayan theory of register is intended to elucidate the aspects of the context of situation which contribute to variation in language use and the ways in which these aspects interact with the metafunctional meaning-making components of language. Halliday (1978) identifies three general aspects which he sees as contributing to the organisation of context: field, tenor and mode. These three variables have to do with the aspects of any situation which are likely to have linguistic consequences. In particular,

> **field** concerns what the language is being used to talk about;
> **tenor** concerns the role relationships between the interactants;
> **mode** concerns the role language is playing in the interaction.

The register of a text is determined by the interaction between these variables, each of which, in turn, interacts with the three meaning-making metafunctions that organise language: the ideational, the interpersonal and the textual. The relationship between the three contextual variables and the metafunctional components of language is illustrated in Table 3.1.

Table 3.1 – The relationship between metafunctions and register variables

Metafunction *(organisation of language)*	**Register** *(organisation of context)*
Ideational meaning:	*Field:*
resources for building content	social action
Interpersonal meaning:	*Tenor*:
resources for interacting	role structure
Textual meaning:	*Mode*:
resources for organising texts	symbolic organisation

Source: Eggins and Martin (1997: 239)

For the purposes of the present study, the three register variables can be considered as factors contributing to the criteria for selection of the source texts (with one in

[4] The definition of accessibility adopted by Lassen (2003: xv) is that proposed by Klare (1963) with reference to *readability*: "ease of understanding or comprehension due to the style of writing".

particular, "field", held constant, as different source texts were chosen that treated roughly the same topic).

However, in line with the initial assumption for the study which sees ST style as a determinant of translator performance, one specific component (i.e. mode) can be seen as particularly significant by virtue of its relevance as far as the stylistic profile of texts is concerned. Mode is the aspect involving such dimensions as the lexico-grammatical configuration and the rhetorical purposes of texts and the channel used for communication (spoken vs. written). It is the aspect which is most immediately responsible for the variation in style as defined in the present work, i.e. as preference for certain lexico-grammatical features over others. The source texts used in this study will thus be seen to vary principally along the dimension of mode and, more specifically, in relation to one particular lexico-grammatical resource determining textual variation: grammatical metaphor (details on the selected texts are given in 3.3.3 below).

3.3.2 Grammatical metaphor

The concept of grammatical metaphor has been introduced by M.A.K. Halliday (1985) as a type of metaphor complementing the more commonly discussed *lexical* metaphor. The notion of metaphor in general is seen by Halliday as referring to variation in the use of words. In lexical metaphor the variation is in meaning, so that "a word is said to be used with a transferred meaning" (Halliday 1985: 321). A lexeme with a literal meaning can thus be said to have the potential for metaphorical, transferred uses or meanings. This traditional perspective on metaphor is equated by Halliday (1996) with a view 'from below', i.e. one taking the word as the starting point. A complementary perspective would be one 'from above', i.e. starting from meaning and looking at the different ways in which this meaning is expressed or realised. Seen from this angle, metaphor can be defined as "variation in the expression of meanings" (Halliday 1985: 342). The label used by Halliday for this type of variation is "grammatical metaphor".

Whereas lexical metaphor is analysed in terms of literalness, so that the meaning or use of a lexeme can be said to be either literal or metaphorical (or, more generally, figurative), variation in the expression of the same meaning can be

defined, according to Halliday, in terms of markedness. A given form can thus be said to be the unmarked expression of a certain meaning, whereas other forms (entailing grammatical but often also lexical variation) are described as the marked expressions of that same meaning. The unmarked expression is the one that conforms to the "typical ways of saying things" (Halliday 1985: 321). This typical ways of expressing a meaning are also called by Halliday "congruent" realisations, as opposed to metaphorical, "incongruent" realisations. An example of a congruent sentence reconfigured as an incongruent realisation is given in Halliday (1997: 46):

> **Congruent realisation**
> *the driver drove the bus too rapidly down the hill so the brakes failed*
>
> **Incongruent realisation**
> *the driver's overrapid downhill driving of the bus caused brake failure*

The two sentences express the same meaning: the first employs an unmarked construction, while the second uses a marked construction resulting from the conversion of the two verbs (*drove*, *failed*) into nouns (*driving*, *failure*). This tendency towards nominalisation is indeed the main result of the operation of grammatical metaphor, so much so that in a large number of cases the two concepts can be seen to coincide.

One relevant implication of the concept of grammatical metaphor for the purposes of the present study is pointed out in Taverniers (2003). Given that various types of lexico-grammatical configurations can be compared as expressions of the same meaning, grammatical metaphor can be seen to give rise to a *scale of congruency*: "some expressions are typical realizations of the given meaning, and are defined as congruent; others are more or less congruent, as compared to the congruent realization(s)" (Taverniers 2003: 7).

The role played by grammatical metaphor in scientific discourse is discussed in Chapter 2 (Section 2.5.4), where it is linked to such notions as syntactic ambiguity and lexical density, both seen by Halliday as grammatical features of texts which are inherently "problematic" (cf. Halliday 1993b). Such features can actually be seen as by-products of grammatical metaphor, which remains the most significant single factor responsible for their emergence in scientific English. But what are the reasons behind the emergence of this 'syndrome' of features? Why, in other words, has grammatical metaphor acquired such prominence in scientific texts (both in English

and in other languages) as to become the preferred mode of expression of the writers of those texts? The evolution of scientific language towards grammatical metaphor can be traced first of all with reference to the emergence of grammatical metaphor in *general* language. Halliday (1996) sees grammatical metaphor as one of the meaning-making mechanisms that have emerged to enable users to produce and interpret abstract meaning. This evolution can be observed both phylogenetically, in relation to the history of language as a system, and ontogenetically, i.e. with reference to the history of the individual language user. There is, moreover, a third dimension in which the realisation of meaning can be seen to unfold in the direction of increasingly incongruent, metaphorical meaning, and this is the history of the individual text, where some meanings may be first introduced in clausal, congruent mode and then be reconstructed in nominal, incongruent mode, sometimes through intermediate steps.[5]

The language of science can be seen as one particular language variety where this evolution towards incongruence, or 'metaphoricity', is realised to its full potential (Halliday 1993a; see also 2.5.4). This is because grammatical metaphor, especially in the form of nominalisation, has established itself as a particularly suitable meaning-making mechanism for the argumentative needs of scientists. Metaphoricity has therefore become a typical, rather than exceptional, feature of scientific discourse, making it inherently complex. The relationship between metaphoricity and complexity, however, should not be seen in simplistic terms:

> It might be assumed that metaphor, while not inherently value-laden, is nevertheless inherently complex, and that the least metaphorical wording will always be the one that is maximally simple. But the concept of *plain and simple* is itself very far from being plain and simple; anything approaching technical language, for example, tends to become noticeably more complex if one tries to *simplify* it by reducing the metaphors. (Halliday 1994: 350; emphasis in the orig.)

This difficulty in simplifying technical language by simply removing grammatical metaphors may perhaps be explained by the very notion of typicality mentioned above (and the same is probably true of scientific language, although

[5] See, for instance, Halliday's (1993a: 56) description of how a grammatical metaphor is constructed step by step within the same text, starting from a verbal group and ending with a nominal group made exclusively of juxtaposed nouns: (glass) cracks – to crack (glass) – a crack (grows) – the crack (has advanced) – make cracks (grow) – (rate of) crack growth – crack growth rate.

Halliday does not explicitly refer to it here). If users of certain text types expect those text types to exhibit a certain degree of metaphoricity, then any dip beneath that threshold could strike users as unfamiliar usage and thus represent a potential source of incomprehension or at least disorientation for the readers of the text. This, of course, implies that, in terms of tenor, the role relationship between text writer and text reader remains stable: that is, that the text exhibiting an unusual decline in metaphorical realisations is one that is directed at readers who *are* expecting a higher degree of metaphoricity as compared to general usage.

3.3.3 Degree of metaphoricity as a selection criterion

Assuming, on the basis of the above, that grammatical metaphor lends an inherent complexity to scientific texts for the purposes of translation, the problem emerges of how texts might be quantitatively differentiated in terms of their degree of metaphoricity. In other words, a system of measurement is needed for placing texts on the "scale of congruence" mentioned above. In particular, for the purposes of the present study, a measurement system was needed that would help place the three different source texts along a scale of decreasing congruence, or increasing metaphoricity – and given that metaphoricity is strongly associated with nominalisation, this scale can be essentially be seen as one of increasing nominalisation.

The three texts that were selected are reproduced in Appendix 1. The selection was based on two sets of criteria: on the one hand, qualitative considerations related to the design of the investigation; on the other, a set of quantitative measurements aimed at determining the degree of nominalisation/metaphoricity found in the texts.

One of the texts (the extract from Michalske and Bunker's 1987 article published in *Scientific American*) was chosen on the basis of a simple but, in this context, decisive criterion: it contains one particular passage that has been presented by Halliday (1993a) as an exemplary specimen of the highly nominalised, highly grammatically metaphorical language used by scientists. From now on this text will be referred to as Source Text 1, or ST 1 in short. Taking ST 1 as 'reference text', two other text were sought where style (equated with lexico-grammar) should be the only

aspect seen to be varying to a significant degree. More specifically, the two texts would have to display different levels of grammatical metaphoricity as compared to the reference text. In particular, it was decided that one text should display a *lower* degree of metaphoricity, while the other should display a *higher* degree of metaphoricity. The reference text was placed 'in between' the other two texts in consideration of the fact that it appeared in *Scientific American*, a publication that has been shown (Hayes 1992; see also 2.5.3) to occupy only a middle position as far as textual complexity is concerned. Other journals, largely identified with those publishing reports of original research, have been shown by Hayes to display considerably higher degrees of complexity. The assumption (which is here left undemonstrated) is that these journals generally contain texts displaying a higher degree of nominalisation or metaphoricity than those published in *Scientific American*.

On the basis of the above considerations, two other publications were sought which treated the same subject matter as the reference text (the development of cracks in glass observed at molecular level) but displayed a different lexico-grammatical profile. At the higher end of the scale, an academic journal article (Michalske and Bunker 1993 in the *Journal of the American Ceramic Society*; henceforth ST 2) was identified as a suitable candidate; at the lower end of the scale, the candidate text (Banks n.d.; ST 3) was chosen from an on-line chemistry tutorial for college students. In each text, an extract was sought that displayed a different degree of metaphoricity from the other extracts. Metaphoricity was assessed using a method presented in Steiner (2002) plus two other measures: lexical density and the average length of the nominal groups contained in the texts.

Steiner's method was originally developed to assess grammatical metaphoricity in translated texts, with a view to testing the hypothesis that a somewhat reduced amount of metaphoricity is a systematic feature of translated texts relative both to non-translated TL texts and to their STs.[6] Steiner (2002: 222) assumes that in grammatically metaphorical texts there is a higher percentage of what he calls 'nominal'-type grammatical units relative to 'verbal'-type grammatical

[6] The underlying assumption is that, in translation, understanding the source text always involves, among other processes, the 'unpacking' of grammatical metaphors.

units.[7] Based on a Hallidayan constituency analysis of texts, Steiner (2002: 223) also puts forward a more specific hypothesis, i.e. that "more highly metaphorized variants should have more groups and phrases per clause and more of these three types of units together per clause complex than the congruent ones".[8] This count can also be presented as in the following formula used by Steiner (2002: 223) himself:

(1) ((number of groups+phrases) / (number of clauses)) / number of clause complexes

which should yield higher values for metaphorical than for congruent texts. According to Steiner (2002), both ratios – the nominal/verbal ratio and the metaphoricity ratio – can serve to give a measure for the degree of metaphoricity of a given text. If, as is the case here, the two ratios are applied to a series of texts, two resulting clines should be obtained which can be taken to express the variation in the degree of metaphoricity of those texts, with lower values representing a tendency towards congruence and higher values representing a tendency towards metaphoricity.

As a way of supplementing Steiner's formulas, two other measures were considered in comparing the three texts for their degree of metaphoricity/nominalisation. Lexical density, calculated by dividing "the number of lexical items by the number of ranking clauses" (Halliday 1994: 351), was chosen on the grounds that highly nominalised texts usually display high lexical density (cf. Halliday 1993b: 79). The other measure, i.e. the average number of lexical items per nominal group, was chosen as a measure of the structural 'compactness' or density of each text.[9]

The figures obtained once these four measures were applied to the three selected STs are given in Table 3.2.

---

[7] The 'nominal'-unit group includes prepositions, nouns, adjectives, and non-finite verbal groups; the 'verbal'-unit group includes conjunctions, verbs, adverbs and finite verbal groups (cf. Steiner 2002: 223).

[8] A "clause complex" is a group of paratactically or hypotactically related clauses (see Halliday 1985: Chapt. 7). Embedded clauses (e.g. relative clauses) appearing in nominal or adjectival phrases are considered as elements of the group or phrase they are in, not as elements of clause complexes.

[9] "Groups" are defined, in a Hallidayan perspective, as those units that come immeditely below clauses as constituents of the language at the lexico-grammatical level. Such units are labelled in terms of their class membership, i.e. what they are on their own. "Nominal groups" are one of the three possible classes of group, the other two being "verbal group" and "adverbial group" (for details

Table 3.2 – Measures of metaphoricity in the STs

	'Nominal' type units / 'Verbal' type units	Metaphoricity (Steiner)	Lexical density	Avg. length of nominal groups
ST 1: Michalske and Bunker (1987) - *Scientific American* (length: 225 words)	46/24 = **1.9**	((74/22)/11) = **0.30**	127/22 = **5.8**	83/38 = **2.2**
ST 2: Michalske and Bunker (1993) - *Journal of the American Ceramic Society* (length: 223 words)	40/15 = **2.6**	((55/21)/9) = **0.29**	140/21 = **6.7**	101/31 = **3.3**
ST 3: Banks (n.d.) - on-line tutorial for college students (length: 246 words)	61/45 = **1.4**	((100/27)/15) = **0.25**	128/27 = **4.7**	94/53 = **1.8**

On three of the four measures, the three texts yield similar increasing clines. When the 'nominal/verbal units' ratio is considered, ST 3 has the lowest ratio (1.4), while ST 2 yields the highest ratio (2.6); ST 1 gives a medial value of 1.9. A similar cline (with ST 2 in top position, ST 3 at the bottom and ST 1 placed in between the other two) can be observed for both lexical density and the average length of nominal groups. Lexical density is highest in ST 2 (6.7) and lowest in ST 3 (4.7). Likewise, the average length of nominal groups is highest in ST 2 (3.3) and lowest in ST 3 (1.8). The exception is represented by Steiner's metaphoricity formula: here ST 1 yields the higher value, i.e. 0.30, immediately followed by ST 2 at 0.29; ST 3 comes further down at 0.25.

Considering that for three out of four measures the same cline is confirmed, and that the difference between ST1 and ST2 is minimal for the remaining measure (Steiner's 'metaphoricity' formula), we conclude that the reference text (ST 1, the

---

on constituent analysis see the "Notes on the criteria for segmentation" in Appendix 2). The length of nominal groups is calculated on the basis of the number of lexical words contained in the group.

extract from Michalske and Bunker 1987) used for this study is, in terms of metaphoricity, on a median level between the other two selected texts, with one (i.e., ST 3) to be considered *less* nominalised and the other (ST 2) *more* nominalised – or one less metaphorical and the other more metaphorical. Style, defined in terms of grammatical metaphoricity/nominalisation, can thus be taken as one variable aspect across the three texts (which, of course, does not rule the other differences existing between them). Because, however, the differences between texts were not consistent on all counts, on this scale of increasing metaphoricity we take ST 1 to be located nearer to ST 2 than it is to ST 3, as visually represented in Figure 3.1.

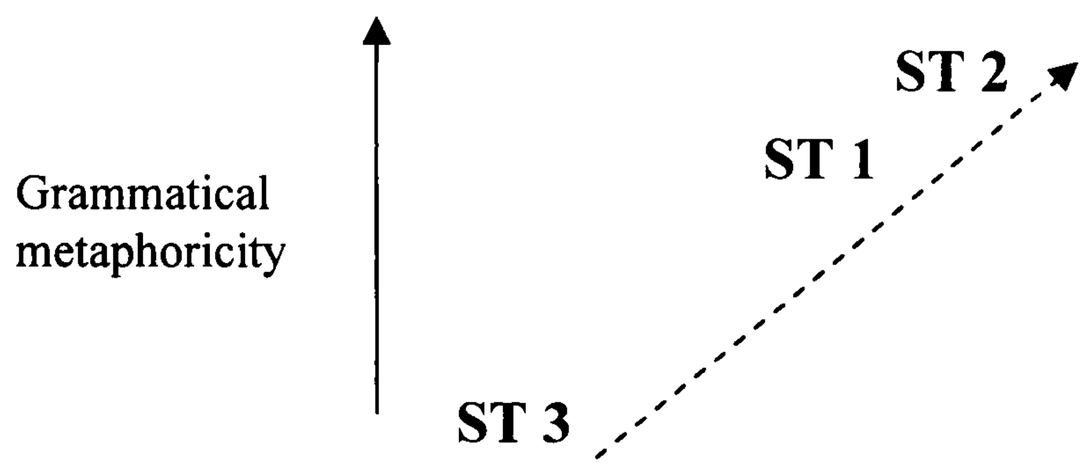

Figure 3.1 – Place of the three STs along a cline of increasing grammatical metaphoricity

In order to carry out the analysis (see 3.6), the STs were segmented – first into clauses and then in components of clauses, i.e. groups and prepositional phrases (see Appendix 2). This segmentation was carried out on the basis of a Hallidayan ranked constituent analysis of the three source texts; its aim was to make possible an eventual mapping of the findings onto the segmentation itself: in other words, results from both the CNA and Error Analysis were to be 'superimposed' on the ST segmentation, so as to see whether they clustered in or around particular units or clauses.

3.4 Subjects

The cohort of participants employed for the investigation consists of five translators, all students from the SSLMIT of the University of Trieste in Italy. At the time they carried out the translation tasks, these students were attending the first year of the two-year specialised degree in translation, roughly equivalent to a Master's diploma, offered by the SSLMIT. The degree is in Translation but it is heavily geared towards technical and scientific translation; the pre-entry qualifications for it are either a first-

level degree in Translation or a degree in Language Mediation ("Mediazione linguitica"). In terms of credits, the proportion of the course which is dedicated to practice on written translation of technical and scientific texts is 46 out of 120. Translation modules are assessed through written translation tasks, some of which are evaluated by experts of the specialised domain dealt with in the source texts.

I am well aware of the potential disadvantages of using translation students for a process-oriented empirical investigation. While the decision to choose translation students was made on essentially opportunistic grounds, following Campbell (1999: 36), I also accept that, in order to claim a reasonable degree of homogeneity in experimenting with translators, "the first-language output of second-language speakers of English who have been screened for entry into a professional translator education course is distinct enough from that of other possible groups of subjects". Moreover, at the time of the study four of the five students held a first-level, three-year degree in Translation, which means that they had already been trained and assessed on their general translation competence. The one student who did not hold a degree in translation is among those who declared they already had some experience in working as a translator. As all of the students had previously obtained first-level degrees in either translation from English or in English language and literature, I also take their level of competence in the English language to be at an advanced stage, particularly as regards reading comprehension skills.

The group of participants can also be considered homogeneous with regard to another crucial aspect of the present study: familiarity (or a lack thereof) with the subject matter of the three source texts submitted to each translator. As was already stressed in 2.4, the difficulty of a given text is not only determined by its lexico-grammatical and semantic features but can be seen as emerging from an interaction between the text and the reader, so that "a text is not accessible in itself but for a particular person in a particular place" (Cook 1995: 13). It follows that in order to assess text difficulty or accessibility, it is necessary to supplement the description of texts with a description of the readers and the interaction between these two factors (cf. Lassen 2003). Difficulty is thus seen as resulting from both variation in text and variation in readers (Cook 1995: 12). For the purposes of the present study, the five readers/translators of the three source texts are assumed to correspond to the same socio-cultural profile as far as familiarity with subject matter is concerned: in

particular, all the participants are considered as having no previous familiarity at all with the topic of the three STs (i.e. glass fracture observed at molecular level). In light of their training, however, the participants are assumed to have acquired familiarity with the language of science in general, both in English and in Italian. With respect to ST subject matter, then, the position of the five translators involved in the present study will be taken as a constant. This means that whatever findings the study arrives at in analysing the difficulty of the STs for this particular group of translators, these findings may be seen as suggestive of characteristics that are common to translators sharing the same socio-cultural profile as the group under study, if not to all translators in general.

Table 3.3 provides a summary of the relevant information for each participant in the study (in order to preserve anonymity, fictional initials are used to identify the translators).

Table 3.3 – Information on the participants to the study

Name	**Age (at time of tasks)**	**1st-level degree**	**Additional information**
CB	24	Translation (English, Dutch)	Had already carried out occasional translation jobs
MS	34	Modern Languages and Literature (English, French)	Had already carried out occasional translation jobs
MT	24	Translation (English, Spanish)	No previous experience of professional translation
SC	23	Translation (English, Spanish)	Had already translated a published book from English (a university textbook)
VM	24	Translation (English, Dutch)	No previous experience of professional translation

3.5 Data collection

As pointed out in 3.1.2, in process-oriented studies the data collected for analysis can be classified according to when and how they are collected in relation to the process. 'On-line' data are those collected from the ongoing process: computer logging of the

translators' keyboard activity, eye-gaze patterns, brain imaging and concurrent verbal protocols all fall within this category. In the present investigation, a special software application (*Translog*) was used to obtain logs of the keystroke activity performed by the translators during the translation tasks. These logs were then analysed to investigate the editing decisions made by the translators, as discussed in Chapter 4. The other category, 'off-line' data, includes data collected after the process of translation has ended: the translated text, retrospective verbalisations and questionnaires are all types of data belonging to this category. In the present context, two types of off-line data were used: the translated texts and the data obtained through an evaluation of the texts, carried out essentially through an analysis of errors. A total of three data sets was thus considered for the present investigation: keystroke logs, the target texts themselves and the data deriving from Error Analysis. Both the target texts and the errors identified in them are reproduced in Appendix 3. The following sections will illustrate in more detail the collection process for these three data sets, before moving on to an overview of data analysis methods (see 3.6).

3.5.1 Organising the translation tasks

The five translators who took part in the study translated the same three texts (ST 1, ST 2 and ST 3; see Appendix 1). The subjects performed their translation tasks in a computer laboratory at the SSLMIT of the University of Trieste. For each subject three different sessions were organised. The first session took place in November 2005; the second and third sessions took place in May 2006. The ST given to all subjects as their first task was ST 1. As a second task, some subjects worked on ST 2 while others translated ST 3. In the third task, those who had already worked on ST 3 received ST 2 and vice versa. The decision to swap texts for some translators was taken as a measure to minimise 'learning effects' on the part of the subjects (other than those deriving from the time span between the first and the second and third tasks).[10]

For each translation task, the subjects were equipped with the following:

[10] In the present context the labels Task 1, Task 2 and Task 3 will be used to refer to the translations of ST 1, ST 2 and ST 3 respectively, regardless of the actual order in which the source texts were translated by the subjects.

- a computer where they typed their translation using *Translog* as a word processor;
- a computer connected to the Internet, where they could conduct whatever searches they wished to;
- a selection of dictionaries, including: an English monolingual dictionary, an English-Italian dictionary, an Italian monolingual dictionary, a dictionary of science and technology in English, a dictionary of science and technology in Italian, a dictionary of synonyms in Italian (the selection remained the same for all tasks).

In each session, the translators were given a copy of the whole text the STs had been extracted from. They were given some background information on the text and were allowed approximately 45 minutes to read it in its entirety. After that, they were told what extract from the text they were expected to translate (which they were given on a separate sheet). Their brief was that they should translate the extracts as if they were addressed at the same audience as the one intended for the STs (see 3.3 above for details on the STs). No time limits were imposed for the completion of the tasks.

3.5.2 Real-time keystroke logging

One data collection method which has gained in popularity among researchers interested in studying the translation process is the logging of the keyboard activity performed by a translator engaged in a written translation task (see 2.2.2.1 for a description of the method and a review of some studies based on data collected through this method). Real-time keystroke logging offers a complementary method of investigating the process, to be used alongside other sources. If "one data source dries up at a certain point, data may be available in the other"; alternatively, if "data from two different sources can be analysed as converging, any finding will be stronger than if based on one data source only" (Jakobsen 2003: 70). In the present context, it was used as a source of data relating to the writing process carried out by each individual translator.

Computer logging can be considered a method with both general and ecological validity (Jakobsen 2003; Englund Dimitrova 2005). High general validity

is ensured by the fact that the method does not seem to influence the process or the characteristics of the final product of translation. As far as ecological validity is concerned, asking translators to use special software such as *Translog* to write their texts on computer is no significant departure from the normal circumstances of a translation task, where typing on computer is the standard method of text generation. It is true that *Translog* does not offer the same formatting options as more advanced word processors, but this is not a particularly significant disadvantage in studies where the research question does not concern the graphic presentation of the target text.

The logging software used for the present study was *Translog*, developed by Arnt Lykke Jakobsen (see Jakobsen 1999) specifically for the analysis of the translation process. Information obtained through keystroke logs produced by Translog can at times resemble that gathered through TAPs in the sense that it gives an indication of false starts, revisions and other text-generating processes. It must be remembered, however, that keystroke logs do not reflect the translation process *per se* but the writing process (Lorenzo 1999: 24; cf. also Englund Dimitrova 2005: 75).

For the purposes of the present study, where the attention is primarily on the phase of the translation process centred on the reading and comprehension of the ST, keystroke logs might be considered of marginal interest, as they reflect first and foremost the TT-generation process. Keystroke logs, however, give valuable information on aspects of the writing process such as false starts and revisions. Although relating primarily to the writing process, these particular data can be treated as a source of evidence of mental processing in the two phases which are assumed to precede TL text generation, i.e. comprehension and transfer. Revisions, in particular, can be seen as signals for ST segments which require more processing effort on the part of the translator. As such, for analytical purposes they can be treated much like the alternative renditions proposed by different translators that Choice Network Analysis was originally based on (see 2.3).

3.5.3 Translation errors

The fifteen target texts produced by the five subjects who took part in the study were submitted to an Error Analysis aimed at identifying inaccuracies in the texts

according to a set of relevant criteria (which are illustrated in detail in Chapter 6). For the identification of such inaccuracies a panel of three external evaluators was consulted. Two evaluators are researchers in materials engineering the University of Modena and Reggio Emilia in Italy. Both are Italian native speakers and possess the necessary English-language skills to understand and produce texts in English in their domain of specialisation. The third evaluator is a teacher of specialised translation (English-Italian) working at the SSLMIT of the University of Trieste. The evaluators were provided with a hard copy of the fifteen translations together with the three source texts. They were instructed to underline what they thought where the errors in the translations (as opposed to simply awkward solutions) and to provide either the right translation or a brief explicative comment relating to the seriousness of the errors. In general, the two domain experts tended to focus their comments on terminology, while the other evaluator, who was not a domain expert, tended to comment on elements that can be considered non-domain specific. ST elements identified by at least two evaluators as errors were all included in the analysis; elements identified by one evaluator only were considered for inclusion on a case-by-case basis (and upon further consultation with the evaluators). Note, also, that for one ST in particular (ST 1, the extract from Michalske and Bunker 1987) an Italian published translation was also available (Michalske and Bunker 1988) and was considered as a further benchmark for the identification of errors in the TTs under investigation.

Once a complete picture of the errors in the TTs was completed (see Appendix 3), the analysis aimed at identifying relevant categories, starting from the identification of two broad groups, "binary" and "non-binary" errors (Pym 1992; see 2.6.3 for background details). Note that inaccuracies deriving from spelling or punctuation errors were not included in the analysis because they were considered irrelevant for the purposes of the investigation (they are generally oriented towards the TL).

3.6 Overall analytical procedure

This last section presents a general overview of the analytical steps which have been followed in examining the data available for the present study. Further details on each individual step will be given in the following chapters.

A key procedure for analysing both the log data and the variation in translation solutions is a slightly modified version of Choice Network Analysis (CNA). Whereas in the original version as proposed by Stuart Campbell (2000) the model only compared the target texts of a group of translators working on the same text, in the present study the comparative analysis of the target texts was integrated with the analysis of the keyboard logs obtained for each task. In particular, this analysis of keystroke-based data looked at the interim solutions proposed by the translators. A CNA was therefore carried out both across participants (by reference to product data, i.e. the translations) and 'within' participants (by reference to process data). The general assumption was that a difference either in target-language renditions (i.e. because of 'inter-translator variation') or in the interim solutions proposed by the same subject (i.e. because of 'intra-translator variation') is the signal for an increase in processing effort. In particular, in both steps of the CNA, the variants, i.e. the different renditions found across the target texts or as interim solutions proposed by each translator, were linked to a specific ST segment. Each rendition was considered as a *token*; each variant of the renditions for the same ST segment was considered as a *type*. The identification of 'types' of solution took into consideration both the paradigmatic and the syntagmatic levels (further details are given in Chapters 4 and 5 respectively).

The analysis based on the Choice Network model was then integrated with results from an Error Analysis of the target texts. The aim was to look for further confirmation of the difficulty associated with certain segments of the texts.[11] In other words, if a given ST segment was found to emerge both in the CNA (at either steps) and in the Error Analysis, then this segment could be regarded as difficult or problematic. Once a group of such segments had been identified, it remained to be

[11] According to Hansen (2003: 35): "An important, if not the most important, source of data in this kind [i.e. process-oriented] research is the evaluation of the target texts. Processes are not very interesting if they cannot be seen in connection with any kind of result, the translation product" (cf. also Hansen 1999: 51 ff.) – a particularly relevant remark where text difficulty in translation is investigated.

seen whether they shared some specific lexico-grammatical features, and whether these features could give clues as to what the exact nature of the difficulty encountered by the translators was. This last step adopted a triangulation approach meant to integrate the findings which emerged from all the previous steps.

To sum up, the analysis of the three translation tasks was carried out in four steps, reported on in the next four chapters:

- **Step 1** (Chapter 4) looks at the keystroke-based data and analyses the editing performed by translators on their own drafts; the aim was to measure and characterise editing effort, seen as emerging from the interim solutions provided for a given ST segment; the analysis was then mapped onto the STs so as to identify a first cluster of ST 'hot spots', i.e. ST segments associated with higher editing effort and therefore with higher processing effort;
- **Step 2** (Chapter 5) looks at the target texts and analyses the degree of variation of the translation renditions proposed for each ST segment, both at the structural and the lexical level; as in the previous step, a cluster of 'hot spots' i.e. the ST segments which led to higher variation in the TTs (with higher variation also assumed to reflect higher processing effort), was identified;
- **Step 3** (Chapter 6) consists of an Error Analysis of the target texts; the aim is to single out inaccurate renditions provided by one or more translators, so as to give additional support to the identification of those renditions as indicative of textual 'hot spots';
- **Step 4** (Chapter 7) integrates the findings which emerge from the previous steps and looks at whether the increase in processing effort observed in steps 1 and 2 and the inaccurate renditions identified in step 3 converge on the same cluster of ST segments.

Besides reporting on the results of the triangulation of findings, Chapter 7 also presents (section 7.4) some final reflections on the methodological implications of the analysis.

Chapter 4
AN ANALYSIS OF EDITING AS AN INDICATOR OF EFFORT ORIENTATION

4.1 Introduction

This chapter reports on the first analytical step taken in investigating text difficulty in the translation of scientific texts. As the research question focuses on the role of the ST in contributing to difficulty, data analysis is carried out with constant reference to the STs employed in the study, i.e. it is conducted by ensuring that, for each translation task, data are observed and analysed as though they were always superimposed on a transparent, background 'grid' constituted by the ST in question, segmented according to the requirements of the analysis itself. The attempt, in other words, is to constantly map data (from both process and product) onto the units observed in the STs. These units have been identified by segmenting texts on the basis of a Hallidayan constituency analysis (see Appendix 2).

The analytical steps taken to close in on the difficult segments of each of the three STs (the 'hot spots' in each text) will therefore be based both on data relating to the writing process carried out by the five translators and on the final target texts (the product). A general overview of the analytical procedure was given in Chapter 3 (see 3.6). The present chapter starts the analysis by looking at process data, and particularly at the keystroke logs for each translation task that were obtained by using the *Translog* software. Partially following Campbell's (1998) approach to the study of editing by translators, this chapter will present an attempt at characterising editing as a reflection of the 'effort' made by the translators and then show how the edits thus characterised are distributed in the tasks performed by each translator.

4.2 Studying target-text editing by translators

The idea of studying the writing process carried out by translators is not new. Other researchers have, in the past, looked at how translators go about writing their drafts and then edit them; the aim was that of gaining more insights into various aspects of the translation process, mostly in relation with the reformulation phase. When computers were not generally available, or when translators were studied who did not

use computers to write their translations (e.g. students in exam conditions). the drafting and editing process was studied using handwritten texts or machine-typed texts. Campbell (1998) studied translation into the second language (L2) and looked, among other things, at editing in handwritten texts (see summary in 4.2.1 below).

Recent advances in software have made it possible to keep detailed records. or "logs", of all the keyboard activity performed by a computer user engaged in the writing of a text. The raw data provided by such software applications consist of the final printout of the text and a detailed register of the typing process giving information about the timing of each keystroke, pauses in typing, deletions, additions and so on. Some studies on the translation process employing keystroke logs have already been mentioned and discussed (see 2.2.2.1 and 3.5.2); in addition to those, another study, i.e. Dragsted (2005), is worth mentioning here which, besides employing keystroke logs, also has findings that are relevant for our present purposes. After submitting two different texts to two groups of translators (one of novices, the other of professionals), Dragsted looked at the translators' keystroke logs, with the aim of studying how the two groups cognitively segmented each text; she found that while differences in segmentation could be detected for the easier of the two texts, for the more difficult text similar patterns of segmentation emerged in the two groups, with professionals taking over the features that were characteristic of the novices in the other task. Dragsted's findings are interesting in that they seem to point to the existence of textual features that are intrinsically difficult, whatever the level of expertise of the translators, which is exactly the phenomenon the present work is seeking to investigate.

Unlike Dragsted, the aforementioned work by Stuart Campbell does not explicitly take difficulty into consideration but provides a useful general framework for the present analysis, especially from the point of view of some of the dimensions that Campbell observes in the editing performed by his subjects. In what follows, Campbell's research is briefly described and some implications it has for the present study are presented.

4.2.1 Editing by student translators working into L2: a summary of Campbell's (1998) study

Campbell (1998: 138-151) develops a framework for the analysis of real-time editing by student translators translating into their second language (Arabic to English). He uses this framework to make qualitative descriptions of the monitoring ability of the students, with monitoring intended as their ability to make effective changes to their output as they are engaged in translating.

The analysis is carried out on the translators' handwritten texts and focuses on the deletions and insertions made by the translators, each instance of them being an "edit" of the text. In his data, Campbell identifies six dimensions of editing: strategy, purpose, level, frequency, economy and effectiveness. Most of these dimensions are of immediate interest for the purposes of the present study, given that Campbell himself considers them in relation to the specific challenges posed by the source text, which he sees as polarised: syntactic at one pole, lexical at the other.

With regard, in particular, to the purpose of editing, a distinction is made by Campbell (1998: 138) between 'correction' and 'revision'. Correction is aimed at remedying a structural target language error and is firmly anchored in lexico-grammar. Revision is made "to the translation itself" (1998: 139): it is relational or intertextual in nature and is anchored either in semantic equivalence (i.e. when the intention of the translator in revising is that of saying what the ST really means) or in the appropriateness of the TT with respect to (TL) textual or genre conventions.

As far as level is concerned, Campbell's analysis focuses on three structural categories (clause, phrase and word) and looks at how edits are distributed in each category. An interesting point is made here by Campbell (1998: 140) as regards methodology: the level at which editing occurs is often hard to determine, the uncertainty essentially deriving from "the attempt to achieve analytical economy by describing a dynamic process with static levels". Using keystroke logs, as in the present study, may even have the effect of increasing such uncertainty on the part of the analyst, as the finer-grained and real-time representation provided by the logs (as opposed to Campbell's handwritten texts) further magnifies the dynamism of the process. To give an example, a false start in the handwritten text may be represented by a word that was deleted and then replaced with a new initial word for a phrase or clause. In the logs, it is not infrequent to have series of letters (i.e. chunks rather than

entire words) as false starts (cases such as these, however, were rare in the logs obtained for the present investigation).

Frequency of editing is counted as number of edits per 100 words of source text, while editing economy, another dimension, is calculated as the mean number of words per edit by each translator.

Finally, effectiveness of editing is measured using the text segment as its basic unit;[1] it is based on a calculation of the percentage of segments in a given text with uncorrected errors and no edits, uncorrected errors and edits, edits and no uncorrected errors, and no edits and no uncorrected errors. By looking at these percentages, a picture may be obtained of the quality of the output and the effectiveness of the editing performed by the translators (most of which, Campbell observes, was aimed at correcting the target text).

A summary of the dimensions of editing in Campbell's study is given in Table 4.1. The implications of Campbell's findings in the context of the present study are discussed in the following section.

Table 4.1 – Dimensions of editing, from Campbell (1998: 142)

Dimension	Features
Strategy	False Start, Bracketed Alternative, Deletion, Insertion and Partial Switch
Purpose	Correction, Revision
Level	Word, Phrase, Clause
Frequency	Low to high; edits per n words
Economy	Low to high; words per edit
Effectiveness	Interaction of uncorrected errors and no edits, uncorrected errors and edits, edits and no uncorrected errors, and no edits and no uncorrected errors

4.2.2 Implications of Campbell's findings for the present work

Comparing the data on these different dimensions of editing across translators, Campbell (1998: 142-150) arrives at some conclusions that can have a direct bearing on the present study, in that they show how a profile of the monitoring ability of the individual translators interacts with what Campbell assumes to be the main source of difficulty in each of the two source texts that he used for his study (one presented as

[1] No further definition of "segment" is provided by Campbell.

posing structural, i.e. syntactic, challenges; the other described as more difficult on the plane of lexis; see below). It should be borne in mind that Campbell's analysis is part of a larger study aimed at identifying facets of translation competence in relation to translation into the second language. His aim and his materials (i.e. the source and target texts) are therefore different from those characterising the present study. Nonetheless, the difference seems to be primarily one of focus or perspective, and the relevance of Campbell's investigation for the present study concerns both findings and methodology.

For the interpretation of editing data, Campbell focuses on three translators in particular and offers a qualitative description of their ability to monitor output in the TL. Data on editing are considered in all of the six dimensions described above and are interpreted in the light of two further dimensions:

1) an evaluation of the translations, which, as acknowledged by Campbell (1998: 140) himself, is heavily "biased" towards correction (as opposed to revision; see above). The better translations, in other words, are those where the number of violations of structural rules in the TL is reduced to a minimum – an approach that may have to be modified when considering translations into the translators' first language: assuming that, in this case, translators are likely to be more confident with TL structures and forms, the 'bias', or rather the focus of attention, could here be towards the two aspects that Campbell sees as the object of editing carried out for revision purposes, i.e. semantic equivalence between the ST and the TT and TL appropriateness;
2) the specific challenge posed by each of the two source texts: syntax for one text, which is described as being a "dour editorial piece with much subordination and transparent lexis"; lexis for the other text, described as using "jokey vocabulary, metaphors and simple sentences" (Campbell 1998: 138).

Campbell's findings offer a useful test-bed for the hypothesis proposed in the present study about the role of lexico-grammar in determining ST difficulty. Certain specific dimensions of editing appear to be particularly relevant in this respect. As regards editing level, Campbell finds that for the text presenting syntactic challenges translators are seen to make more phrase- and clause-level edits, whereas for the texts posing essentially lexical difficulties word- and phrase-level edits are favoured.

This particular finding could be used to provide supplementary confirmation of my hypothesis on the role of ST syntax in posing a challenge for translators: where more phrase- and clause-level edits are observed, the source text could be taken to be more syntactically challenging; where a majority of word- and phrase-level edits are present, the source text could be taken to be more lexically challenging. Taking into account the different directionality of the translations (Campbell's translators worked into L2, those participating in the present study translated into L1), the effect of syntax in determining difficulty may even have to be 'magnified'. Campbell repeatedly makes reference to the different nature of the difficulty posed by each of the two STs employed in his study, which he presents as an assumption rather than an outcome of his study (his aim is to establish how effective different translators are in editing their texts). In particular, the underlying assumption, on his part, seems to be that complex syntax causes difficulty mainly in the TL reformulation stage. In the present context, where translating into L1 is observed, the challenge posed by complex syntax in the ST may well turn out to affect both the comprehension and the reformulation stage or perhaps the comprehension stage only (the translators being mother-tongue speakers of Italian, i.e. the TL, we can assume that they did not experience difficulties in the target texts at the structural level).

4.3 Studying editing in keystroke logs

The log files obtained from Translog, documenting the keyboard activity carried out by the translators during each of the three tasks submitted to them for the present study, can be used to observe the editing process that led to the final version proposed by each translator for each ST. A first analytical step on the keyboard log data can be that of counting the number of changes made to the draft, or "edits", observed in each individual task.

In the present context, where translation into the translators' L1 is being studied, the focus will be on edits aimed at revision, i.e. those motivated either by the translators' intention to reassess the relationship between ST and TT in terms of semantic equivalence or by an attempt to conform to TL textual or genre conventions. In the first scenario, changes are introduced when the translator realises that the what s/he has typed is not an adequate equivalent for a given ST unit and

therefore replaces it with a TL item that she considers more adequate in terms of propositional content. The translator is thus discarding a previous TL solution, which can thus be seen as an *interim* rendition (cf. Toury's "interim solutions"). and arrives at a new TL rendition which, in turn, can become a second interim rendition if the translator decides to replace it once again and so on. In the second scenario, changes are determined by the translator's wish to improve the TT in terms of TL appropriateness. These, too, can give rise to a succession of interim solutions when more, successive changes to the same translation unit are made. Other changes still may emerge as corrections made by the translator to casual typing errors, oversights or misspellings of which the translator might or might have not been aware in the first place (such changes will not be focussed upon in the present analysis)

Correction is assumed to be a secondary dimension, given that, as native speakers of the TL, translators are expected to master its morphology and syntax. This is generally a safe assumption, although in real-life situations there may be occasional lapses in TL competence and cases where an L1 error occurs on the ill-defined border between general language competence and specialized language competence. What if, for instance, a translator is not aware of the irregular plural form of a specialist term? Would this count as a lapse in TL competence or as insufficient knowledge of the specialist domain? Such questions are generally outside the scope of the present study, but they will have to be borne in mind when interpreting aspects such as the errors made by translators.

A dimension not considered in Campbell's study on editing was time. As he was working on handwritten texts, Campbell had no possibility to look at when exactly in the process a translator had made a certain change to his/her draft. Handwritten texts only give a 'flat' representation of the editing process. Using software to keep a log of the writing process allows researchers to observe editing as it unfolds in time. This means that an edit typed by the translator can be observed at the exact moment in which it was introduced during the writing process and that the duration that specific instance of editing ban be measured. In the present context, however, only marginal attention will be paid to these particular temporal dimensions, which are probably of fundamental importance in studies focussing exclusively on the process of translation.

In the following, the categories of edits identified for analysis will first be presented (4.3.1) and an explanation of how each individual edit was counted will be given (4.3.2).

4.3.1 Identifying types of edit as indicators of effort orientation

In the context of the present investigation, the aim in using editing data is twofold. First, editing is used to determine the *orientation* of the effort the translators are seen to make. Orientation is here considered as a pointer to the purpose of editing or, if seen from a complementary perspective, to the motivation or concern behind an individual instance of editing. Second, once edits are assigned to a category based on their orientation, their respective *frequency* is counted in each task, both in aggregate and for each individual translator (see 4.4). In looking at the frequency of edits, the aim is to see whether the three tasks present significant differences in the amount and orientation of the effort required to translators and whether these differences are randomly or consistently distributed across the three tasks.

Edits were placed in three categories:

1) ST-oriented edits;
2) terminological edits;
3) TT-oriented edits.

The first two (ST-oriented and terminological edits) contain edits made for the purpose of reassessing the relationship of semantic equivalence between the ST and the TT, the third (TT-oriented edits) comprises edits that have the purpose of improving the TT in terms of TL appropriateness (all three categories can be characterised as "revisions" in Campbell's terminology; see 4.2.1 above). A few examples will be given to illustrate each of the three categories.

In Task 2, the translation of an extract from Michalske and Bunker (1993), the translation of segment 4 (example 1) was repeatedly edited by one of the translators, VM, as revealed by the keystroke log:

(1) *the mechanism of stress/environment dependent crack extension.*

VM's successive edits on the translation of this segment can be reconstructed as follows, each line representing an interim rendition obtained from what VM typed and then deleted, finally arriving at the translation presented in (d):

(1)' (a) il meccanismo dipendente dalla
[the mechanism dependent on.the]

(b) il meccanismo di estensione della fenditura dipendente dallo stress e
[the mechanism of extension of.the split dependent on.the stress and on.the

dall'ambiente
environment]

(c) il meccanismo di estensione della cricca nel legame tensione/ambiente
[the mechanism of extension of.the crack in.the bond stress/environment]

(d) il meccanismo di estensione della cricca dipendente dalla tensione e all'ambiente
[the mechanism of extension of.the split dependent on.the stress and on.the
environment]

VM's uncertainty in rendering the ST segment in Italian is here apparent both at the lexical and the syntactic level. At the level of lexis, or word choice, she is not sure what the correct equivalent for *crack* is, hence a first rendition with *fenditura* then replaced by *cricca*. The same goes for *stress*, first translated with *stress* and then changed to *tensione*. VM, however, is equally uncertain as to what the correct overall interpretation of the ST segment is at the syntactic level. TL renditions from (b) to (c) show how she is engaged in making a series of successive attempts at decoding the string of nouns the ST has in the prepositional phrase starting with *of*.

Interpreting this uncertainty exclusively in terms of a problematic decoding of the ST at the syntactic (or structural) level is certainly an oversimplification, as there may well be other factors that play a role in the correct interpretation of this particular ST segment. A likely candidate in this respect is the background specialised knowledge expected or presupposed in the readers the ST was originally addressed to. Domain specialists reading the ST would probably bend their syntactic interpretation of the segment so as to make it fit their own conceptual map of the field – a map they were very likely to share with the writers of the text. By integrating specialist knowledge with a "decoding competence" at a purely syntactic level, is seems therefore safe to assume that specialist readers of a segment such as (1) would arrive at its correct interpretation sooner or with less effort than other readers. This entails that other readers (and translators are likely to be among them)

would have to rely, instead, on an essentially syntactic decoding of the text, with little help from their own background knowledge of the field. Where process data testify to the translator's lack of confidence in interpreting the meaning of an ST segment over and above the meaning of single terms, as in (2), this lack of confidence can also be seen as a function of aspects having to do with the syntactic configuration (or lexico-grammar, in Hallidayan terms) of the segments.

As already mentioned above, the editing process represented in (2) also includes instances of edits consisting of changes in word choice, e.g. the change from *fenditura* in (b) to *cricca* in (c) as a TL equivalent for *crack*. In terms of syntagmatic constituency, these changes occur at a lower level than the edits leading to a recasting of the whole segment – only single words or phrases are affected. Specialised lexical competence, i.e. terminological competence, clearly plays a role here: translators are seen to concentrate editing at this level on specialist terms that they are not familiar with. So, while the purpose of editing at this level is once again that of renegotiating the relationship of lexical semantic equivalence between the ST and the TT, a difference can be observed as regards the level at which edits occur. Unfamiliarity with terminology causes editing mainly at word level. Difficulty in interpreting syntax causes editing primarily at higher levels than the word, i.e. phrase and clause levels.

In terms of effort orientation, editing caused by uncertainty on the correct structural interpretation of the ST, as in example 1 above, can be characterised as *ST-oriented*. For editing focussed on terminology, effort orientation is harder to characterise. Unfamiliarity with either terminology or the underlying concepts (and all their implications for the overall interpretation of a text) cannot be linked to any ST alone. Especially for translators who are still completing their training, the lack of domain-specific background knowledge must be taken as a given, a sort of initial 'handicap' that translators can only hope to reduce as they become more acquainted with the specialised domain in question. On the other hand, just as structural interpretation of a specialised text can be assumed to rely to a greater or lesser extent on background knowledge, terminology can be seen as an essentially lexico-grammatical phenomenon - the ultimate example of grammatical metaphor, where the encapsulating power of nouns is observed at its full potential. The difficulty in comprehending scientific texts "lies more with the grammar than with the

vocabulary", according to Michael Halliday, who then goes on to say that "it is the total effect of the wording – words and structures – that the reader is responding to" – "technical terms are part of this overall effect" (Halliday 1993b: 71).

This "overall effect" notwithstanding, editing determined by terminological uncertainty will be distinguished here from ST-oriented editing and included in a second category, called *terminological editing*. The reasons why such a sharp distinction is unsatisfactory should be clear at this point, but it is felt that keeping the distinction will later on present benefits in terms of the explanatory power of the results obtained from the analysis of editing. Grouping genuinely ST-oriented structural edits together with changes made on terminological grounds could, in other words, be justified on a purely theoretical basis, but in operational terms it risks compromising the attempt at assigning a more specific weight to structural aspects of difficulty. As an example of terminological editing, consider the following segment taken from ST 2 and the two successive TL renditions proposed by MT:

(2) *the rate of crack extension in silica glass in humid conditions*

(2)' (a) il tasso di estensione della fenditura nel vetro di silice in condizioni di umidità
[the level of extension of.the split in.the glass of silica in conditions of humidity]

(b) il tasso di propagazione della cricca nel vetro di silice in condizioni di umidità
[the level of propagation of.the crack in.the glass of silica in conditions of humidity]

In particular, rendition (a) was proposed by MT when writing the first draft of the text. Rendition (b), where only the equivalent for *crack extension* is changed, was arrived at during a final revision of the TL text.

The third category of editing considered here is *TT-oriented editing*, by which I mean editing which is linked not to a renegotiation of semantic equivalence but, rather, to the translator's attempt to provide TL renditions conforming to (assumed) TL textual or genre conventions. As an example, let us consider the following ST segment:

(3) *at speeds greater than one millimeter per second.*

CB's translation for this segment is arrived at in three stages:

(3)' (a) a velocità superiori
[at speeds greater]

(b) a velocità che superano un millimetro al secondo
[at speeds that surpass one millimeter per.the second]

(c) a una velocità che supera il millimetro al secondo
[at one speed that surpasses one millimeter per.the second]

The changes made (first to the initial, incomplete draft, then to the second TL rendition) do not alter the meaning of the segment. The only possible explanation for the changes is the translator's desire to improve the TT stylistically. There is no renegotiation of semantic equivalence and the perspective from which editing is considered is entirely 'TT-internal'.

To sum up, based on the keyboard log data collected for each task performed by the translators, three different categories of editing have been identified. These three categories share the same dimension, i.e. purpose, but are differentiated in terms of the orientation of the effort they reflect. ST-oriented editing emerges from successive attempts made by the translators at reaching the correct structural interpretation of the STs. Terminological editing is that performed on specialised terms. TT-oriented editing is that aimed at improving the TT in terms of appropriateness to TL textual and genre conventions. Before moving on to analysing the distribution of each category in the three translation task, a brief explanation will be given of how edits were counted.

4.3.2 Counting edits

Edits were counted with a view to measuring their distribution in each task and for each translator. The unit used for reference in counting the edits was the ST segment (see Appendix 2 for the segmentation of each ST). The first TT draft a translator typed for a given ST segment was considered the first interim rendition for the said segment. Each time a translator was seen to introduce a change to the rendition, this change was considered an edit and counted as such. There would be a number of cases where the changes introduced by a translator on an interim rendition affected different parts or different aspects of the TT segment. In example (1) above, for instance, TT segment (b):

il meccanismo di estensione della fenditura dipendente dallo stress e dall'ambiente

[the mechanism of extension of.the split dependent on.the stress and on.the environment]

is changed by the translator (VM) into (c):

il meccanismo di estensione della cricca nel legame tensione/ambiente
[the mechanism of extension of.the crack in.the bond stress/environment]

with the change entailing both a recasting of the phrase and a replacement of a specialised term (*fenditura*, becoming *cricca*). In such cases, the recasting was counted as one edit (and judged to be ST-oriented) and the change in word choice was counted as another edit, this time assigned to the category of terminological edits. Overall, with reference to ST segment (1) above, 4 edits were counted, as illustrated below.

ST segment	**TT renditions**	**Edits leading to final TT rendition**
the mechanism of stress/environment dependent crack extension	(a) *il meccanismo dipendente* (b) *il meccanismo di estensione della fenditura dipendente dallo stress e dall'ambiente* (c) *il meccanismo di estensione della cricca nel legame tensione/ambiente* (d) *il meccanismo di estensione della cricca dipendente dalla tensione e dall'ambiente*	(a) to (b) = 1 edit, i.e. recasting of phrase (b) to (c) = 2 edits, i.e. recasting of phrase plus substitution of term ("fenditura") (c) to (d) = 1 edit, i.e. recasting of phrase

Figure 4.1 – Counting edits in successive TL renditions typed by VM for segment 4, ST 2

Once edits were counted in all three tasks, they were assigned to one of the three categories discussed in the previous section, which was not always a straightforward decision. In the above figure, the introduction of the word "legame" (literally "bond") in (c) is a case in point. The word, in itself a specialised term, does not seem to have correspondents in the ST and its insertion by the translator was probably an attempt at rendering the idea of "dependence" conveyed in the ST with the adjective *dependent*. In counting and defining the edits, the decision eventually taken in this particular case was to consider the insertion of "legame" in the TT as part of the recasting of the clause, which in turn was motivated by the translator's uncertainty over the correct structural interpretation of the ST phrase. Thus, for the segment

represented in figure 4.1, a total of 4 edits was counted, of which 3 were labelled as ST-oriented edits and the remaining one as a terminological edit.

Example (3) discussed in the previous section provides another instance of uncertainty in deciding how many edits a succession of new renditions entails. this time on changes made exclusively in a TL/TT-perspective. The following figure illustrates editing on this segment as reconstructed from the Translog file obtained for CB.

ST segment	TT renditions	Edits leading to final TT rendition
at speeds greater than one millimeter per second	(a) *a velocità superiori* (b) *a velocità che superano un millimetro al secondo* (c) *a una velocità che supera il millimetro al secondo*	(a) to (b) = 1 edit (b) to (c) = 2 edits

Figure 4.2 – Couting edits in successive TL renditions typed by CB for segment 35, ST 1

The ST segment here was *at speeds greater than one millimeter per second*, for which the translator's very first, partial equivalent was *a velocità superiori*, then changed to *velocità che superano* and completed. The first change is then

superiori (adjective) > *che superano* (relative clause).

which can be taken as a structural, TT-oriented edit. The new, complete rendition for the segment, *a velocità che superano un millimetro al secondo*, is then changed again into *a una velocità che supera il millimetro al secondo*, the changes now affecting the noun *velocità*, the verb following it and the indefinite article *un*, so that

a velocità che superano > *a una velocità che supera*
[at speeds that surpass] [at a speed that surpasses]

and

un millimetro (indefinite article) > *il millimetro* (definite article)

The change from *che superano* to the singular verb form, *che supera*, depends on the choice to render *velocità* in the singular. Therefore, the two changes were here considered as one edit. The other change in the segment (definite to indefinite article) is unrelated to the first two and is thus considered as a different edit. Overall, CB, the

translator, made a total of 3 edits in her way to the final rendition of the segment. all of them TT-oriented.

To sum up, this section introduced the definition of 'edit' taken into account in analysing the logs of keyboard activity obtained through Translog. In particular, edits were equated with changes that the translators made to their drafts with one of two purposes: 1) to reassess the relationship of semantic equivalence between the ST and the TT; 2) to conform to TL textual or genre conventions. (The edits considered here can therefore be compared to what Campbell calls instances of "revision"; see 4.2.1 above). The three categories used to group edits were then illustrated ('ST-oriented'. 'terminolgical' and 'TT-oriented') and examples were given of how each individual instance of editing was counted in the log files.

4.4 Frequency and distribution of editing in the three tasks

Having explained how individual instances of editing were categorised (with respect to purpose) and counted, I will now move on to describe editing frequencies as observed in the three translation tasks, both in aggregate terms and for each individual translator. In particular, looking at editing in aggregate terms will serve to identify possible general patterns in the data; once these patterns are indentified, they can be mapped on to the STs under investigation (see 4.5.)

4.4.1 Aggregate editing in the three tasks

The three tables below summarise, for each of the three tasks, the number of edits performed by the five translators participating in the study and indicate the aggregate figures for each task. Recall that the number of words for the ST used in each task is almost the same (ST 1, N = 225; ST 2, N = 223; ST 3, N = 246), which makes the figures relating to each individual task roughly comparable with one another.

Table 4.2 – Frequency of edits in Task 1, per category and translator

	ST-oriented *f*	Terminological *f*	TT-oriented *f*	Total
CB	1	3	18	22
MS	6	11	21	38
MT	6	1	16	23
SC	9	2	15	26
VM	10	1	20	31
Total	32	18	90	140

Table 4.3 – Frequency of edits in Task 2, per category and translator

	ST-oriented *f*	Terminological *f*	TT-oriented *f*	Total
CB	3	1	5	9
MS	10	1	11	22
MT	4	11	18	33
SC	4	4	7	15
VM	6	4	9	19
Total	27	21	50	98

Table 4.4 – Frequency of edits in Task 3, per category and translator

	ST-oriented *f*	Terminological *f*	TT-oriented *f*	Total
CB	0	0	8	8
MS	3	5	23	31
MT	2	3	30	35
SC	1	0	11	12
VM	1	2	9	12
Total	7	10	81	98

Editing in the three tasks will be considered first in relative terms, i.e. taking into account the different categories of editing as a proportion of the total number of edits observed for a task. Figure 4.3 gives the relative frequency in percentages for each category of edits in the three tasks.

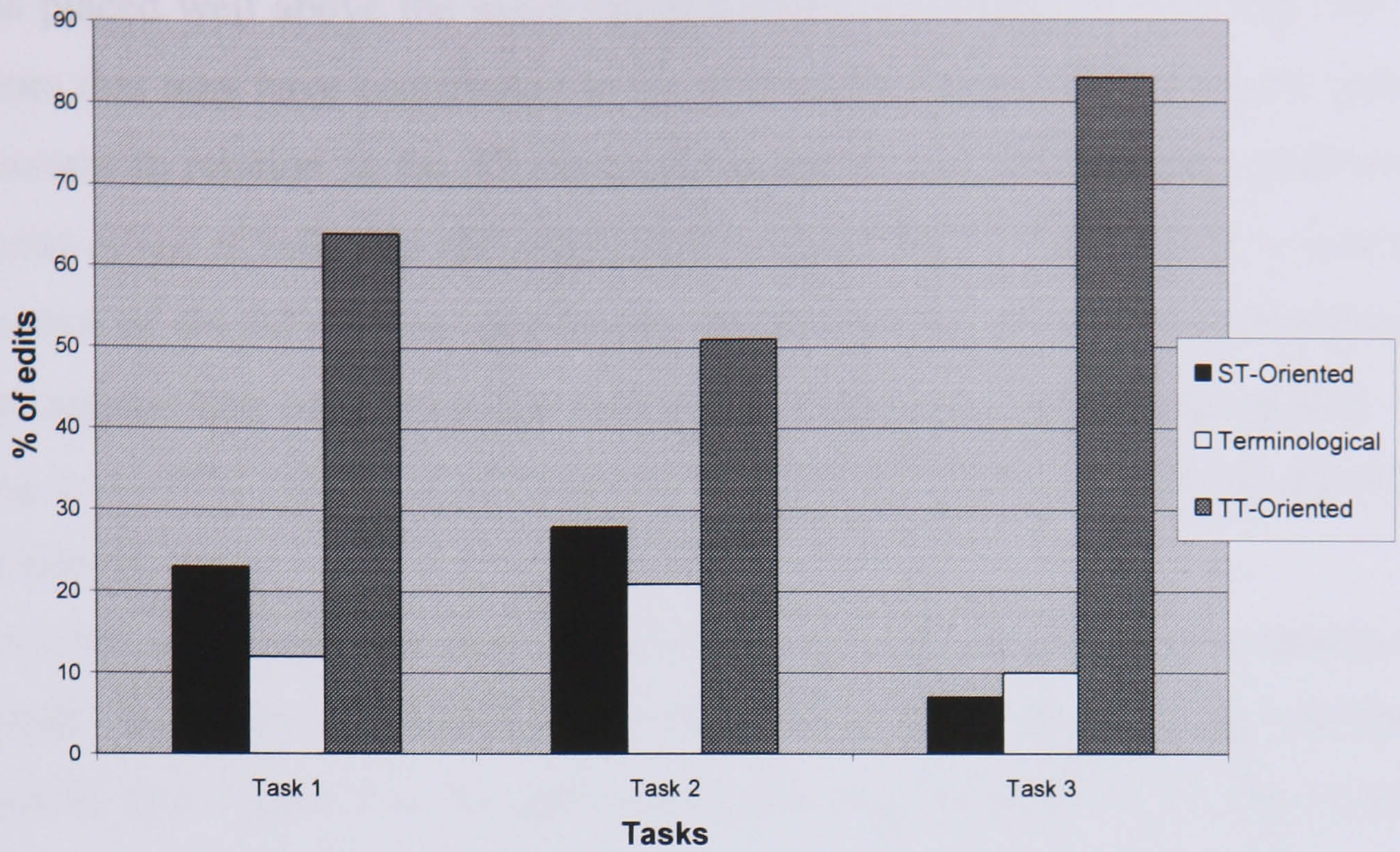

Figure 4.3 – Percentage of each category of edits in the three translation tasks

For all three tasks, editing is, for the most part, target-text-oriented, i.e. most of the editing effort is aimed at ensuring that the TT is appropriate to the textual and genre conventions of the TL. Editing reveals, therefore, that for this group of translators TL appropriateness was the main preoccupation when drafting the translation – not a particularly surprising finding given the nature of editing as a source of process data. Even so, as the figure shows, there is considerable difference among the tasks as far as the overall amount of TT-oriented editing is concerned. For one task in particular, Task 3, TT-oriented editing is by far the largest category, but its predominance gradually decreases in the other two tasks. The role played by the ST in determining this decrease can, in a sense, be seen as the focus of the present study.

Terminological editing ranges from a minimum of 10% of overall editing in Task 3 to a maximum of 21% in Task 2. The difference among tasks may well be due to a different terminological density in the three corresponding STs. In other words, more terminological edits may be observed for one task quite simply because there is a higher number of terms in the ST associated with that task. The different lexical density observed in the three STs (lowest in ST 3, highest in ST 2; see Table 3.2) seems to give some confirmation to this hypothesis.

As regards ST-oriented editing, Task 2 has the highest relative frequency (27%) while Task 3 has the lowest (7%). Task 1 comes in between the two, but at

22% it is placed well above the mean relative frequency (18%) for the three tasks. The factors that may have contributed to the differences observed between the tasks can be sought in relation to the ST themeselves (recall that editing was considered ST-oriented when it reflected the translator's uncertainty on the correct structural interpretation of the ST). More specifically, an attempt can be made at identifying structural aspects that vary from one task to the other – the purpose being that of observing a possible pattern of co-variation between ST structural aspects and ST-oriented editing.

As may be recalled, the three source texts were chosen according to measures that helped to place them along an ideal cline of increasing grammatical metaphoricity (see Figure 3.1). In particular, it was found that ST 3 (i.e. the extract from Banks n.d. used for Task 3) could be placed towards the lower end of the cline, while ST 1 and ST 2 were positioned towards the higher end of the cline, with ST 2 at a slightly higher level when measures based on nominalisation were taken into account (for the measures of metaphoricity see 3.3.3).

As pointed out in Chapter 3, "grammatical metaphoricity" can be seen as the independent variable used in this study. The three STs submitted to translators, in other words, were chosen because they were observed to change along this particular dimension. The cline represented in Figure 3.1 is an illustration of how this variable of metaphoricity is seen to change from one task to the other (note, again, that ST 1 comes considerably closer to ST 2 than to ST 3). Having a global significance for the study, it is worthwhile comparing this cline with variation in editing effort for all categories, i.e. in relation not only to ST-oriented editing but also to terminological and TT-oriented editing. In doing so, the dimension of editing effort will be established as the first component of translator performance to be observed against the background of the STs. More specifically, as variation in editing effort is mapped onto ST structural variation across tasks, this will give the opportunity to observe possible patterns of *co*-variation.

Figure 4.4 gives a visual representation of variation in editing effort across tasks (for each of the three editing categories) and compares this variation with the cline of increasing metaphoricity observed in the three STs.

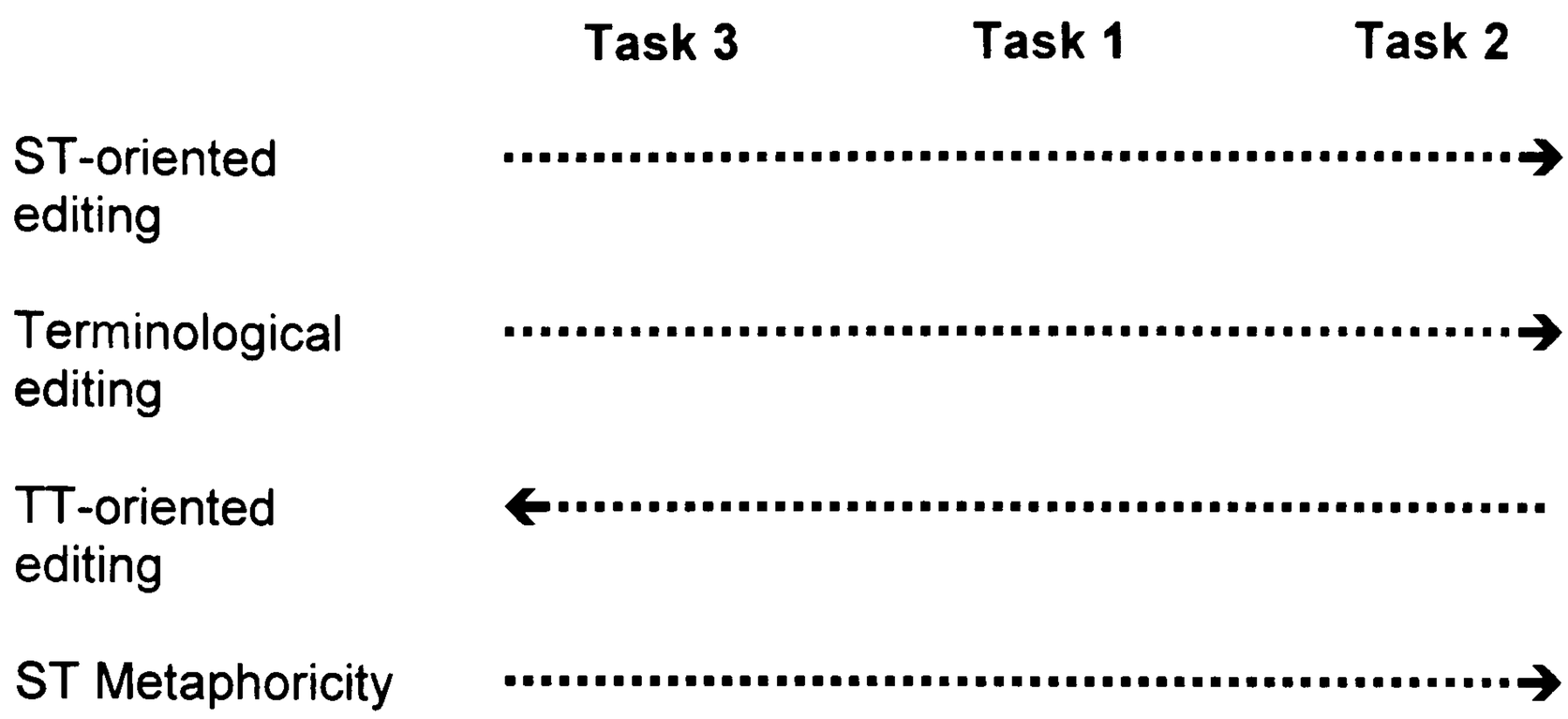

Figure 4.4 – Cross-task variation in editing, per category, against variation in ST metaphoricity (the direction of the arrow indicates an increase)

The order in which ST metaphoricity and all editing categories are ranked is the same: 3-1-2. What differs is the direction of change. In particular, the increase in metaphoricity is accompanied by a parallel increase in both ST-oriented and terminological editing. TT-oriented editing is in inverse relationship with the other editing categories and with ST metaphoricity, i.e. it decreases when metaphoricity and the other two categories of editing all increase. Structural variation in the STs is thus seen to be accompanied by parallel patterns of variation in editing effort for all the categories that have been identified. A pattern of co-variation can be said to obtain when the two dimensions, i.e. ST grammatical metaphoricity and editing effort, are compared. As far as editing effort is concerned, one feature, TT-editing, is the predominant aspect in all tasks but it is consistently seen to diminish as the other features of editing effort increase. Whether this is in any way influenced by the corresponding observed increase in ST metaphoricity remains to be seen and would have to be confirmed by looking at data coming from other sources, as will be done in the following chapters. Before moving on to the other data, however, a last look will be given at data on editing, this time to identify the 'location' of the effort. In other words, for each ST, those segments will be singled out on which translators were seen to perform the highest number of edits.

4.5 Mapping editing effort onto the STs

Data on real-time editing by translators can be subject to an analytical procedure that resembles very closely Choice Network Analysis (CNA; see 2.3). CNA was originally developed (cf. Campbell 2000) as a method for examining the *finished* translations produced by a number of translators working into the same target language and from the same source text, based on the assumption that the differences identified between these translations can be taken as evidence of mental processing by the translators. The first step in CNA, therefore, is that of identifying, in a set of translations of one ST, the spots where these translations are seen to differ the most. For each of these spots, the different renditions provided by translators are compared and inferences are drawn as to what led to one choice instead of another. The basis for the inferential step in the analysis is therefore *variation* across the TTs, so that, to put it simply, the more varied a set of renditions is, the more significant it is taken to be by the analyst in terms of cognitive effort. The analysis of TTs presented in the next chapter closely follows this analytical approach.

The idea of looking at variation in TT renditions, however, can also be applied to the data on editing that are being investigated in the present chapter. As explained in 4.3 above, when the typing process of each individual translator is observed, the interim renditions proposed by the translator for a given ST unit can relatively easily be identified. Editing can therefore be used to observe 'translator-internal' variation, i.e. the units for which a translator proposed more than one rendition before the final one, and this variation can be compared across translators, just as in CNA the final TT renditions are compared between each other. The reason for looking at editing in this way is that, if complemented with CNA on the target texts, an analysis of variation in the editing process can give a better approximation as to which units the translators really concentrated on in carrying out the task. These units are identified in the editing process, that is, in the TL, but are then mapped onto the ST segments to which they refer (segments here correspond to the groups and phrases identified, in the texts, through a Hallidayan constituency analysis).

The way edits were counted was illustrated in 4.3.2 above. An overall picture of edits was presented in 4.4.1, where figures were also given for the number of edits performed in each task by each individual translator. Table 4.5 below lists individual ST segments (i.e. groups or phrases) for which a majority of translators (i.e. 3 out 5

or more) provided interim renditions in the editing process (the segments are numbered according to the ST segmentation presented in Appendix 2). For each segment, an indication is given of the predominant orientation of the edits (ST, terminological or TT). 'Predominant' means that if more than two edits were observed for a given segment, the table only indicates the orientation of the majority of edits. In some cases, no predominant orientation could be identified, and therefore an indication of all types of edits observed for the segment in question is given.

Table 4.5 – ST segments on which editing effort was concentrated*

ST 1 (total edits: 140)				*ST 2 (total edits: 98)*				*ST 3 (total edits: 98)*			
Segm. no.	**ST-orient.**	**Term.**	**TT-orient.**	**Segm. no.**	**ST-orient.**	**Term.**	**TT-orient.**	**Segm. no.**	**ST-orient.**	**Term.**	**TT-orient.**
1	×	×		4	×	×		24		×	×
4		×		7	×			92			×
5		×		12		×					
23	×			22			×				
31			×	32	×						
35			×	35			×				
41		×		41			×				
50-52			×	55		×					
54	×										
63	×										
67			×								
Total (*f*)	4	4	4	**Total (*f*)**	3	3	3	**Total (*f*)**		1	2

* Segm. = Segment; orient. = oriented; Term. = Terminological

Absolute frequencies of editing interventions by the five translators on each of the tasks give an overall indication of editing orientation (see Tables 4.2, 4.3 and 4.4), but do not indicate where that editing was focused. So, for instance, while Task 1 precipitated more edits (140 edits) than either Task 2 or Task 3 (98 edits in each case), Task 3 had fewer editing 'hot spots' (2) than Task 2 (8), indicating that in Task 3 effort was sparser, regardless of its orientation. The CNA approach adopted in analysing editing thus tells us that the ST segments that required more effort of the translators are concentrated in Tasks 1 and 2.

When trying to map editing effort onto specific ST segments, it should also be remembered that the average segment length in the STs used in the three tasks is different. More specifically, the average length, in words, of a segment (i.e. a group

or phrase) in the STs is as follows: ST 1, 3.1 words; ST 2, 4.2 words; ST 3, 2.5 words. This means that many segments in Task 2 are double the length of segments in Task 3. Hence, the potential for edits on the same segment is higher in Task 2 than it is in Task 3. On the other hand, the greater length of segments in ST 1 and especially ST 2 is due to a general tendency, observed in both texts, towards the use of highly 'packaged' constructs, typically nominal groups with heavy pre- or post-modification, leading to more condensed structures. When considering how editing maps onto the three STs, a reasonable expectation would therefore seem to be that editing reflecting ST processing would concentrate on such highly packaged groups which is consistent with the trends observed here. Figure 4.5, giving the full text for each editing 'hot spot' in the ST, seems to provide some confirmation for this.

ST 1	ST 2	ST 3
1: [*t*]*he rate of crack growth* (term./ST) 4: *but also on the magnitude of the applied stress* (term.) 5: [*t*]*he development of a complete model for the kinetics of fracture* (term.) 23: *stress-free silicon-oxygen surface bonds* (ST) 31: *the application of stress* (TT) 35: *at speeds greater than one millimeter per second* (TT) 41: *focuses* (term.) 50-52: *until it reaches* (TT) 54: *within a few atomic dimensions of the tip* (ST) 63: *a silicate tetrahedron* (ST) 67: *the silicon atom in the*	4: *the mechanism of stress/environment dependent crack extension* (term./ST) 7: *the crack-tip bond rupture event* (ST) 12: *the rate of macroscopic crack extension* (term.) 22: *assume* (TT) 32: [*s*]*mall siloxane ring structures* (ST) 35: *the smaller* (TT) 41: *for different size siloxane ring structures* (TT) 55: *the rate of crack extension in silica glass in humid conditions* (term.)	24: *the size of the opening created when the ring structure is broken* (term./TT) 93: *and then* (TT)

center (TT)		

Figure 4.5 – Textual segments in the STs on which editing was concentrated in each task (with indication of editing orientation)

Segments for which editing was observed to be ST-oriented (1, 23, 54 and 63 in ST 1; 4, 7 and 32 in ST 2) can all be described as relatively long, quite densely packaged nominal groups (with the exception of segment 63 in ST 1, *a silicate tetrahedron*). Their interpretation by translators must not have been straightforward, and certainly their lack of familiarity with the subject matter must have provided an additional source of difficulty at the time of 'unpacking' the nominal groups. Unfamiliarity with the subject matter may have rendered problematic even the shorter segment *a silicate tetrahedron* (ST 1), a [N + N] compound term, as suggested, in particular, by the editing performed by two translators, MT and VM. MT's editing, as reconstructed from the Translog files, was as follows:

(4) (a) un tetraedro silicato
[a tetrahedron silicate]

(b) un tetraedro di silicato
[a tetrahedron of silicate]

where the uncertainty seems to be on the function of *silicate* as a part of speech. The first TL rendition interprets it as an adjective (which it is not; note that "silicato" is not used as an adjective in Italian either), the second correctly gives it as a noun. The same uncertainty seems to emerge from VM's editing:

(5) (a) un tetraedro di silice
[a tetraehedron of silica]

(b) un tetraedro che f [sic]
[a tetrahedron that ...]

(c) un tetraedro contenuto in un silicato
[a tetrahedron contained in a silicate]

where *silicate* is at first (see 5a) taken to be the adjectival form for *silica* (which, again, is not the case; in this particular ST segment, *silicate* does a have a modifying function, though).

Some other segments can be dwelt upon in more detail, not only among those for which predominantly ST-oriented editing was observed but also among segments

for which translators performed other types of edits. Based on the assumption that editing concentration suggests an increase in processing effort by the translators. an attempt can be made at interpreting this increase as a function of the lexico-grammatical features shared by the majority of ST segments so identified (many of these segments are highly nominalised constructs).

In particular, even where translators were seen to perform editing on other aspects of the segments, such as terminology, a case could perhaps be made for interpreting the difficulty that this editing reflects in the light of the Hallidayan "overall effect" of wording (i.e. words plus structures) that was mentioned earlier in this chapter (see 4.3.1). In other words, the uncertainty that translators are observed to experience when faced with words such as *rate* (appearing not just once but three times in Figure 4.5; see segment 1 in ST 1 and segments 12 and 55 in ST 2), *kinetics* or *fracture* (both in segment 5, ST 1) can be discussed with reference to their possible characterisation as extreme cases of 'grammatical metaphorisation', meaning that in these words the process of increasing abstraction that leads from more to less "congruent" (i.e. from verb-based to noun-based) meaning realisations is observed at its full potential. Nouns indicating processes or abstract relations are often the result of such metaphorisation. *Fracture*, for instance, is a more abstract, less "congruent" way to indicate either the act of breaking (the process) or the state of being broken (or even a specific split or division in a material) - an instance of polysemy that may represent a further source of difficulty for someone who is not familiar with the subject matter of the text.

The noun *rate* is, in this respect, extremely interesting. Defined in the CED as "a quantity or amount considered in relation to or measured against another quantity or amount", it is in itself so abstract that trying to find a more "congruent" realisation for it in clausal form appears to be an all but impossible task. For a translator faced with words such as these, deciding what the relevant meaning is could well be a first hurdle in itself (the CED has 9 meanings for *rate* as a noun, most of them not too distant from the one quoted above). And the difficulty can be compounded, once again, by the lack of familiarity with the subject matter.

The steps taken in the next chapters will complement the study of editing presented here. The final aim is to develop an analytical framework capable of both

closing in on the 'hot spots' in the STs and providing firmer support for discussing and interpreting those hot spots in terms of the difficulty they posed for translators.

Chapter 5
ANALYSING TARGET TEXTS FOR CLUES ON SOURCE-TEXT DIFFICULTY

5.1 Introduction

The studies in which Stuart Campbell has, occasionally together with Sandra Hale, developed, discussed and applied Choice Network Analysis (CNA) have been reviewed in Chapter 2 (see 2.3). CNA is an attempt at using clues in the translations by multiple subjects in order to draw inferences about the processes that typically operate "in particular types of subjects translating particular texts under specific conditions" (Campbell 2000: 31). In particular, by looking at the different target texts they have produced for the same source text, CNA is intended to investigate the mental processes of translators. It is presented by Campbell as a complement or an alternative to research methods based on introspection, such as think-aloud protocols and word-based experimental techniques. In the context of the present work, it complements data obtained from both keystroke logs and an analysis of errors.

The principles CNA is based on are as follows (cf. Campbell 2000: 32): 1) target texts are a tangible source of evidence of mental processing in translation; 2) the texts produced by a sample of subjects translating the same texts will reveal a range of differences and similarities in the behaviours of the subjects; 3) the larger the sample, the more likely it is that the complete range of behaviours is approached; 4) a model of the mental processing underlying the translation of a given text in a given language combination can be inferred by comparing and classifying the behaviours of the subjects; 5) the analyses of specific texts and language combinations can help in extracting general principles about mental processing, and these principles can be used as hypotheses for investigating other texts and language combinations.

Of these principles, two are worth discussing in more detail in relation to the present work. One is the possibility to extract general considerations about mental processing from the analyses of specific texts and language combinations. Given the modest size of the sample employed in this study, it should be pointed out that whatever findings are arrived at (both in the CNA and in the other analytical steps)

they should only be seen as relevant for the specific group of translators that participated in the study. Although they cannot be characterised as "general principles of mental processing", such findings could be used as hypotheses for investigating other texts and language combinations, possibly with larger samples, as Campbell suggests. In the present context, findings obtained through CNA will be discussed in relation to findings obtained from other sources of data (i.e the data on editing presented in Chapter 4 and the Error Analysis presented in Chapter 6).

The other principle of CNA worth briefly elaborating on a this stage is the possibility of using ever larger samples so that the complete range of behaviours on the part of translators is approached. As previously acknowledged, the present study is based on a modest sample and therefore corrective measures had to be taken to compensate for this limitation. More specifically, the decision was taken to complement CNA with data from other tangible sources of evidence, and keystroke logs were judged to provide a tangible enough source of evidence. The data obtained for keystroke logs, were, in addition, amenable to an analytical method which closely resembles that of CNA. Just as CNA is based on inference "from the analysis of translations of the same source texts by multiple subjects" (Campbell 2000: 30), so keystroke logging was deemed capable of providing data from which the behaviour of a single individual translator could be inferred (cf. 4.3 to see how this was done).

In what follows, I will first illustrate the criteria that guided the analysis of the TTs in terms of the network of choices made by the five translators participating in the study and then propose an overall picture of each of the three tasks in terms of variation of choices across translators.

5.2 Analysing choice networks

CNA is presented by Campbell as being theory-free: "[w]hat it generates are models, i.e. representations of complex systems" (2000: 38). The starting point for building a network of choice nodes can vary depending on the specific question the researcher wants to investigate. For example, one might want to look at the translation of ST metaphors, in which case the network will reflect the strategies employed by translators to deal with these (e.g. reduction to sense, preservation or modification). Another possibility would be to look at ST clauses containing passive verbs: in this case the networks will emerge from the different options chosen by translators as

regards the TL verb forms that translate these SL passives. Whatever the ST element the researcher decides to concentrate on, CNA only provides a 'template' for constructing a model of mental processing. The researcher is free to make the models generated through CNA more or less complex and to incorporate theory in a network. In the present context, the theoretical basis for the analysis is provided by the Hallidayan description of scientific discourse (see 2.5.4).

A clarification regarding terminology (cf. Hale and Campbell 2002: 18) will also serve to illustrate more clearly how a network is constructed. Based on a given ST element, a set of *options* is identified for the translation of that element; this set is arrived at empirically by analysing the translations provided by a group of subjects. An *alternative rendition* (or simply a rendition) is a selection made from this set of options. The term *choice* is the overarching term that covers the process whereby translators select alternative renditions from options. The inferential step implied in CNA is brought into the picture at the time of characterising what exactly this "choice" is based on or how it can be more profitably described in terms of the research question one is proceeding from. To return to an example given above, if the researcher is building a network to see how different translators treat ST metaphors, "choice" can be characterised in terms of the selection from a set of strategic options, so that translators are observed to, say, reduce metaphors to sense, preserve them or modify them. In such a scenario, the researcher may or may not be interested in looking at the accuracy of the TL renditions, whereas in other scenarios (e.g. the translation of complex noun phrases) a consideration of accuracy may have a direct bearing on how choice is characterised (because, say, the researcher wants to see whether such elements are problematic or not).

In the present context, the inferential step taken in interpreting the choice networks takes into consideration the continuum implied by the research question, which sees structure at one pole and lexis at the other. Alternative renditions for a given ST element are, in other words, compared with each other and described as varying at either the structural or the lexical level (or both). A subsequent, interpretative step tries to relate such variation to the lexico-grammatical features of the ST element in question. Before looking at how this was done in more detail, I will first illustrate how analytical units were identified so that choice networks could be constructed and how TL renditions for a given ST segment were differentiated.

5.2.1 Identifying units for analysis

In performing Choice Network Analysis on the translations provided by the five participants in the study, TT units were identified on the basis of the segmented STs, in an attempt to identify "coupled pairs of target- and source-text segments" (Toury 1995: 89; the criteria for ST segmentation are illustrated at length in Appendix 2). This would seem to run counter to Toury's own suggestion that, in a retrospective study, no unit should be postulated in either the ST or the TT; rather, units should be established ad hoc, with the pairing of segments subject by a "heuristic principle" whereby the segments defined themselves simultaneously, "determining each other" (Toury 1995: 89). In the present context, however, as the research question focuses on the way lexico-grammatical features of the ST influence the degree of difficulty of a given translation task, data analysis is carried with constant reference to such features. The emphasis on ST features implies that, although different sources of data are considered and data themselves may be submitted to different analytical procedures, the initial step for analysis is always a consideration of the lexico-grammatical configuration of the STs. Hence the decision to postulate units for analysis. Heuristic methods, however, were not discarded altogether, and indeed were employed at various stages of the analysis, e.g. at the moment of establishing what constitutes variation in the different TTs proposed by the translators.

In line with my hypothesis (which is that grammatical metaphor and its most immediate manifestation, i.e. nominalisation, contribute to an increase in the degree of difficulty of an ST), an obvious decision would have been to identify grammatical metaphors in the STs and then construct choice networks based on those. These networks might have been interpreted in terms of what led to specific renditions instead of others, and each individual rendition might have been checked for accuracy. Focussing on grammatically metaphorical units alone, however, would have left out a possibly important comparative element. By only looking at the units that were assumed to be significant, other units might have been ignored that could actually be as significant as the former in terms of the difficulty they pose to translators. Therefore, instead of isolating specific ST units, I opted for a segmentation of the STs carried out according to relevant, Hallidayan-based criteria and, from that, proceeded to construct choice networks that would than be interpreted and compared both within and across the three translation tasks under investigation.

5.2.2 Differentiating between TL renditions

The first step in CNA is that of looking at how an ST element is rendered in the TL by a group of translators. The renditions provided by translators are compared to one another and a description is given of what distinguishes them. This description may be made from a number of perspectives, depending on what the researchers is specifically interested in (as was seen earlier, one possibility could be that of making inferences as to what led to one choice instead of the others). The basis for this descriptive or more overtly inferential step in the analysis is therefore *variation* across the TTs, so that the more varied a set of renditions is, the more complex the set of choices available to the translators is assumed to be. Such choices can then be interpreted by the researcher according to the specific aspect s/he is interested in.[1]

In order to obtain a picture of variation, TT units corresponding to ST segments were here analysed using a 'type/token' descriptive approach. Each individual TT unit corresponding to an ST segment was considered as a *token*, while renditions were grouped into *types* according to criteria established with reference to two distinct levels of linguistic description. One is the syntagmatic or structural level, i.e. the level concerned with how words (or other units of linguistic description) are arranged. The other is the paradigmatic level, i.e. the level of word choice. Different renditions of the same ST segment were compared to one another and a resulting type/token ratio was obtained for each level. Segments for which the type/token ratio was higher were considered as having a higher degree of TL variation. The TL renditions of a given segment would therefore be seen to yield a type/token ratio of, say, 2/5 at the structural level and 3/5 at the paradigmatic level (referred to as the "lexical" level from now on).

In the following sections, some examples will be discussed to illustrate how the target texts were analysed with a view to distinguishing between the renditions provided by each of the five translators. In particular, section 5.2.2.1 illustrates the criteria that were adhered to in analysing structural variation between renditions, while section 5.2.2.2 explains how TT renditions were distinguished at the lexical

[1] Campbell (2000: 38) himself suggests that "the complexity of choices available to the translator to select from" can be used as a measure for "the relative difficulty of parts of source texts" (cf. also Campbell 1999).

level. On some occasions, the analysis of variation across TT renditions relied on criteria established on a case-by-case basis; relevant examples are given in section 5.2.2.3. The analysis illustrated in these sections is the followed by a quantitative overview (5.3.) of how variation is distributed in the three tasks.

5.2.2.1 Variation across TTs at the structural level

At the structural level, the criteria used to distinguish between renditions reflect those that were taken into consideration for segmenting the STs. TL renditions were thus subject to a constituency analysis aimed at identifying units comparable to those identified in the STs. At the group/phrase rank, the closest comparable unit in Italian traditions of language description is the *sintagma*, defined as a grouping of words that acts as unit within the clause (cf. Serianni 1988: 89; Dardano 1996: 316; Bonani *et al.* 2003: 120).[2] Different classes of *sintagma* are normally identified (cf. Renzi *et al.* 1991):

- *sintagma nominale*, e.g. "la velocità di frattura" ("the rate of fracture");[3]
- *sintagma verbale*, e.g. "ci si avvicina" ("one gets closer to")
- *sintagma preposizionale*, e.g. "con una molecola d'acqua" ("with a water molecule")
- *sintagma aggettivale*, e.g. "più reattivo" ("more reactive")
- *sintagma avverbiale*, e.g. "molto lentamente" ("very slowly")

In the present context, TL renditions of the same ST segment at group/phrase rank were analysed according to these categories of *sintagma* (but from now on the term "phrase" will be used) with the purpose of discriminating one rendition from another. (More finely-grained analyses of the Italian texts were felt to go beyond the scope of the study.) In particular, any two TL renditions of the same ST segment were judged to be different from each other when one of the following conditions was met:

---

[2] The Italian term *sintagma* was introduced in the language as a translation for the Saussurean term *syntagme*; it is often used as an equivalent for the term "phrase" as used in Anglo-Saxon structuralist traditions (cf. Beccaria 2004: 702-703). Note that, strictly speaking, a *sintagma* is usually taken to indicate a unit "deriving from the combination of two or more lexical or grammatical items" (Dardano 1996: 316; my translation). This means that a one-word *avverbio* (adverb) would not coincide with a *sintagma avverbiale*, which consists of at least two words (e.g. "più frequentemente"; En.: "more frequently"). For the present purposes, however, this distinction is not taken into account: an adverb and an adverbial phrase are considered as the same type of unit.

[3] All examples in Italian are taken from the translations obtained for the study; a back-translation into English is given for each example.

1) the two renditions belonged to different classes of phrase;
2) the two renditions belonged to the same class of phrase but their internal structure was considerably different.

Let me now illustrate some examples relating to each of the conditions.

Segment 20 in ST 1 is a non-finite verbal group (*to dissolve*) appearing in the context reproduced in Figure 5.1 below, where the renditions by the five translators are also given. (Note that, in terms of constituents, the ST reproduced in the figure spans across two clauses; the segment I am focussing on and its translations are underlined.)

The dissociative reaction [...] causes the surface of silica glass to dissolve in water at a rate of...
CB: La reazione dissociativa [...] porta la superficie del vetro di silice a dissolversi nell'acqua secondo un rapporto di...
MS: La reazione dissociativa [...] provoca lo scioglimento della superficie del vetro di silice nell'acqua ad una velocità di...
MT: La reazione dissociativa [...] provoca lo scioglimento della superficie del vetro di quarzo in acqua ad una velocità di...
SC: La reazione dissociativa [...] fa sì che la superficie di un vetro di silice si sciolga in acqua alla velocità di...
VM: Tramite le reazioni [...] , la superfice del vetro di silice si scioglie in acqua alla velocità di...

Figure 5.1 – Translations of segments 17-22 in ST 1

In terms of a type/token description, the translations of *to dissolve* provided by the five translators participating in the study can be said to correspond to three different types, thus leading to a ratio of 3/5. The first type is the non-finite verbal group in CB's rendition, *a dissolversi*. The second type is the nominal group *lo scioglimento*, proposed by both MS and MT. Finally, the third type is the finite verbal group proposed by SC and VM. In particular, SC has the verb in the subjunctive (*si sciolga*), while VM has an indicative present tense (*si scioglie*); the difference in mood, however, is not taken into account in the analysis and the two verbs are described as expressing the same type of TL rendition. To be sure, this particular example already illustrates how certain decisions taken in differentiating between renditions contain a discretionary element, in this case associated with the decision to

differentiate between the finite and non-finite verbal groups (which were felt to bring about fundamental structural differences between the clauses that contained them); other examples of ad hoc decisions that were taken in differentiating between translations will be discussed in 5.3.2.3 below.

Figure 5.2 gives an example of renditions that were all differentiated according to the second condition indicated above, i.e. on the basis of the internal structure of phrases.

the strain-enhanced chemical reactivity of silicate materials
CB: la reattività chimica che aumenta la deformazione dei materiali di silicato
MS: la reattività chimica intensificata dalla deformazione dei materiali silicati
MT: la reattività chimica aumentata dalla deformazione dei materiali a base di silicato
SC: la reattività chimica dei silicati aumentata in seguito a una sollecitazione
VM: la reattività chimica aumentata dalla deformazione dei silicati

Figure 5.2 – Translations of segment 29 in ST 2

All TL renditions of the ST segment can be described as a nominal group. In particular, MS, MT and VM have an embedded clause introduced by the verb that translates *enhanced* (*intensificata*, *aumentata* and *aumentata*, respectively). Their renditions were judged to realise the same type. CB's translation was considered a different type, in that the embedded clause is introduced by a finite verb, and so was SC's rendition, where the specification (*dei silicati*, translating *of silicate materials*) is moved immediately after the head of the nominal group. The internal structure of the group varies here. The type/token ratio for this segment was therefore established at 3/5.

A higher type/token ratio was found, structurally, in the translations shown in Figure 5.3, which were judged to realise four different types. In translating this segment all translators opted for an expansion. The rendition which is structurally more similar to the ST segment is MS's, the expansion here being *un qualche tipo di* (En.: "some sort of") preceding the equivalent for *stress*. In the case of CB and SC, the expansion consisted in choosing a noun group including an infinitive (*il fatto di esercitare*) with the equivalent of *stress* (*pressione* for CB, *tensione* for SC) becoming the object of this infinitive. Given the presence of the verb, CB's and SC's translations were judged to be a different type from MS's rendition (based on

condition 2 indicated above). The third type is MT's rendition, which translates *application* with a non-finite verb (*esercitando*), thus turning the unit into a dependent clause. VM's translation is also a dependent clause, this time a conditional introduced by *qualora* ("if; in case"), but with a finite verb (*venga applicata*, a passive). MT' and VM' rendition are different from each other according to condition 2; at the same time, they are different from the other three renditions according to condition 1.

the application of stress
CB: il fatto di esercitare pressione
MS: l'applicazione di un qualche tipo di stress
MT: esercitando una certa pressione
SC: il fatto di esercitare tensione
VM: qualora venga applicata tensione

Figure 5.3 – Translations of segment 31 in ST 1

In analysing the TTs for variation at the structural level, not all ST segments could be mapped individually onto TT units. In a few instances, some translators opted for TL solutions which could be seen as spanning more than one ST segment; to differentiate between renditions, joint segments were therefore considered, as illustrated in Figure 5.4 below, where renditions by MT and VM 'merge' the two ST segments into one verb.

are so unreactive
CB: non presentano nessuna reazione
MS: risultano talmente solidi
MT: non reagiscono
SC: reagiscono talmente poco
VM: non reagiscono

Figure 5.4 – Translations of segments 24-25 in ST 1

In such cases, the boundaries between ST segments were considered as a single segment, a fact whose relevance will be more apparent at the moment of presenting an overall picture of inter-translator variation for each task (see 5.3).

5.2.2.2 Variation across TTs at the lexical level

At the lexical level, the analysis of variation across translators took into account the content words found in the various renditions for a given ST segment. Establishing the extent to which renditions were different from one another in terms of word choice was more straightforward with shorter segments (typically verbal groups or adverbial phrases), as such segments contained fewer content words. More problematic was the differentiation between longer segments (especially the more densely packed nominal groups or segments containing embedded clauses), where there was a higher concentration of content words. In such cases, segments were sometimes further split into constituents and the comparison was made on the basis of these. Examples that illustrate this problem and discuss the solutions that were adopted to solve it are given in the following section. In this section I will briefly present some examples of lexical variation observed in the TL renditions of shorter segments.

ST verbal groups are among the elements for which lexical variation in TL renditions could more easily be observed. Consider for instance the translations of the verb *occurs* (segment 63, ST 3) reproduced in Figure 5.5, where only two translators are seen to give the same equivalent for the verb. The figure gives the wider context for both the ST and the TTs (segment 63 is omitted as it is not relevant to the example being discussed). As regards the renditions for *occurs*, the type/token ratio is 4/5.

Stress [...] occurs when...
CB: La pressione [...] ha luogo quando...
MS: La tensione [...] ha i suoi effetti...
MT: La tensione [...] ha luogo...
SC: le sollecitazioni [...] si riscontrano quando...
VM: La tensione [...] si ha quando...

Figure 5.5 – Translations of segments 62 to 65 in ST 3 (segment 63 omitted)

In considering lexical variation, the morphological difference between words was disregarded. In other words, TL content words were considered similar or different on the basis of their lexical root. To go back to an earlier example. in Figure 5.1 the verb *to dissolve* has, lexically, only two different equivalents in the set of TL

renditions: *dissolversi* (verb) is one, *sciogliere/scioglimento* (verb and noun respectively) is the other. The type/token ratio is thus 2/5.

5.2.2.3 Examples of variation identified on a case-by-case basis

The examples discussed in the previous two sections were among the more straightforward cases of differentiation between TL renditions. For a number of segments, however, the analysis of renditions at either the structural or lexical level went beyond the criteria illustrated above and was conducted on a case-by-case basis, thus reflecting "a process of trial and error" (Campbell 1998: 110) or, in Toury's (1995: 89) terms, the operation of a "heuristic principle" supported, but not constrained, by a characterisation of the texts in terms of constituency.

It would sometimes be the case, for example, that the TL unit under consideration consisted of a rather dense or long stretch of text (say because it was the equivalent of an ST segment containing a complex nominal group or an embedded clause). In such cases, TT renditions were very likely to present a high degree of mutual variation deriving precisely from their length (the more words an ST segment contains, the more varied its possible TL equivalents can be expected to be). Internal portions of these segments, however, often turned out to be quite similar in terms of word choice. Whether in such cases the renditions could be seen as fundamentally different or fundamentally similar was often decided on an ad hoc basis.

Figure 5.6 provides an example of such a long ST segment (a nominal group with a particularly high density of content words) and describes the procedure whereby an overall type/token ratio was assigned to the segment at both the structural and the lexical level. Similar procedures were adopted for the other segments sharing the same features in all three translation tasks.

the mechanism of stress/environment dependent crack extension
CB: al meccanismo riguardante l'estensione della fenditura che dipende dal rapporto pressione/ambiente
MS: il meccanismo di dipendenza dell' estensione di una frattura rispetto allo sforzo esercitato su di essa e rispetto all' ambiente
MT: il meccanismo di propagazione della cricca dipendente dal rapporto sforzo/ambiente
SC: il meccanismo dell'estensione dell'incrinatura che dipende da condizioni

ambientali o di sollecitazione
VM: il meccanismo di estensione della cricca dipendente dalla tensione o dall'ambiente

Figure 5.6 – Translations of segment 4 in ST 2

Structurally, all TL segments consist of a nominal group with "il meccanismo" as Head, but they differ in the way the rest of the nominal group is constructed. CB, MT, SC and VM all follow the Head with a series of embedded structures: a nominal group ("riguardante l'estensione della fenditura", "di propagazione della cricca", "dell'estensione dell'incrinatura" and "di estensione della cricca") and an embedded relative clause (introduced by "che" in CB and SC and by a non-finite verb, "dipendente", in MT and VM). These four TL renditions, therefore, have a similar overall structure, which can be represented as follows:

Head + embedded nominal group + embedded clause

Note, however, that within the embedded clause the four renditions diverge; in particular, CB and MT have one prepositional phrase ("dal rapporto pressione/ambiente" and "dal rapporto sfrozo/ambiente"; note that "pressione/ambiente" and "sforzo/ambiente" could be seen in itself as elliptical prepositional phrases, whose full-length equivalent would be "tra pressione e ambiente" and "tra sforzo e ambiente"). SC and VM, on the other hand, opt for two prepositional phrases joined by a disjunction ("o"; En.: "or"). Based on the difference between the embedded clause, CB's and MT's translation can be seen to correspond to one structural type and SC's and VM's to another.

MS's rendition of this segment is structurally different from the others in that it contains more than one ranking group, as can be seen by observing the segment in the wider context of the clause (underlining identifies the segment in question):

> Alcuni modelli fenomenologici sono stati sviluppati con lo scopo di studiare il meccanismo di dipendenza dell' estensione di una frattura rispetto allo sforzo esercitato su di essa e rispetto all' ambiente.[4]

If bracketing were performed on this clause, one possibility would be:

[4] Cf. the ST: *Phenomenological models have been developed to address the mechanism of stress/environment dependent crack extension.*

(Alcuni modelli fenomenologici) (sono stati sviluppati) (con lo scopo di) (studiare) (il meccanismo di dipendenza dell'estensione di una frattura) (rispetto allo sforzo esercitato su di essa) (e rispetto all' ambiente.)

Note that the number of constituents is greater here than that in the other TL renditions; in particular, the nominal group *il meccanismo di dipendenza dell'estensione di una frattura* is followed by two other, coordinated ranking groups: *rispetto allo sforzo*...and *rispetto all'ambiente*, this being the main difference between this translation and the others discussed above (note, also, the different internal structure of the nominal group itself). At the structural level, therefore, the five renditions of the segment can be seen to correspond to three different types, bringing the type/token ratio for this segment to 3/5.

Lexically, characterising variation across the five renditions for this segment is perhaps more problematic, as the ST segment represents a particularly dense stretch of text. Out of the 8 words making up the nominal group, 6 are content words (*mechanism*, *stress*, *environment*, *dependent*, *crack*, *extension*). Of these, 2 are certainly terms (*stress*, *crack*) and another is also likely to be classified as such (*environment*). With such a high proportion of content words, a high degree of variation in the TL renditions is perhaps to be expected, as lexical items are likely to have various accurate TL equivalents. Terms could be an exception in this regard, in that they are more likely to have only one accurate equivalent, but variation concerns them too, here – perhaps reflecting the translators' unfamiliarity with the domain. On the other hand, the five TL renditions do show some degree of similarity on portions of the segment. Restructuring the segment without altering the overall meaning, such differences and similarities may be seen to emerge more clearly; the ST segment, i.e.

the mechanism of stress/environment dependent crack extension

could be rewritten as

the **mechanism of crack extension** (which is) **dependent** on (the relationship between) **stress and environment**

so as to obtain three more clearly defined chunks (highlighted in bold). As regards the first chunk, TL word choice for *the mechanism of* [...] *crack extension* is quite consistent across the five translators (all have *meccanismo* for *mechanism* and four have *estensione* for *extension*, the exception being MT's choice of *propagazione*), while renditions for *crack* present more variation (*fenditura*, *frattura*,

incrinatura and, twice, *cricca*). The chunk could be treated as if it was one self-contained segment and, on the basis of the different renditions for *crack*, a type/token ratio of 4/5 can be assigned to it. The second chunk (*dependent*) is the only one for which translators opted for consistent lexical choices: all TL renditions derive from the same root (*dipend-*, realised as either a verb or an adjective). The third chunk shows, again, considerable variation, in particular as far as *stress* is concerned. The type/token ratio for this chunk is 4/5, since the following different equivalents are proposed: *pressione*, *sforzo* (twice), *sollecitazione* and *tensione*. Note that variation on this chunk is distributed differently from variation on the first chunk (the one including crack). In other words, no two translators consistently provided the same TL equivalents. To conclude, based on the high degree of variation observed for two out of the three chunks the segment could be split into, an overall type/token ratio of 5/5 was assigned to the segment in its entirety.

Another example of how TL renditions had to be differentiated ad hoc is provided by the translations reproduced in Figure 5.7, all of them nominal groups. In particular, in order to identify what may be called the internal structural profile of each rendition, attention was paid to the way translators had interpreted the noun phrase *macroscopic crack extension*. Some considered *macroscopic* as a qualifier for *extension* (see VM's translation); others saw it more closely associated with *crack* (see MT and SC). CB's translation does not make it clear whether it is one way or the other (if anything, *macroscopico* in her rendition refers to what is ostensibly the equivalent for *rate*, i.e. *livello*) and can therefore be considered as yet another different type of rendition. As a result, the renditions for this segment were assigned a type/token ratio of 3/5.

the rate of macroscopic crack extension
CB: il livello macroscopico della fenditura
MS: la velocità di estensione della frattura macroscopica
MT: il tasso di propagazione della cricca macroscopica
SC: l'estensione della incrinatura macroscopica
VM: il tasso dell'estensione macroscopica della cricca.

Figure 5.7 – Translations of segment 12 in ST 2

Other problematic cases of differentiation between TT renditions emerged where translators heavily restructured, in their TTs, portions of the ST spanning more

than one segment, leading to sequences of TT renditions ordered differently from their ST counterparts.[5] In such cases, at the moment of establishing a one-to-one correspondence between segments (so as to measure cross-translator variation) TT segments had to be re-ordered and mapped onto the sequence of ST segments.[6] Moreover, in the process of restructuring their target text, translators often ended up 'omitting' a given segment, whose content was incorporated in that of other segments. The ST clause complex and its translations reproduced, segmented, in Figure 5.8 below provide examples of both reordered segments and 'omitted' segments in the TTs. In particular, for the stretch of text spanning segments 16 to 19, all TTs present renditions that, in one way or other, depart from the sequence of segments observed in the ST (this is reflected, in the figure, by the TT cells having different borders than in the ST).

<table>
<tr><th colspan="2" rowspan="2">Source Text (segmented)</th><th colspan="5">Translations</th></tr>
<tr><th>CB</th><th>MS</th><th>MT</th><th>SC</th><th>VM</th></tr>
<tr><td>13</td><td>This view of the fracture process</td><td>Questa considerazione del processo di frattura</td><td>Questa interpretazione del processo di frattura</td><td>Questa analisi del processo di frattura</td><td>Questo modo di intendere il processo di fratturazione</td><td>Questa visione del processo di frattura</td></tr>
<tr><td>14</td><td>requires that</td><td>richiede che</td><td>vuole che</td><td>implica che</td><td>presuppone che</td><td>richiede</td></tr>
<tr><td>15</td><td>the rate of chemical reaction at the crack tip</td><td>il livello della reazione chimica all'estremità della frattura</td><td>la velocità di reazione chimica nel punto di frattura</td><td>il tasso di reazione chimica in corrispondenza dell'apice della cricca</td><td>la quantità di reazioni chimiche all'estremità dell'incrinatura</td><td rowspan="4">un aumento del tasso di reazione chimica sulla punta della cricca</td></tr>
<tr><td>16</td><td>increase</td><td rowspan="4">sia direttamente proporzionale alla pressione esercitata sul legame fenditura-estremità</td><td rowspan="4">aumenti proporzionalmente allo sforzo esercitato sul legame chimico presente nel punto di frattura stesso</td><td rowspan="4">cresce all'aumentare dello sforzo sul legame cricca-apice</td><td rowspan="4">diventi maggiore con l'aumentare della sollecitazione sui legami all'estremità dell'incrinatura</td></tr>
<tr><td>17</td><td>as</td><td rowspan="3">all'aumentare della tensione sul legame della punta della cricca</td></tr>
<tr><td>18</td><td>the stress on the crack-tip bond</td></tr>
<tr><td>19</td><td>increases.</td></tr>
</table>

Figure 5.8 – Translations of segments 13 to 19 in Task 2

[5] An example in which TT renditions could be seen to span more than one ST segment was already discussed (see Figure 5.2). In that case, however, the relevant strecth of ST segments was short and there was no reordering in the sequence of TT renditions.

[6] Needless to say, the establishment of one-to-one correspondences in observing a ST and its translation is more often than not an ideal abstraction and largely depends on the criteria used to identify units of translation (here, a grammatical analysis of the ST).

Trying to neatly map TT elements onto ST segments for this clause complex, and for segments 16 to 19 in particular, is not a straightforward task. For most translators, some TT renditions can only be made to match a *grouping* of ST segments, of which they act as an equivalent. Note, for instance, how in CB's translation the one verbal group *sia direttamente proporzionale* is provided as the equivalent for a grouping of three non-adjacent segments in the ST: *increase as ... increases* (the same happens in MS's translation: *aumenti proporzionalmente*). The apparent effect, in this case, is that segment 19 is omitted in the translations, where in reality its content is incorporated elsewhere in the text. An ostensible equivalent for segment 19 (the infinitive *aumentare*) is present in MT and SC's translations, but it is relocated: *cresce all'aumentare* and *diventi maggiore con l'aumentare*. VM's translation is still different in that both segment 16 and segment 19 are relocated, with the former turned into a nominal group (*un aumento del...*). For all translators, then, the "coupled pair" (Toury 1995) of ST and TT segments spans a number of different segments originally identified in the ST.

Cases of structural mismatch between ST and TT such as this were, admittedly, rather infrequent in the translations investigated for the present study. In other words, translators mostly avoided departures from the order of segments they found in the STs. The need to realign the texts so as to check what portions of a given TT could be seen to match with a given ST segment (or grouping of segments) will become more apparent later on the analysis (5.3), when the procedure is illustrated of how cross-translator variation was measured in the three tasks.

5.2.3 Summing up

Before moving on to consider an overall picture of cross-translator variation for each of the three tasks, let me briefly sum up the main points which emerged in this description of how the basic analytical procedure provided by CNA was applied.

CNA is intended to study variation in the way a group of translators render, in the TL, a given ST element. The ST elements referred to in this study correspond to the units identified thorough an analysis of the source texts at the lexico-grammatical level. In line with the research question, such an analysis was meant to describe the texts according to formal criteria. More specifically, it was meant to segment the text into constituents at different ranks of the lexico-grammatical stratum. Variation in

the TTs was then analysed, through a CNA approach, with reference to such ST segments. In particular, variation was characterised on the basis of the two converging perspectives of structure and lexis. Occasional adjustments had to be made in matching "coupled pairs" (Toury 1995) of ST and TT segments, since – for reasons that should be clear – the structural profile of the segments in the two languages was not always comparable. Frequent ad hoc decisions were also taken in analysing how the various TL renditions for the same ST segment differed from each other. To ensure consistency, however, such decisions were always taken on the basis of lexico-grammatical criteria. Note, also, that cross-translator variation observed for a given TT unit at one level (say, structure) did not exclude the possibility that the same unit varied at the other level as well. Some sets of TT renditions, in other words, were seen to vary for both structure and lexis, albeit at different degrees (i.e. with different type/token ratios).

The CNA approach adopted here could perhaps be described as 'lightweight' if compared to the way CNA was deployed by Campbell himself. In Campbell's application of CNA there is a more explicit attempt at using it to uncover the (inferred) set of strategic options employed by a group of translators for the rendition of a given unit. CNA, in other words, is attributed considerable explanatory power in terms of the motivation behind the choice made by a translator for a specific TT option (the extent to which the translator is *aware* of such choice is a question that CNA alone can probably not investigate; for examples of choice networks taken from Campbell's studies see 2.3). In the present context CNA is used essentially as a method for representing variation across translations (expressed as a 'type/token' ratio), the more genuinely inferential, interpretative step being deferred to a later stage in the investigation, i.e. after data from the three different sources have been integrated.

Note, finally, that no evaluative element has been introduced yet to describe the translations. As explained in 2.3.1, when CNA is used to study text difficulty in translation, the accuracy of TL renditions also needs to be considered. It was felt, however, that the introduction of an evaluative dimension at this stage would be premature. Translations will be considered for their accuracy later on in the study (Chapter 6), and this will be used to give further support to the identification of ST 'hot spots', thus complementing both the study of editing and the findings arrived at through CNA.

5.3 An overall picture of variation in the three translation tasks

Data obtained through a comparison of the target texts along the lines described in the previous section can be submitted to various analytical procedures with the aim of building up an overall picture of variation. According to the way data are looked at, variation can be seen to relate to different aspects of the tasks, such as the incidence of segments resulting in a high number of different TT renditions or the (dis)similarity among renditions.

Observing different renditions provided by a group of translators for a given ST unit is the starting point of CNA and can serve as a basis for quantitative analyses: "[b]y comparing each translator's version of each word with every other translator's version it is possible to build up a purely mathematical picture of the similarity among them" (Campbell 1998: 103-104).[7] Needless to say, such an approach works best when the sample is large (say, at least 15-20 translators). Nevertheless, even with a small sample such as the one used for the present study, it is felt that a quantitative picture of variation is, if not amenable to generalisations, at least capable of giving insights to be confirmed by looking at the other data gathered on the three translation tasks.

A quantitative measure of the similarity among individual translators' renditions is obtained, applying a procedure adopted by Campbell (1998: 105-107), by comparing each translator's version of a unit with every other translator's version of the same unit. Each translator's rendition of each segment was thus entered in two tables (each similar to that represented in Figure 1): one table aims to represent variation at the structural level, the other variation at the lexical level. The basis for entering renditions in the two tables was the comparison between renditions as described in the previous section. This gave a total of six tables (i.e. two for each of the three tasks), which will from now on be referred to, for brevity's sake and for want of a better term, as "comparison tables" (the six tables in their entirety are

[7] Campbell (1998: Chapter 6) compares, qualitatively and quantitatively, the renditions provided by a group of 16 translators for a sample of 40 individual SL words contained in a text that the subjects had translated in its entirety. The aim of the study was to probe the psychological motivations behind the choice of lexical equivalents made by the translators (who were translating into their second language). It is of interest to note that the sample of SL words was chosen in a "syntactically straightforward" (Campbell 1998: 105) section of the ST and that each word would be the first occurrence in the text - criteria that in the present study have not been taken into account at the moment of mutually comparing TT renditions, neither at the lexical nor at the structural level.

reproduced in Appendix 5). Thus, for instance, a comparison table would show that for a given ST segment (say, segment 2 in Figure 5.9), CB, MT and VM provided structurally similar renditions (all marked as "a"), while MS and SC provided renditions that were different from any other (marked as "b" and "c" respectively).

	CB	MS	MT	SC	VM
1	a	b	b	b	b
2	a	b	a	c	a
3	a	a	a	a	a
4	a	b	c	b	d
...	...	...	...	...	...

Figure 5.9 – "Comparison table" used to differentiate between renditions (the first column represents ST segments, identified by numbers; cells contain a symbol for the renditions provided by translators: the same symbol indicates that renditions are identical according to the aspect considered, i.e. either structure or lexis)

Each symbol as entered in the table identifies a given rendition as a type. The table can therefore be used to give a complete picture of the type/taken ratio for all segments in each task, for both levels considered. In the following, data will be presented in more detail regarding the relative frequencies of renditions with different type/token ratios. These data will help in building up an overall picture of variation for each task, based on which a comparative analysis of the tasks will be proposed.

5.3.1 Measuring variation in TT renditions

Based on the comparison tables, a count was made of ST segments that displayed the higher degrees of TT variation in each task. More specifically, for each individual task (and for both levels of variation, i.e. structural and lexical) segments were counted that resulted in renditions characterised by type/token ratios of 3/5 or higher, thus obtaining, for each task, a subset of segments to be submitted to further analysis. Segments with renditions having type/token ratios below 3/5, i.e. with fewer than three different types of translation solution chosen by the five translators, were excluded from the count because they were considered less significant in terms of variation.

In general, the higher the number of solution types, the greater the variation. Hence, a type/token ratio of 3/5 is a relatively weak indication of variation, as it implies that three out of five translators chose the same solution. A type/token ratio of 4/5, is a stronger indication of variation, as it means that only two translators chose the same solution. Finally, a type/token ration of 5/5 is stronger still as an indication of variation, as it implies that all translators chose a solution that was different from the others.

The result of the segment count is given in Tables 5.1 and 5.2 below, referring to the structural and lexical level respectively.

Table 5.1 – ST segment count according to type/token ratio of TL renditions at structural level

	Task 1		Task 2		Task 3	
5/5	1	(<1%)	1	(<1%)	1	(1%)
4/5	9	(12%)	5	(9%)	1	(1%)
3/5	14	(19%)	8	(15%)	13	(13%)
Sub-total	**24**	**(32%)**	**14**	**(25%)**	**15**	**(15%)**
1/5 + 2/5	50	(68%)	41	(75%)	85	(85%)
Total	74	(100%)	55	(100%)	100	(100%)

Table 5.2 – ST segment count according to type/token ratio of TL renditions at lexical level

	Task 1		Task 2		Task 3	
5/5	3	(4%)	7	(13%)	2	(2%)
4/5	9	(12%)	7	(13%)	11	(11%)
3/5	16	(22%)	10	(25%)	22	(22%)
Sub-total	**29**	**(39%)**	**24**	**(44%)**	**36**	**(36%)**
1/5 + 2/5	45	(61%)	31	(56%)	64	(64%)
Total	74	(100%)	55	(100%)	100	(100%)

The two tables can be discussed first separately and then in comparative terms. As regards the structural level (cf. Table 5.1), the task in which renditions are seen to vary the least is Task 3, where only 15% of all segments show at least three solutions for any given segment. Of these, a very clear majority are at the lower end of the variation scale; higher type/token ratios are observed in only two segments (one

occurrence of four solutions and one of five). In the other two tasks, the percentage of segments showing variation of at least three solutions for a segment is quite higher: 25% in Task 2, rising to 32% in Task 1. What distinguishes these two tasks from Task 3 is the presence of a higher number of 4/5 segments: 12% in Task 1 and 9 % in Task 2. Note, however, that 5/5 segments only appear once in these tasks, too. Overall, the three tasks exhibit, at the structural level, a predominance of what I have labelled "weak" variation, i.e. variation resulting from renditions with a type/token ratio of 3/5. Such predominance, however, is less marked in Tasks 1 and 2. One possible explanation for this difference between Tasks 1 and 2 on the one hand and Task 3 on the other could be that segments in Tasks 1 and 2 are, on average, longer than those in Task 3.[8] More specifically, if we assume that the longer an ST segment, the more likely that its TL renditions are different from one another (both structurally and lexically),[9] this could explain the higher degree of strong variation in Tasks 1 and 2.

At the lexical level (cf. Table 5.2), variation concerns a higher proportion of the segments in all tasks. Task 2 is the highest, with almost half (44%) of its ST segments showing at least three TT solution types. Task 1 comes second (at 39%), with Task 3 a not too distant third (36%). The different ratios are more evenly distributed than at the structural level, especially in Tasks 1 and 2. Segments with a 5/5 ratio are rare in Task 3, but more frequent in the other two tasks (see particularly Task 2, where such segments constitute 13% of the total segments, thus equalling the percentage of 4/5 segments). As in table 5.1, 3/5 segments account for the majority of cases in all tasks but the presence of segments with "strong" variation (i.e. with a type/token ratio of 4/5 or 5/5) is more pronounced.

Looking at the two distributions comparatively, the first thing to be noted is the more marked variation occurring at the lexical level, where all tasks have at least a third of their segments corresponding to renditions with a minimum type/token ratio of 3/5 (Task 2 has well over a third). As already remarked, the lexical level also has a higher proportion of strong-variation segments; 5/5 segments, in particular, have a higher incidence than at the structural level.

---

[8] The average ST segment length, in number of words, is as follows: Task 1 = 3; Task 2 = 4.1; Task 3 = 2.5. This count considers all words in the texts, i.e. "lexical" words as well as "function" words (for the distinction between the two categories, see Biber *et al.* 1999: 55).

[9] Cf. 5.3.2.3, where a long ST segment (no. 4, ST 2) and its translations are discussed.

More than one hypothesis can be put forward as regards the factors that may have contributed to the differences between the patterns of variation emerging in the two tables. In comparative terms, the fact that variation is rather more pronounced at the lexical level than at the structural level (not only because more segments result in different renditions but also because there are, lexically, more cases of strong variation among renditions) could be, on one hand, a natural consequence of the language-typological similarities between English and Italian, leading to minimal changes to the structure of texts. The lower variation rate for the structural level may, however, also be due to the translators' particular "disposition" (Campbell 1998: 104) vis-à-vis the tasks, that is, their attitude characterised in terms of persistency and willingness to take risks (or a lack of either). More specifically, translators might have been wary of introducing structural changes, limiting themselves to those changes only which they considered strictly necessary to achieve morpho-syntactic correctness in the TL.[10]

At the lexical level, on the other hand, translators might not have been as wary, since they were, in a sense, forced to provide TL equivalents, even when they were conceivably not sure of their appropriateness or correctness (note that none of the translators left words untranslated). Regarding in particular the subject-specific terminology contained in the STs, a further hypothesis is that, not having enough familiarity with the domain, translators ended up proposing a wide spectrum of possible TL terminological equivalents, thereby contributing to an increase in the variation rate for the lexical level. Such a spectrum would include correct/appropriate items, less appropriate but still not completely incorrect items and, finally, incorrect/inappropriate items.[11]

To sum up, when variation between TT renditions is observed in the three tasks, the level at which variation is more marked is the lexical level. Structural

---

[10] This was probably to be expected of student translators; recall, however, that some of them already had some experience in translation at professional level prior to the experiment. SC, in particular, had already translated a university textbook (see 3.4 for details on the subjects).

[11] As an example (and based on the suggestions made by the evaluators consulted for this study; cf. 3.5.2), consider the ST element *crack tip*, variously rendered by the translators as *apice della cricca* (a correct translation and one appropriate to the genre), *estremità dell'incrinatura* (a correct translation from the point of view of denotative meaning, but one that experts in the field would not have provided, thus inappropriate to the genre) and *punto di frattura* (which was judged by expert evaluators to be an incorrect translation, especially as *punto* – considered too vague a term here – is concerned). An evaluative dimension in the analysis of the TT renditions will be introduced in a more systematic fashion later on in the study (Chapter 6).

variation occurs to a smaller degree in all tasks, although there are differences between one task and the other, with Task 3 varying the least and Task 1 the most.

5.4 Using CNA to identify textual 'hot spots'

So far, texts have been compared to each other in their entirety. One of the reasons for now undertaking a CNA, is the identification of *specific* parts of each of the three source texts, i.e. the so-called 'hot spots', arrived at by using the measure of variation described in the previous section.

Such an approach was already applied earlier on the study, notably when identifying ST hot spots on the basis of editing data (see 4.5). Hot spots will now be identified based on the analysis of variation carried out on the target texts. As with the editing data, highly varying units are identified in the TL first, but are then mapped onto the ST segments to which they relate. The criterion adopted for the identification of hot spots is similar to that used in the analysis of editing data and is based on a fairly straightforward quantitative measure: where a majority of translators (i.e. 3 out 5 or more) are seen to provide a distinct type of TL rendition, then the corresponding ST segment is identified as a hot spot. In terms of the 'type/token' approach to describing variation in TT renditions, the majority in a group of 5 translators means considering type/token ratios of at least 4/5, i.e. segments for which there are four or five different translation solutions, as illustrated in Figure 5.10.

Translator 1 provides rendition *a*
Translator 2 provides rendition *b*
Translator 3 provides rendition *c*
Translator 4 provides rendition *d*
Translator 5 provides rendition *a*

(Type/token ratio = 4/5)

These three translators, i.e. the majority, have provided a TT rendition which is different from any other

Figure 5.10 – Relationship between type/token ratio of 4/5 and number of translators providing distinct renditions

Table 5.3 below presents the ST segments associated with strong variation across TT renditions for both the structural and the lexical level. Some segments are listed as pairs, indicating that their TL equivalents were consistently observed to act as single segments in all TTs (as discussed in 5.2.2.1).

Table 5.3 – ST segments resulting in strong variation across TT renditions (normal font indicates 4/5 segments; bold indicates 5/5 segments; bottom row gives totals)

ST 1		ST 2		ST 3	
Structure	**Lexis**	**Structure**	**Lexis**	**Structure**	**Lexis**
17	**1**	7	**3**	**23**	18
24-25*	18	16-17*	**4**	59	20
31	23	25	**6**		22
36	34	32	7		33
41	**36**	**41**	12		39
54	37		**13**		40
56	46		14		**49**
59	56		15		50
63	**58**		16		**53**
66	64		**17-18***		56
	68		22		64
			25		69
			34		94
			55		
10	*11*	*5*	*14*	*2*	*13*

* ST segments treated as single segments in the TT by all translators

In terms of TL variation, the segments listed in table 5.3 can be seen as the hot spots for the three STs under investigation. The number of structural hot spots and lexical hot spots varies across and within texts. In ST 1 there is almost no difference between the two groups: the text has 10 structural and 11 lexical hotspots. In ST 2 the number of hot spots for lexis is nearly three times the number for structure. The diffference is even greater in ST 3, with 2 structural and 13 lexical hotspots.

The higher number of hot spots resulting from lexical variation (both overall and in STs 2 and 3 in particular) should, however, be reconsidered in light of an aspect which has so far not been taken into account in differentiating between TT renditions, namely the possibility that some lexical items are repeated in different segments of the same ST. An ST lexical element (say, a term) occurring in one segment and repeated in a later segment may have been translated in the same way in both segments. If that particular element contributed to TT variation in its first occurrence, then it is likely to contribute to variation in its second occurrence as well. Yet, if we take variation as an indicator of an increase in effort on the part of the translators, the significance assigned to the first instance of variation is likely to be

greater than that assigned to the second instance of variation. In other words, if a particular element was really the trigger for an increase in effort the first time it occurred, for its later occurrences translators can be assumed to have spent much less effort in providing a TL rendition. Therefore, if we take variation as the basis for identifying textual hot spots in terms of effort, we may well decide to reconsider the status of segments containing repeated lexical items and eliminate them from the group. Such segments are listed in Table 5.4 below, where an indication is also given of the lexical item repeated in the segment and of the segments where the items appeared first. Note that only "lexical" words, as opposed to "function" words (Biber *et al.* 1999: 55), are considered here.

Table 5.4 – Segments containing lexical items repeated from previous segments (previous segments are indicated in brackets after the items)

ST 1	ST 2	ST 3
36: *crack* (1) 46: *crack* (1)	15: *rate* (12); *crack* (12); *tip* (7) 18: *stress* (4); *crack* (12); *tip* (7) 25: *stress* (4) 55: *crack* (12); *rate* (12)	33: *tip* (20)

The catalogue of segments associated with high lexical variation in TT solutions can therefore be pared down. A final list of hot spots is thus obtained (Table 5.5; note that in this new list 4/5 and 5/5 segments have been aggregated).

Table 5.5 – ST hot spots in terms of TT variation (bottom row gives totals)

ST 1		ST 2		ST 3	
Structure	**Lexis**	**Structure**	**Lexis**	**Structure**	**Lexis**
17	1	7	3	23	18
24-25*	18	16-17*	4	59	20
31	23	25	6		22
36	34	32	7		39
41	37	41	12		40
54	56		13		49
56	58		14		50
59	64		16		53
63	68		22		56
66			34		64
					69
					94
10	*9*	*5*	*10*	*2*	*12*

* ST segments treated as single segments in the TT by all translators

In this new list, the difference between structural and lexical hot spots has been reduced. Lexical hot spots are no longer the dominant category for all texts. The difference between lexical and structural is still most marked in ST 3, but has been reduced in ST 2.

The most interesting result is the difference between the tasks as regards the number of structural hot spots. When the number of hot spots is considered as a percentage of the total number of segments in each ST, the resulting figures are those given in Table 5.6.

Table 5.6 – Hot spots as a percentage of total segments in the STs, per category

	Structural level	Lexical level
ST 1	13%	12%
ST 2	10%	18%
ST 3	2%	12%

These figures indicate a more marked difference in terms of structural than lexical profile. In particular, the increase in effort at the lexical level characterises a comparable number of segments in all three STs, with a peak in ST2. At the structural level, on the other hand, one text (ST 3) stands out for its low number of

hot spots: while both ST 1 and ST 2 show an increase in effort at the structural level for roughly one tenth of their total segments, ST 3 sees an increase in effort in a much smaller number of cases. It is now time to have a closer look, in the STs, at the group of segments observed to vary at the structural level, so as to see what particular structural features they display and whether such features are recurrent across the group.

5.4.1 A closer look at ST 'structural' hot spots

The three figures below reproduce the ST segments which were identified in the previous section as hot spots in terms of structural variation observed in their respective TT renditions (segments listed as pairs indicate those that were treated as single segments in the TT by all translators; the double slash marks the segment boundary).

Segment no.	
17	*The dissociative reaction [we have discussed]*
24-25	*are // so unreactive**
31	*the application of stress*
36	*A crack tip*
41	*focuses*
54	*within a few atomic dimensions of the tip.*
56	*to these large stresses,*
59	*from its normal bonding configuration.*
63	*a silicate tetrahedron*
66	*on the corner oxygen atoms*

Figure 5.11 – Hot spots in ST 1 (structural level)

Segment no.	
7	*the crack-tip bond rupture event*
16-17	*increase // as**
25	*with applied stress.*
32	*Small siloxane ring structures*
41	*for different size siloxane ring structures,*

Figure 5.12 – Hot spots in ST 2 (structural level)

Segment no.	
23	*roughly*
59	*Two factors, stress and the chemical environment,*

Figure 5.13 – Hot spots in ST 3 (structural level)

A few remarks can be made about the segments in Figures 5.11. 5.12 and 5.13 in order to identify features that may have contributed to the increase in effort that we are assuming they are a signal of. The first thing to be noticed is perhaps the diversity between segments in terms of length. It was not only very long segments (such as number 41 in ST 2) that led to high variation in the TT but also one-word segments such as the adverbial *roughly* (number 23, ST 3). As a crude measure of length, we can consider the number of words per segment and establish, arbitrarily, that short segments are those of up to three words and long segments contain more than three lexical words. This gives two fairly balanced groups of segments: one of 8 short segments (24-25, 31, 36, 41 and 63 in ST 1; 16-17 and 25 in ST 2; 23 in ST 3) and the other of 9 long segments (17, 54, 56, 59 and 66 in ST 1; 7, 32 and 41 in ST 2; 59 in ST 3), which suggests that segment length is not a candidate factor leading to an increase in processing effort.

A second remark to be made about the set of segments included in the three figures has to do with the predominance of nominal groups. Of the 19 individual segments reproduced in the three figures, 13 are constituted by or contain nominal groups (17, 31, 36, 54, 56, 59, 63 and 66 in ST 1; 7, 25, 32 and 41 in ST 2; 59 in ST 3), 3 are verbal groups (24 and 41 in ST 1; 16 in ST 2), 2 are adverbials (17 in ST 1 and 23 in ST 3) and 1 is an adjective phrase (25 in ST 1). Nominalised forms thus present themselves as strong candidates in terms of the possible factors that may have contributed to an increase in processing effort for these segments.[12] More support for this hypothesis will have to be sought in the findings related to the other data that have been collected for this investigation: the data on editing, discussed in Chapter 4, and those given by an evaluation of the TTs, which will be the subject of the next chapter.

[12] Are there segments with high nominal content which did *not* lead to an increase in processing effort as signalled by TT structural variation? To answer this question the structural hot spots identified here should be checked against the whole set of highly nominalised ST segments. As this issue is relevant for the whole of the investigation, it will be addressed in Chapter 7 (see 7.2.3 in particular).

Chapter 6
ERRORS AS SIGNALS OF DIFFICULTY

6.1 Translation errors

The basis for the way errors are employed as data within the framework of the present study was laid out in Chapter 2 (see Section 2.6). Errors can be seen as problems which the translator was not able to solve and, as such, as one of the indicators of the difficulty of a translation task. Within the methodological framework of the study, errors are to be seen as data to be combined with information and findings obtained through other observational methods. In particular, errors are treated as a measure of accuracy and, as such, used to complement the picture of variation across TTs that emerges from Choice Network Analysis. While CNA is used as a method to investigate text difficulty in translation, errors are a means of establishing a relationship between difficulty and accuracy. As already pointed out in 2.3.1, it was found (Hale and Campbell 2002) that it is possible for different subjects to translate the same ST element identically but inaccurately or, conversely, that subjects can translate an element differently but accurately. Establishing the nature of the relationship between the number of alternative renditions and the accuracy of these renditions is thus felt to be a crucial issue. Therefore it was decided to submit the translations in the study to an error analysis. The basic aim was to single out inaccurate renditions provided by a majority of translators, which should give additional support to the identification of those renditions as textual 'hot spots'.[1]

6.2 Criteria for Error Analysis of the TTs

The fifteen target texts (five for each of the three STs) produced by the participants to the study were submitted to a panel of three evaluators (see 3.5.3). Of these, two

[1] A brief note on the terminology employed in this chapter: the term 'inaccurate' will be used throughout as a synonym for 'incorrect' or 'erroneous'. When discussing non-binary errors (i.e. errors that are relative to factors having to do with the communicative purpose of the translation), erroneous TT solutions may also be labelled 'inappropriate' so as to stress the nature of the error. The different labels, however, are neutral with respect to the seriousness of the errors they refer to.

were domain experts (researchers in materials engineering from the University of Modena and Reggio Emilia) and one was a teacher of specialised translation (from the SSLMIT of the University of Trieste). They were asked to identify what they considered errors in the TTs and to distinguish between these errors in terms of seriousness. No predefined grid was given to the evaluators but they were asked to be as specific as possible in their comments. Note that for one of the three STs (ST 1, the extract from Michalske and Bunker 1987) a published Italian translation was also available (Michalske and Bunker 1988) and was taken into consideration in evaluating the corresponding TTs.

The comments provided by the evaluators were collated so as to obtain a uniform picture of the errors found in the TTs. In particular, errors were included in one of two broad categories: "binary errors" and "non-binary errors" (a distinction first proposed by Pym 1992; see 2.6.3). Binary errors are outright mistranslations or mismatches in the meaning of a TT rendition and its ST counterpart. Non-binary errors derive from the choice of a TT rendition felt to be less appropriate than another possible rendition, this last being more appropriate to the requirements of the task (e.g. in terms of register or textual function). In Pym's (1992: 282) words: "for binarism, there is only right and wrong; for non-binarism there are at least two right answers and then the wrong ones". Non-binary errors are reminiscent of Gouadec's (1989) "relative" errors, i.e. solutions that do not conform to the requirements of a given translation project. As such, they appear less relevant for the purposes of the present investigation, in that their orientation is more clearly towards the target text.[2]

Once the errors had been identified by the evaluators, they were submitted to further analysis, this time with the explicit aim of identifying factors that might have contributed to the translator's proposing an erroneous rendition. At this stage, the analysis focussed on binary errors, and particularly on a subset of errors that were considered to be more clearly oriented towards the ST (see 6.2.2.2 below). Focussing on binary errors does not imply considering them more serious than non-binary

[2] The distinction between binary and non-binary errors can also be related to Toury's (1995: 56-57) distinction between "adequacy" (adherence to source norms) and "acceptability" (adherence to target norms). Thus (and even at the risk of oversimplifying Toury's concepts), a binary error could be seen as a rendition which is not adequate with respect to the ST, whereas a non-binary error could described as a solution which is not acceptable in terms of the target language.

errors: the attention paid to them in the present context is to be related to the ST orientation implied by the research question.

A complete picture of the binary and non-binary errors found in the TTs is given in Appendix 3. In the following sections, after briefly illustrating some examples of non-binary errors, a more detailed discussion of binary errors will be presented. As pointed out earlier, the initial identification of errors is based on the comments made by the three evaluators. Note that the two categories of errors considered here do not take into consideration all the errors that may be found in the TTs under analysis. Inaccuracies deriving from spelling or punctuation errors were considered irrelevant for the purposes of the present investigation, because they are generally oriented towards the TL, and therefore excluded from the analysis (with one exception: CB's and MT's renditions of segment 23, ST 2; see 6.2.2.2 for details).

6.2.1 Non-binary errors (N=79): some examples

As noted earlier, the notion of non-binary error is best considered in relative terms and by establishing as a point of reference the requirements of the translation task. This means that labelling an error as "non-binary" depends on factors such as the translation brief, the intended TL readers and ultimately all those aspects that some translation theorists have subsumed under the notion of *skopos*.[3] Thus, if the purpose of the translation is, say, that of providing a mere gloss for the source text, the criteria for considering an error as non-binary may be different from those adopted when the translation serves a different purpose. For the texts at hand, the brief given to translators was such that the *skopos* implied for their TTs was to be the same as that of the original texts. In other words, they were asked to act as if their translations had the same communicative purpose and the same type of intended audience as the source texts (and relevant background information was provided to them regarding

[3] The term (Greek for 'purpose' or 'goal') has been introduced in translation theory by the German scholar Hans J. Vermeer. For a brief overview of the notion, see Hatim (2001: 73-80) or Baker (1998: 235-238), where references to Vermeer's works and to those of other scholars working within the same paradigm are given.

the STs).[4] The evaluators were also informed of what the translation brief was: they, too, were given detailed background information about the STs.

It is not easy to give a quantitative picture of the non-binary errors found in the TTs since in some cases the evaluators limited themselves to highlighting a certain TT element that they considered noteworthy but did not explicitly label it as incorrect or erroneous. Table 6.1 gives a breakdown of the number of non-binary errors per translator and task, obtained by a conservative count (i.e. leaving out less clear cases).[5]

Table 6.1 – Non-binary errors in the three tasks (N=79)

	Task 1	Task 2	Task 3
	f	*f*	*f*
CB	9	5	10
MS	5	4	8
MT	5	5	3
SC	9	4	1
VM	3	3	4
Total	31	22	26

Some of the erroneous solutions included in the above count can be characterised as stylistic lapses; see for instance the following extract taken from CB's translation of ST 1 (the relevant portion of text is underlined):

(1) *the application of stress can cause cracks...*
il fatto di esercitare pressione può portare a fenditure che ...
[the fact of applying pressure can lead to splits that ...]

where the TT paraphrase containing an infinitive ("il fatto di esercitare") was felt to be less appropriate to the genre than a possibile TL nominal ("l'applicazione di"). Similarly, MS's s lexical choice for *shatters* (segment 10, ST 3) in the following example was considered inappropriate for this particular text (as opposed to *si frantuma*, a more likely choice in this text type in Italian and, incidentally, the one proposed by all other translators).

[4] This, of course, could not control for the 'putative' idea of translation entertained by each individual translator, that is to say, her own view of translation as a special communication situation.
[5] Note that, in this and in the other counts of errors given later on, each individual error is counted only once, i.e. at its first occurrence. Repetitions of errors throughout a text were left out of the count.

(2) *when glass* <u>*shatters*</u>
quando il vetro <u>va in mille pezzi</u>
[when the glass goes into a.thousand pieces]

Appropriateness (or a lack thereof) was also noted by evaluators with regard to collocational restrictions. Figure 6.1 below reproduces all TT renditions for a stretch of segments in ST 3. All the equivalents provided for the verb *occurs* were judged not to collocate appropriately with the verb's subject (one evaluator proposed *la tensione si sviluppa*, 'stress develops', as an alternative rendition).

Stress [...] <u>*occurs*</u> *when...*	
CB:	La pressione [...] <u>ha luogo</u> quando ... [The pressure [...] takes place when]
MS:	La tensione [...] <u>ha i suoi effetti</u>... [The stress [...] has the.its effects ...]
MT:	La tensione [...] <u>ha luogo ...</u> [The stress [...] takes place ...]
SC:	le sollecitazioni [...] <u>si riscontrano</u> ... [the stresses [...] are observed ...]
VM:	La tensione [...] <u>si ha</u> quando... [The stress [...] is found when ...]

Figure 6.1 – Translations of segments 62 to 65 in ST 3 (segment 63 is omitted)[6]

In terms of appropriateness or acceptability, another interesting example is the following TT rendition provided by CB for the segments 73 to 76 in ST 3:

(3) *glass does not conduct heat well*
il vetro non conduce bene il calore
[the glass does not conduct well the heat]

Of the five translators, CB was the only one who closely reproduced the ST's verbal clause. In all other TTs, an alternative rendition was consistently provided which represented a shift to a nominalised structure: *il vetro non è un buon conduttore di calore*. One of the evaluators was quick to notice the difference between CB's rendition and the alternative solution provided by the other translators, which was judged to be more appropriate to the register of the text. CB's translation is discussed here because it provides a good example of how the selection among possible TT solutions can occur along a scale of increasing appropriateness, even with all the possible TT solutions securing an effective transfer of denotative meaning. What

[6] The same stretch of segments is also discussed in Chapter 5 (see 5.2.2.2) as an example of variation across TTs at the lexical level.

probably led the evaluator to prefer the nominalised TT rendition was exactly its formulaic nature: the clause *X è un buon conduttore di* (followed by such nouns as *calore*, 'heat', *elettricità*, 'electricity', and so on) could be described as an extended, "prefabricated unit" of meaning such as those postulated by Sinclair (1991).

A number of solutions judged to be inadequate by the evaluators occupy a grey area in between incorrectness/inaccuracy (i.e. a binary error) and inappropriateness/unacceptability with respect to the task (i.e. a non binary error). Some of the evaluators highlighted awkward formulations likely to lead to ambiguity in a text that, by its nature, should strive for clarity and precision. See, for instance, the following rendition by MT for segment 32 in ST 2, a nominal group:

(4) *small siloxane ring structures*
la struttura dei piccoli anelli di silossano
[the structure of.the small rings of siloxane]

in which there seems to be an assumption, not necessarily implied in the ST, that siloxane rings are always "small". To be sure, what exactly "small" refers to is not clear in the ST either (is it "rings" or "structures"?), but further information on siloxane rings given elsewhere in the text makes it clear that they can have various sizes (i.e. they re not necessarily "small"). For such TT solutions, inclusion in one or the other category of errors we are considering was decided on a case-by-case basis (this one was considered a binary error)..

This brief overview was principally meant to show how the majority of what we have called non-binary errors identified in the translations had to do with aspects closely related with the target language, which makes such errors less relevant then binary errors for the purposes of the present study. It was also shown, however, that a few borderline cases emerged where clear-cut assignment to one or the other category of errors was not possible.

6.2.2 Binary errors (N=120)

A binary error is a rendition of a given ST element that mistranslates that element – a wrong TT solution in which the propositional content is different from that of its

ST counterpart. The result of a binary error may be what traditional classifications of errors called "misinterpretation", "incorrect meaning", "nonsense" or "interference".[7]

Overall, based again on the evaluators' comments, a total of 120 binary errors were identified in the 15 TTs under analysis. As with non-binary errors, dubious cases were left out of the count (some of them will be illustrated later on). A breakdown of binary errors per translator and task is given in Table 6.2.

Table 6.2 – Binary errors in the three tasks (N=120)

	Task 1 *f*	Task 2 *f*	Task 3 *f*
CB	15	16	8
MS	13	4	6
MT	11	6	2
SC	6	7	4
VM	8	9	5
Total	53	42	25

Without making assumptions concerning their possible causes, the binary errors identified in the TTs can be grouped in two subcategories, identified on the basis of the aspect commented on by the evaluators, i.e. what they thought the TT rendition in question was erroneous for:

1) inaccurate TL lexical equivalent (n=85);
2) inaccurate ST content transfer (n=45).

The first subcategory includes TT terms or other lexical items judged as inaccurate by the evaluators (of the three evaluators, the two domain experts tended to focus their comments on terminology, while the other evaluator, who was not a domain expert, tended to comment on elements that can be considered non-domain specific). The second subcategory of binary errors includes TT renditions whose propositional content was judged to be different from that of their ST counterparts. To be sure, a difference in meaning or propositional content is the result of errors included in the other category as well (a wrong TL term has a different meaning from both the ST

[7] These four types of errors are listed and defined in Delisle *et al.* (1999) alongside others ("inappropiate paraphrase", "undertransaltion" and "hypertranslation") which would in the present context fall within the category of non-binary errors. Note that if traditional classifications of errors are mapped onto the binary/non-binary distinction some categories may be seen to cut across this distinction: "interference" for instance may result either in an inappropriate rendition (i.e. one still conveying content) or an incorrect rendition (i.e. one inaccurately conveying content).

term and the correct TL term); this second category, however, groups renditions for which no specific term or lexical item was identified by the evaluators as the erroneous element – inaccuracy was in such cases ascribed to a whole target-text unit rather than individual words.

Introducing subcategories of errors within the general binary/non-binary distinction would at first sight seem to run counter to the very simplicity the distinction was aiming at in the first place. With a view to using errors as data to be combined with information obtained through other observational methods, however, a reconsideration of Pym's distinction appears inevitable, since a more fine-grained differentiation between types of errors is assumed to be capable of providing a more accurate picture of the patterns emerging in the performance of translators and, ultimately, of the specific elements that can be characterised as difficult or problematic for them. In the following sections, a few examples will be discussed for each of the two subcategories and their distribution in the three tasks will then be presented.

6.2.2.1 Inaccurate lexical equivalents in the TTs (n=75)

A large number of the comments made by evaluators relate to individual lexical items that were judged to be inaccurate with respect to their ST counterparts. Most of these comments concern inaccurate renditions of ST terms. For example, the words *fenditura* and *incrinatura*, provided by CB and SC respectively as equivalents for *crack* (occurring in all STs), were considered erroneous. Other examples of terms translated inaccurately by some translators include *stress* (occurring in all STs) and *kinetics* (ST 1).

In other cases, erroneous lexical equivalents proposed by translators concerned not so much specialised terms as specialised senses of some ST words. A recurrent example of such an error concerns the word *rate*, which was present in all three STs and turned out to be consistently problematic for most translators. This word seems to play a crucial role in all texts. It occurs frequently[8] and always appears in lexical environments that are particularly relevant to the topic of the texts (e.g. *rate of crack growth*, ST 1; *rate of hydrolysis*, ST 2; *rate of crack propagation*,

[8] *Rate* occurs as follows: twice in ST 1; 5 times in ST 2; 3 times in ST 3.

ST 3). As pointed out by the evaluators, the correct Italian equivalent for *rate* in all its occurrences would be *velocità*, but most translators failed to provide it. Examples of incorrect equivalents for one particular occurrence of *rate* are given in Table 6.4 below. MS was the only translator who provided the correct equivalent, while all the other renditions were considered incorrect, including SC's decision to omit the element.

The rate of crack extension ...	
CB:	il livello dell'estensione della fenditura ... [the level of.the extension of.the split...]
MS:	la velocità di propagazione della frattura ... [the rate of propagation of.the fracture...]
MT:	il tasso di propagazione della cricca [the percentage of propagation of.the crack ...]
SC:	l'estensione dell'incrinatura ... [the extension of.the crack ...]
VM:	il tasso di estensione della cricca ... [the percentage of extension of.the crack ...]

Figure 6.2 – Translations of *rate* in segment 55, ST 2 (the segment is only partially reproduced)

The word *rate* would probably hardly qualify, in itself, as a domain-specific term, but it could be argued that its sense was very domain-specific in the three STs under investigation.[9] A similar observation can be made for another 'ordinary' word that seems to acquire a very specific sense in the three STs under investigation: the word *tip*, as in *the tip of the crack* (segment 21, ST 3). In most of its occurrences,[10] *tip* was rendered by translators as *estremità* ('end', 'termination'), a correct translation from the point of view of denotative meaning, but one that experts in the field pointed out was inappropriate to the genre – in our terms, a non-binary error and therefore not included in the count in Table 6.2. In Task 2, MS consistently translated *tip* as *punto*; for segment 7, for instance, her rendition was

[9] The three STs basically talk about the "kinetics of fracture" (ST 1) in glass; "kinetics" is defined in the CED as "the branch of chemistry concerned with the *rates* of chemical reactions" (emphasis added).

[10] Other occurrences of *tip* within segments are: *crack tip* (segments 36 and 46, ST 1), *crack-tip bond rupture event* (segment 7, ST 2), *at the crack tip* (segment 15, ST 2).

(5) *the crack-tip bond rupture event*
la rottura del legame chimico nel punto di frattura
[the rupture of.the bond chemical at.the point of crack]

which was judged by expert evaluators to be an inaccurate translation (i.e. a binary error).

6.2.2.2 Inaccurate transfer of ST content (n=45)

This subcategory of binary errors groups TT renditions which were observed to distort ST content without a clear association with a particular term or word. Based, again, on the comments proposed by the evaluators, a total of 45 errors were identified and included in this subcategory. These errors were then submitted to further analysis with the aim of identifying factors that might have contributed to the renditions being erroneous. In most cases TT renditions were observed to misinterpret ST content and the error was related to a malfunctioning arising in the ST comprehension stage.[11]

As a first example, consider this TT rendition proposed by CB for segment 29, ST 2:

(6) *the strain-enhanced chemical reactivity of silicate materials*
la reattività chimica che aumenta la deformazione dei materiali di silicato
[the chemical reactivity that enhances the strain of.the materials of silicate]

where the equivalent for *chemical reactivity* is made to act as subject of the relative clause in the TT (i.e. the chemical reactivity enhances the strain rather than being enhanced by it). The ST segment was here clearly misinterpreted by the translator at the syntactic level – with the proviso that the incorrect decoding of the segment may have been favoured by unfamiliarity with content.

A similar failure in correctly decoding the syntax of ST segment is observed in CB's and MT's structurally similar translations, given in Figure 6.3 below.

[11] The identification of the two separate stages of comprehension and reformulation in the translation process is the basis for Gile's (1992) operational model for the analysis of errors. In the last few years many researchers engaged in empirical studies implicitly or explicitly accepted this two-stage model and have associated problems (and errors) to one or the other stage, although they recognise that a clear-cut assignment is often impossible (see, for instance, Dancette 1992; Kussmaul 1998; Markus 2000). With regard to the comprehension phase, research has focussed especially on the cognitive mechanisms that result from the interaction between the information presented in the ST and the translator's pre-existing knowledge.

the crack-tip bond rupture event	
CB:	la rottura del legame fenditura-estremità [the rupture of.the bond crack-tip]
MT:	la rottura del legame cricca apice [the rupture of.the bond crack-tip]

Figure 6.3 – Inaccurate TT renditions of segment 7, ST 2

The ST segment and the two translations can be contrasted by paraphrasing them with more explicit formulations which show more clearly the difference in content between the ST unit on the one hand and the two TT units on the other (for the translations, the paraphrase is given in English):

paraphrase of ST:
the rupture event occurring at the tip of the crack;

paraphrase of CB's rendition:
the rupture of the bond between crack and tip;

paraphrase of MT's rendition:
the rupture of the bond between crack and tip.

VM's rendition of this same segment was among the cases considered to be at the borderline between binary and non-binary errors:

l'evento di rottura del legame della punta della cricca
[the event of rupture of.the bond of.the tip of.the crack]

This rendition was eventually assigned to the non-binary category on the grounds that, while admittedly ambiguous, it could nevertheless still be interpreted as reproducing the ST content correctly.

Another example of inaccurate decoding of ST syntax is provided by the translations of segments 65-66, ST 1 (see Figure 6.4 below). In this particular case, none of the translators was able to work out the correct interpretation of the ST.

by pulling on the corner oxygen atoms	
CB:	spingendo gli atomi di ossigeno nell'angolo [pushing the atoms of oxygen into.the corner]
MS:	con una tensione ai lati degli atomi di ossigeno [with a stress at.the sides of.the atoms of oxygen]
MT:	spostando in un angolo gli atomi di ossigeno [moving to a corner the atoms of oxygen]
SC:	perché gli atomi di ossigeno vengono spinti in un angolo [because the atoms of oxygen are pushed into a corner]
VM:	estraendo atomi di ossigeno da un angolo [extracting atoms of oxygen from a corner]

Figure 6.4 – Translations of segments 65-66, ST 1[12]

Looking at the TT renditions, it seems likely that for all translators the misinterpretation originated in the way they segmented the ST. Instead of considering *atoms* as the Head of the nominal group *corner oxygen* atoms in which both *corner* and *oxygen* function as modifiers, all translators seem to have considered the source text as made up of three segments (or "ranked constituents", to use the Hallidayan term): *by pulling // on the corner // oxygen atoms*, thus taking *on the corner* to be a separate segment (possibly a prepositional phrase acting as a circumstantial element).[13] Also, some translators seem to confuse pull with push ('spingere' in Italian; see CB and SC).

Erroneous TT renditions were provided by some translators for yet another segment (no. 4, ST 2) containing a densely packed nominal group (see Figure 6.5 below).

[12] Here is the larger context in which segments 65-66 appear in the ST: *Theoretical calculations show that if a silicate tetrahedron is distorted by pulling on the corner oxygen atoms, the silicon atom in the center is more likely to bond with a water molecule.*

[13] Regarding segments 65-66, which explain how a silicate tetrahedron is distorted when stress is applied to a glass crack, it is interesting to note that, while they were translating this part of the text, none of the translators referred to the several diagrams contained in the copy of the original *Scientific American* article they had been provided with. The diagrams show, in quite vivid detail, how a silicate tetrahedron becomes gradually distorted: their potential as visual aids for the correct interpretation of the ST was, however, left unexploited by the translators.

the mechanism of stress/environment dependent crack extension	
MS:	il meccanismo di dipendenza dell'estensione di una frattura rispetto allo [the mechanism of dependence of.the extension of a crack with.respect to.the sforzo esercitato su di essa e rispetto all' ambiente stress applied on of it and with.respect to.the environment]
SC:	il meccanismo dell'estensione dell'incrinatura che dipende da condizioni [the mechanism of.the extension of.the crack that depends on conditions ambientali o di sollecitazione environmental or of stress]
VM:	il meccanismo di estensione della cricca dipendente dalla tensione o [the mechanism of extension of the crack depending on.the stress or dall'ambiente on.the environment]

Figure 6.5 – Translations of segment 4 , ST 2

The three renditions reproduced in Figure 6.5 were all judged to distort content. In particular, the distortion can be considered stronger in SC's and VM's renditions, where a disjunction is proposed which has no counterpart in the ST (see "dipende da condizioni ambientali o di sollecitazione" in SC and "dipendente dalla tensione o dall'ambiente" in VM). MS's translation, on the other hand, does not make it clear enough that *stress* and *environment* have to be considered as being in a mutual relationship such that crack extension in glass depends on how they interact. As in the previous cases, it can be argued that the segment was rendered problematic by the combined effect of dense syntactic packaging in the ST and the translators' unfamiliarity with the subject matter. We return to this question in section 6.4 below.

Note that this combined effect may also have been operating for shorter stretches of text. A particularly interesting example is given by the nominal groups containing the term *silicate* (see an earlier discussion in 4.5). The term, a noun in English, appears in two source texts, in both cases functioning as a modifier: *silicate tetrahedron* (ST 1) and *silicate materials* (ST 2). It was translated inaccurately by CB and VM in Task 1 and by MS in Task 2. In Task 1, CB's translation was

(7) *a silicate tetrahedron*
un tetraedro silicato
[a tetrahedron silicate]

In Task 2, MS provided the following, similar rendition:

(8) *silicate materials*
materiali silicati
[materials silicates]

For both TT renditions, what the evaluators objected to was the use of *silicato* as an adjective in Italian (as attested by dictionaries, it is only used as a noun). In both cases, the alternative renditions recommended by evaluators were *silicatico* or *silicico* (the correct adjectival forms, both attested in dictionaries). From our particular viewpoint, these two renditions are interesting because they represent an instance in which the translation error may have been induced by the converging of a variety of factors, namely:

- the translators' unfamiliarity with the domain and with the ST term in particular;
- the ST structure, a short string of two juxtaposed items: while the leftmost item can safely be regarded as a Head, the grammatical function of the first item is more uncertain, especially for someone who is not familiar with the item itself;
- possibly – and very tentatively – the form of the Italian noun *silicato* itself, which the translators may have confused with a past participle (*-ato* being the Italian suffix for past participles, which can often be used with an adjectival function in Italian) – again, unfamiliarity with the term may have contributed to the mistake.

VM's erroneous rendition of *a silicate tetrahedron* was "un tetraedro contenuto in un silicato" ('a tetrahedron contained in a silicate'), where *silicate* was correctly identified as a noun; the relationship between *silicate* and *tetrahedron* (which remains implicit in the ST) is, however, wrongly constructed in the TL.[14]

Segments containing densely packaged syntactic structures account for 23 errors out of the 45 inserted in the subcategory of binary errors being discussed. These 23 errors relate to 12 different ST segments, which are reproduced in Figure 6.6. Note that none of these segments is from ST 3.

[14] To be sure, VM's rendition is one of those cases where the error is at the borderline between the binary and non-binary categories: strictly speaking, a tetrahedron *is* contained in a silicate, in the sense that silicates have a structure formed of tetrahedrons. The TL wording proposed by VM, however, is highly unlikely and was judged to be inaccurate rather than just inappropriate.

ST 1	ST 2
11: *the bond-rupture reaction* (MS) 23: *stress-free silicon-oxygen surface bonds* (CB, MT) 59: *from its normal bonding configuration* (VM) 63: *a silicate tetrahedron* (VM) 65-66: *by pulling // on the corner oxygen atoms* (all translators)	4: *the mechanism of stress/environment dependent crack extension* (MS, MT, SC, VM) 7: *the crack-tip bond rupture event* (CB, MT) 29: *the strain-enhanced chemical reactivity of silicate materials* (CB) 31: *through the use of model compounds containing small siloxane rings* (VM) 32: *Small siloxane ring structures* (CB, MT, VM) 53: *a quantitative chemical-kinetics-based model* (CB, VM)

Figure 6.6 – ST segments with syntax as contributing factor to erroneous renditions (authors of erroneous TT renditions indicated in brackets)

Unfamiliarity with the subject matter, which could have signalled some kind of cognitive dissonance between the syntactic interpretation and propositional content, can be seen as a contributing factor in other cases of ST misinterpretation. For some of these, once the ST segment and its TTs counterparts are compared, additional linguistic aspects can however also be identified as factors leading to an erroneous TL solution. Consider for instance the translations of segment 54, ST 1, reproduced in Figure 6.7 below.

within a few atomic dimensions of the tip	
CB:	in alcune dimensioni atomiche dell'estremità [in some dimensions atomic of.the tip]
MS:	all'interno delle poche dimensioni atomiche del punto d'origine [inside of.the few dimensions atomic of.the point of origin]
MT:	all'interno delle poche dimensioni atomiche della frattura (?)* [inside of.the few dimensions atomic of.the crack]
SC:	ad una piccolissima distanza dall'estremità dell'incrinatura [at a very.small distance from.the tip of.the crack]
VM:	in alcune dimensioni atomiche dell'estremità [in some dimensions atomic of.the tip]

Figure 6.7 – Translations of segment 54, ST 1
* Bracketed question mark as found in the TT

In four out of the five TT renditions (the exception being SC's translation) the translators apparently did not recognise the correct meaning of *within*, i.e. "not beyond the limits of" (CED). In this particular case, therefore, a lack of second

language competence interacted with unfamiliarity with content and led to the TT error. Note, incidentally, that SC's translation was also considered erroneous, but on different grounds, i.e. the vagueness of "ad una piccolissima distanza" ('at a very small distance') as an equivalent for *within a few atomic dimensions*.

There are still other erroneous TT renditions that can be associated with the ST comprehension stage, although for these an identification of specific syntactic traits or other linguistic elements of the ST was not possible. Unfamiliarity with the subject matter certainly played a role in all of these cases as well. A list of the segments in question is given in Figure 6.8 below, together with the corresponding erroneous renditions.

ST	Segment no.	Segment	TT rendition	Translator
1	1	*The rate of crack growth*	Il livello di fragilità del vetro [The level of fragility of.the glass]	CB
1	5	*The development of a complete model for the kinetics of fracture*	Lo sviluppo di un modello completo per [The development of a model complete for lo studio della cinetica legata alla frattura for the study of.the kinetics linked to.the fracture	MS
2	25	*with applied stress*	se verrà applicato uno sforzo [if will.be applied a stress]	MT
2	41	*for different size siloxane ring structures*	per diverse strutture di anelli di [for different structures of rings of silossano siloxane]	CB
3	13	*on the applied stress*	dal tipo di tensione [on.the type of stress]	MS

Figure 6.8 – ST segments misinterpreted on no apparent syntactic grounds

All the errors discussed so far have been characterised as misinterpretations of the ST and related to the ST comprehension stage. For other errors, attribution of a definite orientation towards either the ST or the TT is more problematic. As pointed out when introducing the criteria for error analysis, inaccuracies in spelling or punctuation as such, for instance, were excluded from the analysis because they were deemed irrelevant for the purposes of the present investigation. In the target texts at hand, however, there was at least one case where inaccurate punctuation in the TT was seen as actively contributing to distorting the transfer of ST content. Consider the following ST segment (no. 23, ST 1) and its rendition by CB:

(9) *stress-free silicon-oxygen surface bonds ...*
i legami di superficie silicio-ossigeno, che non sono stati sottoposti a pressione. ...
[the bonds at surface silicon-oxygen, which not have been subjected to pressure, ...]

where the Italian has a relative clause in between two commas which make the clause non-defining. In other words, in the TT all silicon-oxygen bonds are understood to be "stress-free", where the ST presupposes some are *not* stress-free; note that MT has a similarly erroneous rendition). It is difficult to say, however, whether this instance of inaccurate punctuation has to do with the translator misinterpreting the ST or whether it is to be discussed in terms of poor TL competence.

To sum up, in the vast majority of cases, misinterpretations of the ST were favoured by the translators' lack of familiarity with the subject matter. However, specific lexico-grammatical aspects of the STs were identified that could be seen to contribute to some TT renditions resulting in errors. In particular, a large subset (19 errors) of the 45 binary errors included in this category occurred in connection with segments containing nominal groups or highly nominalised constructs. Of the remaining errors, a few still had to do with linguistic aspects related to the ST, while for others an identification of specific syntactic traits or other linguistic elements of the ST was not possible.

6.3 Distribution of errors within and across the three tasks

Table 6.3 below illustrates the distribution of the two main categories of binary and non-binary errors in the three tasks and across the group of five translators.

Table 6.3 – Distribution of binary errors and non-binary errors in the three tasks (per translator)

	Task 1		Task 2		Task 3	
	Binary	Non-binary	Binary	Non-binary	Binary	Non-binary
CB	15	9	16	5	8	10
MS	13	5	4	4	6	8
MT	11	5	6	5	2	3
SC	6	9	7	4	4	1
VM	8	3	9	3	5	4
Total	53	31	42	22	25	26
Distribution within each task	63% (53/84)	37% (31/84)	66% (42/64)	34% (22/64)	49% (25/51)	51% (26/51)
Distribution across tasks	44% (53/120)	39% (31/79)	35% (42/120)	28% (22/79)	21% (25/120)	33% (26/79)

Looking across tasks, the task with the highest overall number of errors is Task 1 (84 errors). Task 2 has a total of 64 errors, while the total for Task 3 is 51. Non-binary errors are slightly more evenly distributed in the three tasks: Task 1 has 39% of all non-binary errors, Task 2 has 28% and Task 3 has 33%. For binary errors the difference across tasks is more marked: for Task 3 the number of binary errors (25) is less than half the number for Task 1 (53). Task 2, with 42 binary errors, comes closer to Task 1. Overall, there are more binary (N=120) than non-binary (N=79) errors across the three tasks. However, Task 3 differs from the other two for two reasons: 1) it has almost the same number of errors in the two categories, and 2) it has fewer overall errors than in the other two tasks. Compared to Task 3, in both Task 1 and Task 2 there are more binary errors relative to the total number of errors for the task. Considering that the subject matter of the three source texts is very similar, there must be something else in the texts that led the translators to commit more errors in ST 1 and 2 than in ST 3, and particularly more binary errors. To attempt an explanation for this difference, attention can be turned to the two subcategories that were identified within the binary errors: 'inaccurate TL lexical equivalent' and 'inaccurate ST content transfer'. Table 6.4 gives the distribution of the two subcategories (shortened to 'TL lexis' and 'ST content' respectively) in the three tasks and, again, for each individual translator.

Table 6.4 – Distribution of binary errors (N=120) in the tasks (per subcategory and translator)

	Task 1		Task 2		Task 3	
	TL lexis	ST content	TL lexis	ST content	TL lexis	ST content
CB	10	5	8	8	8	0
MS	7	6	3	1	4	2
MT	6	5	3	3	1	1
SC	4	2	6	1	2	1
VM	3	5	5	4	5	1
Total	30	23	25	17	20	5
Distribution between categories by task	57% (30/53)	43% (23/53)	60% (25/42)	40% (17/42)	80% (20/25)	20% (5/25)

The lower number of binary errors in Task 3 is due to a large extent to the lower number of errors deriving from what we have called 'inaccurate ST content transfer'. In other words, the vast majority of binary errors in Task 3 (i.e. 80%) have to do with instances of inaccurate lexis in the TL.[15] In the other two tasks, although 'TL lexis' errors still represent the majority of binary errors (57% in ST 1 and 60%), the relative difference with 'ST content' errors is much smaller than in Task 3.

6.4 The interaction between lexico-grammar and domain knowledge

What is it that led translators to make more 'ST content' errors in Tasks 1 and 2? Recall that in discussing examples from this subcategory of errors a number of cases were identified in which the error made by the translator could be linked to specific syntactic traits of the corresponding ST segment. In particular, a set of segments were identified that, alone, accounted for half the 'ST content' errors found in the two Tasks (23 out of 45 errors; see Figure 6.6 for the complete list of segments). The common trait among these segments is their highly nominalised style and, for some

[15] A closer look at the erroneous renditions included in the 'ST content' subcategory for Task 3 reveals that for 3 of them (out of the total 5) the error had actually to do with the same term, *commercial glass*, contained in segment 81: in their renditions, three translators (MS, MT, SC) omitted the second word of the correct equivalent in Italian (i.e. *vetro commerciale*). As with other cases of omission in the TL, this error was treated as an instance of ST content distortion. Debatable as this decision may be, it does not invalidate the argument we are pursuing here. If anything, it is another reminder of the uncertainty linked to attempts at classifying translation errors.

segments in particular, their heavy reliance on compounding. Considering an error as a certification, as it were, of difficulty, these segments leading to errors can be labelled as 'syntactically difficult'.

Extensive space was devoted in Chapter 2 to a description of nominal groups in general and compounding in particular as a preferred mode of designation of specialised concepts (see 2.5.1). It is now worth returning to some of the remarks made there, as they appear particularly relevant to a discussion of the 'syntactically difficult' ST segments identified here.

One of the distinguishing features of specialised language is not so much the presence of nominal groups *per se* as the tendency to expand such groups through the insertion of different types of modifiers. The possibilities for modification (carried out through extension, combination or sequencing) are limited only " by the ability of the intended recipient easily to comprehend the resultant text" (Sager *et al.* 1980: 219). Modification of a nominal group can occur either to the right (post-modification) or the left (pre-modification) of the Head. Post-modification gives rise to sequences of prepositional phrases (e.g. *the development of a complete model for the kinetics of fracture*), at times combined with non-finite verb forms (e.g *the use of model compounds containing small siloxane rings*). Pre-modification can give rise to compounds, combinations "of two or more words into a new syntagmatic unit with a new meaning independent of the constituent parts"; these units are formed on the principle that "one, usually the first, element determines the other" (Sager *et al.* 1980: 265).[16] Compounding can be cumulative and combinable with affixation: the number of constituents of multiple compounds can increase to five or six elements.

In nominal groups, and in compounds in particular, elements are often omitted that would make the relations between the constituents more explicit at the conceptual level (a typical feature of many LSP texts in English). Faced with such elliptical constructs, translators often find themselves, unlike domain specialists, unable to provide the missing elements that would make the relations explicit. So, for instance, *the bond-rupture reaction* (a segment from ST 1) may not be interpreted correctly as 'the reaction causing a rupture of the bond', as only subject-matter

[16] Note that nouns are not the only elements that can act as the Head of a compound (in *heat-resistant*, for instance, the Head is an adjective).

knowledge can disambiguate the possible semantic interpretations: e.g. 'a reaction leading to' or 'a reaction caused by'.

Reference to the global meaning of the text may help in correctly interpreting such constructs, but trainee translators, such as those involved in the present investigation, are often seen to focus on the microstructure if not on individual words of the ST (Séguinot 1991; Nord 1997). In other words, translators often rely on a 'bottom-up' approach to text processing, i.e. one starting from the elements that are present in the text. In 'top-down' processing, by contrast, the translator's pre-existing knowledge is projected onto the text. A similar distinction is proposed by Dancette (1992), who talks about "centrifugal" and "centripetal" comprehension – the former based on a close scrutiny of the texts' microstructure aimed at clarifying its macrostructure, the latter using the translator's knowledge and intuition. Dancette sees the centripetal mode as the one leading to the better results, in that it ensures that the test is understood globally even though 'local' errors may be made.

The evaluation of the target texts under investigation did not rest on a global, 'holistic' assessement of the texts – it only focused, in other words, on the microstructure, and looked at errors in the target texts, considered as a signal of difficulty at the 'local' level. When such errors are taken into consideration, a distinction can nevertheless be made, globally, between ST 1 and 2 on the one hand and ST 3 on the other. In particular, the analysis has revealed how the difficulty linked to ST syntactic traits emerged more for STs 1 and 2 than for ST 3. Thus, although the domain knowledge gaps on the part of the translators can be assumed to be the same for all three tasks (given that three texts represented roughly the same 'knowledge space'), these gaps interacted differently with the lexico-grammatical profile of each source text. Where highly nominalised, densely packaged constructs were present in the text, these were seen as magnifying the negative effect exerted by the gaps in domain knowledge.

In terms of our research question (which considers nominalisation as an effect of grammatical metaphor), the analysis of errors presented in the present chapter seems to give some support to the hypothesis that an increase in grammatical metaphoricity represents an obstacle (although by no means the only one) to ST comprehension. As regards each individual ST, although each was presented as occupying a different, successive point on a scale of increasing metaphoricity (with

ST 3 as the least and ST 2 as the most metaphorical texts), the findings obtained here only confirm an increase in difficulty from ST 3 to ST 1 and 2, but do not allow further differentiation between these two.

Errors, however, are only an overt signal of difficulty. The difficulty attached to a given ST may also manifest itself in ways that do not find their way into the final product of the translation. Therefore, the findings of the error analysis presented in the present chapter must be compared with those arrived at by analysing the other data sets available for the present investigation. This triangulation of data will be the topic of the next chapter.

Chapter 7
TRIANGULATING FINDINGS

7.1 Introduction

The purpose of this final chapter is twofold. First, a review and integration of the findings which have emerged in the previous analytical steps will be presented (7.2). Here, the aim is to observe instances of the possible convergence of such findings in relation to the STs under investigation. In other words, findings from the keystroke-based study of editing (Chapter 4), Choice Network Analysis (Chapter 5) and Error Analysis (Chapter 6) will be compared and contrasted (sections 7.2.1 and 7.2.2), and special attention will be paid to cases where they 'overlap' in relation to the STs (section 7.2.3). The ST segments where findings are seen to converge will be characterised as the 'hot spots' of the ST in question. Second, this review will serve to provide an answer to the research question that the study is investigating (7.3). More specifically, once the three datasets employed in the study have been compared and contrasted, the stated hypothesis that the difficulty posed by a source text increases with the level of grammatical metaphoricity displayed by the text will be addressed.

From a methodological point of view, this final step in the analysis relies on the assumption that collecting data through different methods and then interpreting each set of data in relation to the others enhances their validity. Known as "triangulation", this approach has established its own recent tradition in process-oriented studies of translation. It emerged as an attempt to overcome some limitations posed by methods based on verbal reports (either introspective or retrospective), which some researchers felt should be complemented by other tangible sources of evidence. As a complement to verbal reports, keystroke logs are one of the most common sources of data (other methods of data collection used in recent years in process-oriented studies include the logging of search processes in electronic environments, video recording and eye-tracking.) Keystroke logs have been used mostly to carry out quantitative analyses on particular aspects of the translation process (see details and references in 2.2.2.1). In O'Brien (2006), which reports on a study of the effort associated with post-editing Machine Translation

output, keystroke logs are combined with CNA; in particular, while keystroke logs are used to measure and analyse the pauses occurring in post-editing (pauses being equated to signals of cognitive effort), CNA is used to observe variation in the post-editing solutions offered by a group of subjects.

The present investigation can be seen as introducing still another variant of the triangulation approach (see Figure 7.1 below). To recap: keystroke logs were here exploited for the information they give on the editing performed by the subjects during the drafting phase; Choice Network Analysis was then performed *both* on the editing data and on the target texts, focussing on the translation process and ST-oriented data respectively. Findings from the analysis of both editing and the target texts were then integrated with those emerging from the Error Analysis conducted on the TTs – see the two diagonal sides of the ideal triangle in Figure 7.1. The third side of the triangle could have been considered, too (i.e. a comparison between the keystroke-log data and the TT-variation data), but priority was given to a comparison of each dataset with errors because we assume variation, either intra- or inter-subjects, to be a stronger indicator of difficulty if it is associated with inaccurate renditions.

Other ways in which keystroke-based data might be exploited to investigate ST difficulty are discussed later on in the chapter (7.4), where some final reflections on the methodology employed by the study are presented. These are followed by a short section illustrating the implications of the study for training, evaluation and research (7.5) and by some concluding remarks (7.6).

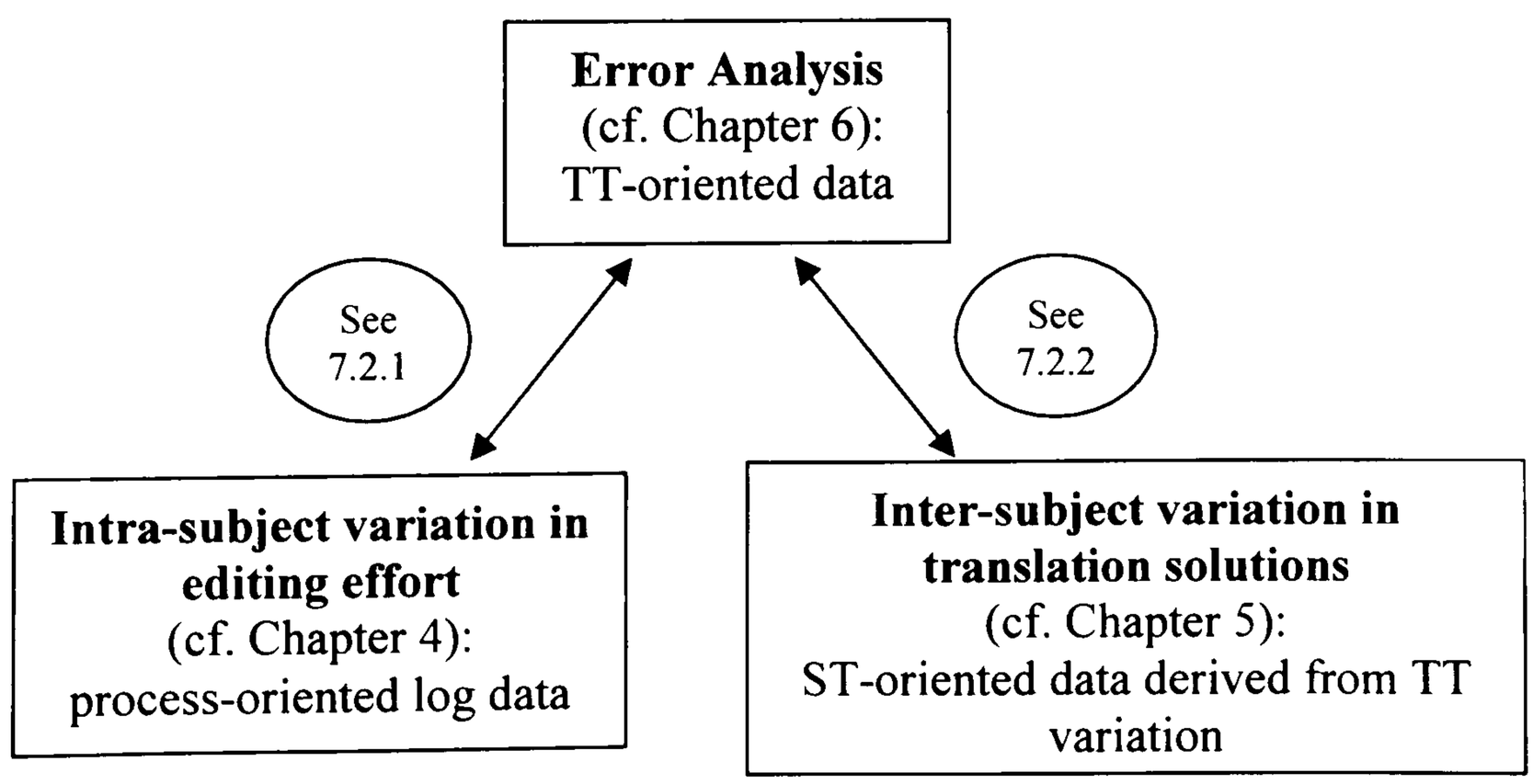

Figure 7.1 – Triangulation of findings

7.2 Integrating findings from the previous analytical steps[1]

In the first two analytical steps of the present investigation (i.e. those reported on in Chapters 4 and 5) a similar analytical procedure (Choice Network Analysis) was followed although on different datasets, i.e. keystroke-based data and the target texts respectively. In both cases patterns of variation were first observed across the three tasks (of interim solutions for individual translators in one case and final TT solutions for all translators in the other). and the focus then moved to the source texts, where a CNA approach was adopted, in both cases, so as to identify a set of 'hot spots', i.e. a set of ST segments where the feature under analysis was seen to emerge for a majority of the translators.

More specifically, in the first step (reported in Chapter 4) the changes introduced by the translators to their own draft during the task, i.e. editing, were characterised as a reflection of the processing effort made by the translators. Edits were grouped into three different categories based on the orientation of this effort ('ST-oriented', 'terminological' and 'TT-oriented') and their distribution was observed across the three tasks. ST-oriented and terminological editing were found to be more frequent in the STs displaying a higher degree of grammatical metaphoricity. In particular, the higher frequency of ST-oriented and terminological edits turned out to be parallel to the increase in ST metaphoricity. Using Choice Network Analysis, the analysis then focussed on each individual edit, treated as an alternative TT rendition; the aim was to identify, across translators, the units on which editing effort was concentrated; these units were mapped onto the ST segments they related to and a first set of ST hot spots (see Figure 4.5) was obtained. These hot spots were seen as segments associated with an increase in processing effort. Of the 21 segments thus identified, 7 were observed to be ST-oriented.

---

[1] The history of triangulation within Translation Studies is perhaps too recent for elaborate distinctions to have been proposed between terms such as 'integration', 'combination' or 'mix' of either methods or findings. In fields where triangulation has a longer history, the importance of paying attention to the practices involved in bringing multiple methods and datasets together has been extensively discussed and terminological distinctions have been introduced (in social science. for instance, see Moran-Ellis *et al.* 2006). In the present context, however, no further elaboration of the notion of triangulation is proposed and reference to the 'integration' of data or to the 'complementing' of methods should not be taken as an attempt at introducing finer methodological distinctions. Much as these would perhaps be needed in studying the process-related aspects of translation, they fall outside the scope of this investigation.

In Chapter 5, Choice Network Analysis was again used, this time applied to the target texts, as an established method of observing increases in processing effort in relation to certain ST segments. In particular, the increase in effort was takne to be reflected by the extent to which TT solutions for a given ST segment varied between subjects. Overall, the analysis found that variation across TTs was higher at the lexical than at the structural level. When mapping variation on to the STs, another set of hot spots was obtained (see Table 5.5); in particular, it was found that, of the three texts, ST 3 (the text with the lowest metaphoricity) stood out for its low number of structural hot spots. A closer look at the structural hot spots in ST1 and ST 2 revealed that the majority were constituted by nominal groups. This lexico-grammatical trait was characterised as one of the factors that contributed to the increase in processing effort for such segments.

The target texts were then submitted to an Error Analysis (Chapter 6) so as to identify inaccurate TT renditions. A subset of errors was identified that consisted in distortions of ST content and it was found (6.2.2.2; cf. also Figure 6.6) that half of such errors corresponded to ST segments containing densely packaged syntactic structures – these segments constituting a third set of ST hot spots.

An integration of these findings was deemed necessary in view of the fact that, within a CNA approach, the relationship between the number of alternative renditions and the (in)accuracy of these renditions is felt to play a crucial role in judging the difficulty posed by the texts (cf. Hale and Campbell 2002; see also 2.3.1). In the rest of this section the findings resulting from CNA will therefore be compared with those obtained through Error Analysis, with a view to progressively consolidating the group of hot spots already indicated in Chapters 4, 5 and 6. Errors will be taken as additional indicators of difficulty and segments will be identified which emerged as both CNA hot spots and erroneous TT solutions. The distribution of this consolidated group of segments in the three tasks will finally be determined and the lexico-grammatical features of the members of this group will be described and compared. As in the previous steps, the analysis is here conducted in essentially qualitative terms and focuses on aspects that have to do with the structural profile of the STs.

7.2.1 Intra-subject variation hot spots and TT errors compared

The editing hot spots that are the focus of this section are the segments that were identified from among segments that gave rise to ST-oriented edits, i.e. edits made for the purpose of reassessing the relationship of semantic equivalence between the ST and the TT.[2] The segments, a subset of those listed in Figure 4.6, correspond to those listed in Figure 7.2 below:

ST 1	ST 2	ST 3
1	4	NIL
23	7	
54	32	
63		

Figure 7.2 – Segments emerging as ST-oriented editing hot spots (intra-subject variation in editing based on log data)

A way of comparing these segments with TT errors is to see whether any of them also gave rise to erroneous solutions in the TTs and, in particular, to errors showing a similar ST orientation, i.e. those deriving from the 'inaccurate transfer of ST content' (or 'content errors' in short; see 6.2.2.2). The subset of segments for which at least two translators[3] were seen to provide TT solutions containing content errors is the following:

ST 1	ST 2	ST 3
23	4	81
54	7	
65-66	31	
69	32	
	53	

Figure 7.3 – Segments resulting in ST-oriented content errors (Error Analysis data)

[2] 'Terminological' edits, another of the categories considered in Chapter 4, shared the same purpose; however, they were differentiated from ST-oriented edits because of the essentially lexical nature of the element they modified (see, for instance, the change from "fenditura" to "cricca" as an equivalent for *crack* in the ST), which is also the reason why they are not considered in the present context.

[3] Considering at least two translators is intended to avoid possible biases deriving from the fact that one translator in particular (CB) committed, in all tasks, many more errors than the other translators.

Once this list of segments is compared to the previous one, the two are seen to converge on the segments listed in Figure 7.4, which we take to represent a first 'consolidated' cluster of ST hot spots in terms of structural difficulty.

ST 1	ST 2	ST3
23: *stress-free silicon-oxygen surface bonds* 54: *within a few atomic dimensions of the tip*	4: *the mechanism of stress/environment dependent crack extension* 7: *the crack-tip bond rupture event* 32: [*s*]*mall siloxane ring structures*	NIL

Figure 7.4 – ST-oriented editing hot spots matching segments leading to content errors in the TTs

7.2.2 Inter-subject variation hot spots and TT errors compared

The same set of ST segments resulting in content errors considered above (Figure 7.3) can be compared with the segments that emerged from the CNA carried out on the target texts in Chapter 5, more specifically the set of segments associated with high structural variation in TT solutions. The relevant segments are shown in Figure 7.5:

ST 1	ST 2	ST 3
17	7	23
24-25	16-17	59
31	25	
36	32	
41	41	
54		
59		
63		
66		

Figure 7.5 – Segments associated with high structural variation in TT solutions (inter-subject variation data based on TT solutions)

The consolidated set of hot spots resulting from this new comparison is given in Figure 7.6. Note, once again, that no segment emerging from the comparison belongs to ST 3.

ST 1	ST 2	ST 3
54: *within a few atomic dimensions of the tip*	7: *the crack-tip bond rupture event*	NIL
66: *on the corner oxygen atoms*	32: [*s*]*mall siloxane ring structures*	

Figure 7.6 – Inter-subject structural hot spots matching segments leading to content errors in the TT

7.2.3 Collating results

The sets of ST segments given in Figures 7.4 and 7.6, each representing what we have called a 'consolidated' set of ST hot spots, can now finally be combined, i.e. the two lists can be added to each other so that a final set is obtained including all the elements appearing in either list. The resulting cluster of ST segments (some of which actually appear in both lists) thus includes the following segments: no. 23, 54 and 66 from ST 1 and no. 4, 7 and 32 from ST 2 (see Figure 7.7). These segments are therefore those which the five translators found most difficult at the structural level:

ST 1	ST 2	ST 3
23	4	NIL
54	7	
66	32	

Figure 7.7 – Structurally difficult ST segments

In the identification of this cluster of segments, the threshold for inclusion considered for both editing and inter-translator variation data was stricter than that considered for errors. In other words, whereas CNA hot spots emerged as such because variation was observed for a majority of translators, in the case of errors a lower threshold was taken into account (two out of five translators) because errors were treated as additional indicators only. It would however also be interesting to see, at this point, how the cluster in Figure 7.7 compares with the set of ST segments for which errors can be seen as particularly strong indicators of structural difficulty *on their own*, i.e. the ST segments which resulted in content errors for a majority of translators, or those having a particularly 'error-engendering potential'. Five ST segments in all meet this condition. Of these, four segments have been identified as 'structurally difficult' (Figure 7.7), which can be taken as further confirmation of the validity of the method employed to arrive at identifying the segments in that set.

ST 1	ST 2	ST 3
54 65 66	4 32	NIL

Figure 7.8 – ST segments with high error-engendering potential (Error Analysis data)

Segment 65 from ST 1 is the only one not to have emerged in the previous steps. Recall, however, that the adjacent pair of segments no. 65 and 66 (*by pulling // on the corner oxygen atoms*) was already discussed in detail when analysing errors (see 6.2.2.2); as was already pointed out there, the generalised misinterpretation of these two segments was probably due to the translators' failed attempt at correctly locating the syntactic boundary between them. Now that a final cluster of difficult ST segments is sought, it seems appropriate that these segments should be included in the group as a pair. A final set of hot spots is thus obtained (Figure 7.9) and this can be regarded as the one that represents the segments that turned out to be the most difficult, at the structural level, in the three tasks under investigation.

ST 1	ST 2	ST 3
23: *stress-free silicon-oxygen surface bonds* 54: *within a few atomic dimensions of the tip* 65-66: *by pulling // on the corner oxygen atoms*	4: *the mechanism of stress/environment dependent crack extension* 7: *the crack-tip bond rupture event* 32: [*s*]*mall siloxane ring structures*	NIL

Figure 7.9 – Final consolidated cluster of structurally difficult ST segments based on the 3 datasets

A few general remarks can be made about this cluster of segments, some of which have already been discussed in detail in the previous chapters[4]:

1. All the segments can be described as densely packaged groups; some of them (the segments from ST2) come from clauses that are particularly dense lexically,

[4] For segment 54, ST 1, see 6.2.2.2 (example 9). For segment 4, ST 2, see: 4.3.1 (example 1); 5.2.2.3 (Figure 5.6); 6.2.2.2 (Figure 6.5). For segment 7, ST 2, see 6.2.2.2 (Figure 6.3). For segment 32, ST2, see 6.2.1 (example 4).

a factor which may have exerted a compounding effect in the 'unpacking' effort on the part of the translators.

2. The difficulty for the translators mainly consisted in their inability to correctly decode the constructs. As far as terms contained in the segments were concerned, translators could (and did) resort to reference works – they had access to dictionaries of various kinds and to the Internet. Faced with lexico-grammatically highly complex constructs such as those listed here, however, translators could find no pre-packaged solutions. It is interesting to note, in this respect, that the difficulty of some of these constructs even passed unnoticed for some translators: a construct such as *by pulling on the corner oxygen atoms* led to an error in the target texts of 4 out of 5 translators. However, for some of them no trace of this difficulty was found in the keystroke-based data in terms of either a longer processing time[5] or a succession of interim solutions.
3. Some of the segments could have been expressed in a more "congruent" form, i.e. a less grammatically metaphorical construct; so, for instance, *the mechanism of stress/environment dependent crack extension* (segment 4, ST 2) might have been phrased as 'the mechanism of crack extension that depends on the relationship between stress and environment'. For most of these segments, however, the "instantial" effect of grammatical metaphor (Halliday 1997: 30-31) can be seen at work in the source text: the nominalised form is being employed by the writer as a temporary construct set up to meet the argumentative needs of the discourse. For the translator, however, the effect is one of disorientation and the construct could not be easily unpicked on the basis of sound subject knowledge as this was not present.

The fact that some of these segments were already identified at previous stages of the analysis can be considered noteworthy in two respects. First, from a methodological viewpoint it testifies to the validity of the CNA approach followed in studying both intra-subject variation as evidenced by editing and inter-subject variation in the target texts. Even with this small sample of subjects, the analysis led to the identification of largely overlapping sets of segments, the salience of which

[5] Pauses, which are usually taken as signals of mental processing in keystroke-based data, were not considered systematically for the present investigation; some suggestions on how they could be analysed are given in section 7.4 below.

was then confirmed by the evaluation carried out on the TTs.[6] Second, it means that the claim to psychological validity for the difficulty attached to the segments in Figure 7.9 can be made on firmer ground, as it has been 'sieved' through more than one analytical filter. Needless to say, had a larger group of translators been involved in the study, the validity of both the method and the results would have been strengthened.

It may be legitimate to ask, at this stage, whether the ST segments that have been identified as structurally difficult segments – largely by virtue of their highly nominalised structure – are representative of the whole set of highly nominalised segments contained in the STs. In other words, it would be interesting to know why some nominalised segments have *not* been identified in the present data as 'structurally difficult' – part of the reason being, possibly, the fairly strict thresholds which have been established for filtering the data.

Let us therefore consider the segments identified above as the final cluster of ST hot spots (cf. Figure 7.9) together with all the other segments that emerged as hot spots at each of the previous steps in the analysis and let us check these segments in relation to the whole group of nominally dense segments contained in the three STs.

Tables 7.1 and 7.2 below give, for each ST, the number of nominally dense segments that, throughout the analysis, were observed to lead to either an increase in ST-oriented processing effort or a content error (referred to as 'higher-effort' segments). In particular, these 'higher-effort' segments are checked against two increasingly restricted sets of highly nominalised segments found in the source texts: the first set (Table 7.1) includes all ST segments containing at least two nouns (e.g *the atomic structure of silica*), while the second set (Table 7.2) only includes segments containing at least three nouns (e.g *the silicon atom in the center*). These two different, increasingly restricted groups are considered in order to establish how the incidence of 'higher-effort' segments changes as segments become nominally denser. As a further element for comparison, for the higher-effort segments an

[6] If the hot spots emerging from keystroke-log data are compared with those given by TT-variation data (i.e. if we consider the base of the triangle in Figure 7.1), four segments emerge as hot spots in relation to both datasets: no. 54 and 63 in ST 1 and no. 7 and 32 in ST2. Three of the four (the exception is segment 63, ST 1) also appear in the final consolidated cluster of hot spots given in Figure 7.8.

indication is also given of the subsets (one per table) of segments that contain a compound (as in *small siloxane ring structures*).

Table 7.1 – Nominalised ST segments leading to increase in processing effort or TT content error against total of nominalised ST segments (threshold = 2 nouns per segment)

	Total of nominalised ST segments (threshold for inclusion = 2 nouns)	No. of higher-effort segments	Subset of higher-effort segments containing a compound	Number of higher-effort segments as % of all nominalised segments	Number of higher-effort segments containing a compound as % of all nominalised segments	% of higher effort segments containing a compound
ST 1	21	8	6	38% (8/21)	29% (6/21)	75% (6/8)
ST 2	22	6	6	27% (6/22)	27% (6/22)	100% (6/6)
ST 3	22	0	n/a	0% (0/22)	0% (0/22)	n/a

Table 7.2 – Nominalised ST segments leading to increase in processing effort or TT content error against subtotal of nominalised ST segments (threshold = 3 nouns per segment)

	Total of nominalised ST segments (threshold for inclusion = 3 nouns)	No. of higher-effort segments	Subset of higher-effort segments containing a compound	Number of higher-effort segments as % of all nominalised segments	Number of higher-effort segments containing a compound as % of all nominalised segments	% of higher-effort segments containing a compound
ST 1	10	5	4	50% (5/10)	40% (4/10)	80% (4/5)
ST 2	18	5	5	28% (5/18)	28% (5/18)	100% (5/5)
ST 3	7	0	n/a	0% (0/7)	0% (0/7)	n/a

Of the three source texts, ST 1 is the one where high nominal content is more strongly associated with an increase in processing effort: 38% of its segments containing at least two nouns lead to an increase in effort (Table 7.1), and the percentage is still higher (50%) when nominally denser segments are considered (Table 7.2). In ST 2 the association between high nominalisation and high processing effort is weaker, although Table 2 shows that over a quarter of the nominally denser segments lead to an increase in effort. No association between nominal density and processing effort is evident in ST 3.[7] One interesting aspect that emerges from both

[7] The vast majority of nominally dense segments in ST 3 (14 out of 22) contain no more than two nouns. Could it be that no increase in effort is associated with the nominally denser segments of the text because the overall lexical density is not high enough to exert a cumulative effect on the

tables, however, is the strong association between the presence of compounds and the increase in processing effort. As the left-most column in either table shows, almost all higher-effort segments are seen to contain a compound (for ST 2, this is the totality of segments), even when the nominal density is lower (cf. Table 7.1), which implies that even some compounds containing only two elements lead to a higher processing effort (e.g., *a silicate tetrahedron*, segment no. 63 in ST 1). In conclusion, the figures in the two tables above seem to indicate that the hot spots identified in this final step of the analysis (Figure 7.9) can be taken to be representative if not of the whole set of highly nominalised segments found in the STs at least of those segments where high nominal density comes in a highly 'packaged' form.

7.3 Back to the Research Question

It now remains to be seen to what extent the findings obtained at each step (from the analysis of editing through to the identification of the final cluster of ST hot spots in the present chapter) lend support to our initial hypothesis regarding the role of grammatical metaphor as a contributing factor in shaping ST difficulty (see 1.4).

The study set out to establish whether an increase in grammatical metaphoricity in the source texts would be associated with an increase in structural processing effort on the part of the translators participating in the study. Three STs were thus chosen which reflected a progression from a lower to a higher level of grammatical metaphoricity, largely equated with nominalisation (see 3.3). The increase in processing effort, it was further expected, would be revealed by looking at three dimensions: editing, TT variation and errors. In particular, it was expected that a higher degree of ST metaphoricity would lead to:

- increasing variation in the interim renditions proposed by translators, especially for ST-oriented edits;
- increasing inter-translator variation in the final TT renditions, especially at the structural level;

'unpacking' effort that the translators make at the syntactic level? Testing this highly speculative hypothesis goes beyond the scope of the present work. Notice, however, that once again, ST 3, i.e. the less nominalised, more "congruent" (in Hallidayan terms) specimen of the three STs, stands distinctly apart from the other two STs.

- an increasing number of TT errors associated with structural (as opposed to lexical) aspects of the STs (cf. the 'content' errors discussed in Chapter 6).

A visual representation of this hypothesis is given in Figure 7.10.

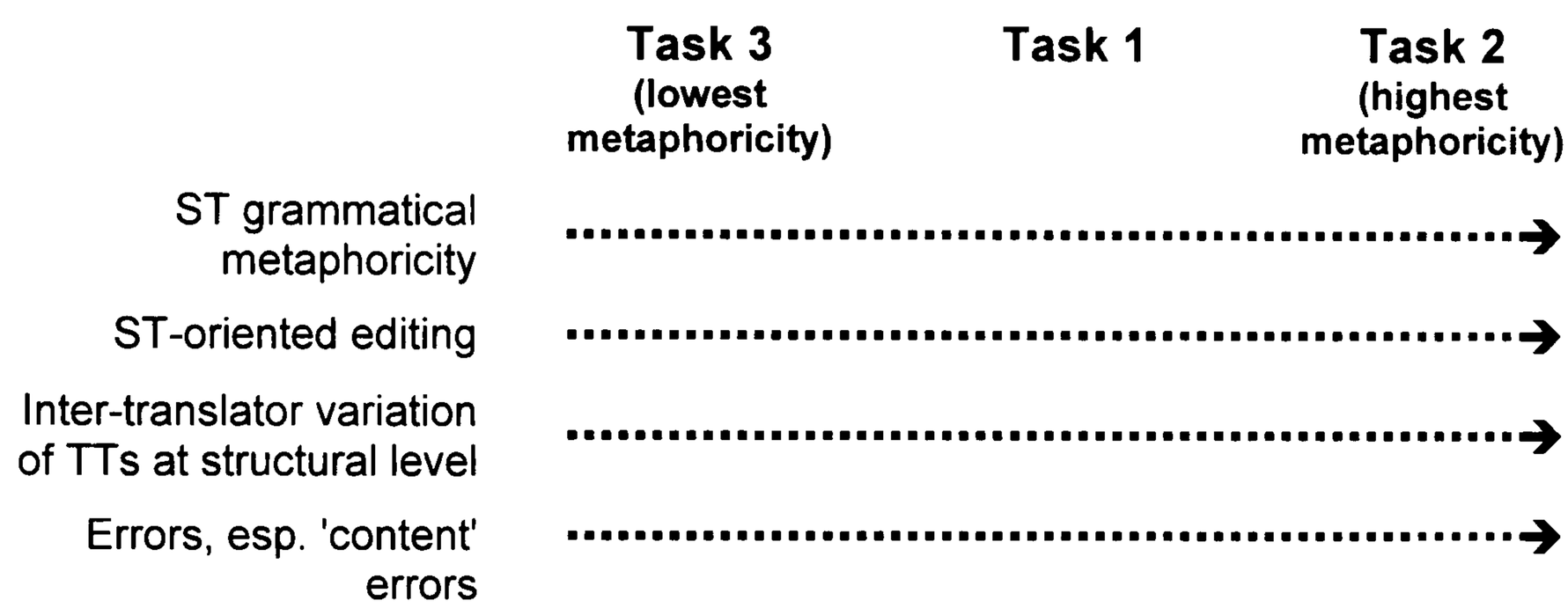

Figure 7.10 – Hypothesised association between increase in ST grammatical metaphoricity and the three dimensions observed in the study

How do the findings that emerged at each step in the analysis compare with this hypothesis? The only dimension for which a distinct pattern of association between cross-task variation and the increase in ST metaphoricity was observed was editing (Chapter 4). In particular, it emerged (cf. 4.4.1) that the increase in metaphoricity was accompanied by a parallel increase in both ST-oriented and terminological editing, while TT-oriented editing was seen to enjoy an inverse relationship with both the other editing categories and ST metaphoricity. Structural variation in the STs was thus seen to be accompanied by parallel patterns of variation in all three categories of editing effort. A clear pattern of co-variation was thus observed when ST grammatical metaphoricity and editing effort were compared.

No such distinct patterns of co-variation between the increase in metaphoricity and the observed dimension emerged in the other two steps in the analysis, i.e. the CNA conducted on the basis of the target texts (Chapter 5) and the analysis of errors (Chapter 6). According to the stated hypothesis, structural inter-translator variation would be lowest in Task 3 and highest in Task 2. By the same token, the number of content errors expected in Task 1 was lowest in Task 3 and highest in Task 2. CNA showed (cf. 5.3.1) that variation at the structural level was indeed lower in ST 3 than in STs 1 and 2 but that no further differentiation could be

made between ST 1 and ST 2. As far as errors are concerned, the analysis showed that they could be taken as signals of the difficulty linked to ST syntactic traits and that such difficulty emerged in STs 1 and 2 more than in ST 3 (cf. 6.4) but, again, no further differentiation could be made between ST 1 and ST 2 – possibly as a reflection of the fact that, along the cline of increasing metaphoricity that was considered in order to select the three STs, ST 1 was observed to be considerably closer to ST2, the most metaphorical text, than to ST 3, the least metaphorical.

In conclusion, while there was no overall confirmation of the expected patterns of co-variation between findings on processing effort and the increase in ST grammatical metaphoricity, a subset of the data (i.e. those on editing) provided partial support for the hypothesis. As regards the other two datasets, instead of the hypothesised three-step progression, the findings seem to be pointing to a polarisation that sees ST 3 at one end and ST 1 and ST 2 at the other – the former representing the pole where processing effort is lower, the latter two constituting the pole where processing effort is higher. If we take processing effort, in turn, to be a reflection of ST difficulty, we can say that, of the three texts under investigation, ST 3 proved to be the least difficult as far as structural aspects were concerned; STs 1 and 2, on their part, turned to be more difficult than ST 3, but no further differentiation between the two was possible, given the present analysis. As regards the factors that contributed to the difference in difficulty between the texts, the cluster of textual hot spots identified by the analysis points to highly nominalised, grammatically metaphorical constructs as one such factor. As was repeatedly stressed throughout the analysis (see, in particular, 6.4), the presence of such constructs in texts such ST 1 and ST 2 can be seen as one of the hurdles facing translators when, given their lack of specialist knowledge and the resulting impossibility of projecting any pre-existing knowledge onto the text, they are forced to adopt a 'bottom-up' approach to ST processing.

7.4 Some reflections on methodology

The findings which emerged at each successive analytical step of the present investigation were seen to give support to the validity of the methods employed to collect and interpret data. In particular, the CNA approach followed in studying both

editing and the target texts led, even on this small sample of subjects, to the identification of aspects that turned out to be equally prominent when a different analytical method (i.e. Error Analysis) was applied. Even so, with hindsight, a few further considerations can be made regarding some methodological aspects of the investigation.

The first set of considerations has to do with experimental design and particularly with the subjects who took part in the investigation. Their homogeneity as a group, in terms of both second-language competence and translation experience, contributes to whatever validity might be attached to the results that have been obtained, although the results clearly cannot be generalised beyond the present data. Future studies could focus on professionals who specialise in scientific translation, or professionals who specialise in non-scientific domains, or a combination of the two. Also, other language pairs might be considered in order to observe the extent to which the typological proximity/distance of the languages involved interact with the (assumed) difficulty of ST lexico-grammatical aspects.

An obvious remark to be made about the subjects of the present investigation regards the size of the sample, which is likely to be considered modest. The choice of a triangulation approach was a means of overcoming this limitation but of course a larger, randomised sample (perhaps combined with an inferential statistical analysis of the findings) could yield more generalisable results.

Further considerations relate to the selection of materials (the STs) and the choice of data types. Considering the nature of the hypothesis underlying the study, the selection of the source texts was a crucial aspect of the study. Since at least some of the data employed for the investigation (i.e. those on editing) were actually seen to yield distinct patterns of variation in relation to each ST, it might be asked whether a clearer differentiation between the texts would have led to similar patterns in the other datasets as well. In other words, the three source texts could have been chosen (or amended) so as to be located further apart from each other on the cline of increasing grammatical metaphoricity that was taken as the basis for the investigation (recall that, on this cline, ST 1 was observed to be closer to ST2, the top text, than to ST 3, which was positioned at the lower end of the cline). As an answer to the question whether "special texts" should be written for an experiment that tries to elicit information about processing at text level, the answer provided by Campbell

and Wakim (2007: 15) is "unequivocally yes, if the special text will help us to provide a more effective data set". For the texts at hand, the introduction of some changes (aimed, for instance, at artificially lowering the degree of nominalisation of ST 1 so as to further differentiate it from ST 2) might have been considered.

As regards data, a few remarks can be made either on the way the available datasets were employed or on the possibility of introducing supplementary datasets. Keystroke-based data could be analysed taking into account not only edits but also the time-coding information provided by the logs. For the present investigation, a 'flat', non-sequential representation of the editing process was constructed, starting from the two dimensions of purpose and frequency. Pauses in the writing process could, however, be profitably combined with a study of the interim solutions. In particular, the pause-to-keyboarding ratio (cf. O'Brien 2006) could be observed, so as to see whether it differs significantly for segments which emerged as difficult in the analysis of the other data sets. Recent research (e.g. Breedeveld 2002) has shown how descriptions of the translation process should probably also take into account its inherent dynamism, such that the activities carried out by translators during the process (reading the ST, writing the TT, evaluating the TT) do not occur in successive, compartmentalised stages. Rather, they are seen to recur each time the translator runs through the text and, more importantly, they end up being "functionally dependent" (Breedveld 2002: 231-233) on one another, meaning essentially two things. On the one hand, an activity is attributed a certain function according to the moment when it is performed. Reading the TT, for instance, may have the purpose of assessing semantic equivalence with the ST when performed at an initial stage in the process, but it may be aimed at assessing the fluency of the translated text if performed towards the end of the process. On the other hand, an activity performed at a certain moment may entail another activity (or its absence) later on in the process. Such aspects may also have to be taken into account at the moment of identifying textual hot spots. For instance, considering the *moment* when one edit occurs could be seen as relevant. If an interim rendition is only changed by the translator during the final run-through, the change could be attributed either to her careful re-reading of the ST or to an assumption made on the basis of the 'text-internal' logic of the TT (an assumption that may or may not resolve the translator's uncertainty as regards the exact meaning of that particular ST segment).

As a supplementary dataset, Think-Aloud Protocols and retrospective interviews on log data could also be considered (for the purposes of this investigation, which already included 3 complementary data sets, they were deemed too time-consuming and work-intensive). Once an appropriate set of "problem indicators" (Krings 1986: 137; cf. also Bernardini 2001) had been identified, it would be interesting, for instance, to see how introspective or retrospective data could enrich our understanding of what translators find difficult and why in relation to what we have called textual hot spots.

7.5 Implications for training, evaluation and future research

As pointed out in Chapter 1, in the context of translator training, source-text classification schemes are rarely related to the degree of difficulty of the texts. This is, however, an aspect of fundamental importance if the criteria for text selection are to aspire to some kind of pedagogical progression. The findings of the present investigation give some indications as to how the structural profiles of source texts could be selected (or manipulated) so as to obtain texts displaying increasing degrees of difficulty, at least from a prospective point of view. In particular, aspects such as lexical and nominal density, the presence of compounds and the length of clauses or nominal groups (in short, the various features subsumed under the concept of grammatical metaphor) can all be taken into account at the moment of choosing source texts to submit to trainees, either for practice or for assessment purposes.

This leads naturally to the link between text difficulty and evaluation. In contexts where translation competence must be assessed (for training but also for accreditation purposes) the need to categorise source texts according to degrees of difficulty is evident, but no universally accepted parameters have been established. Research on the sources of difficulty can provide a basis for establishing such parameters, although it should not be forgotten that the source text as text is bound to remain one element only of the difficulty posed by a translation task. Further investigation of how text-external elements interact with the lexico-grammatical features of source texts would be necessary for the development of reliable 'difficulty indicators'.

From the specific point of view of research, however, the only viable approach for the identification of the elements contributing to the difficulty of a translation task seems to be one that tackles such elements one by one, while building on assumptions derived from previous, related research (and testing these assumptions in the process). For an aspect such as difficulty, in other words, the only feasible methodological procedure appears to be one based on the formulation of "more and more small-scale, and hence more testable hypotheses" (Toury 1995: 239). Seen in this perspective, the present investigation would be but a small brick of a larger analytical edifice devoted to studying what makes texts difficult to translate.

7.6 Concluding remarks

The present study set out to investigate the translation of scientific texts (in a specific language pair) with the aim of shedding light on one particular aspect, i.e. the difficulty linked to certain ST structural features, that has received scant attention on the part of researchers in the past – most probably, on account of the high variability of the factors involved in ST processing by translators. Choosing to focus on one specific lexico-grammatical aspect (grammatical metaphor), the investigation has shown how it played a far from marginal role in contributing to the degree of difficulty of the source texts used for the experiment – texts belonging to a genre where terminology *per se* is usually presented as the main and foremost source of difficulty for lay readers. The analyses based on the Choice Network approach and their integration with results from an Error Analysis of the target texts, have helped identify difficulty 'hot spots' in the source text where terminology does not seem to play a decisive role. These ST segments are often seen to display one or more of the features that Halliday sees as characteristic of scientific discourse, most of them subsumed under the label of grammatical metaphor. Their role in contributing to text difficulty as far as translation is concerned should perhaps be given more attention than it has received so far, which could have implications for areas such as translator training and translation evaluation.

The reassessment of the relative weight of terminology and lexico-grammar that the findings point to should not necessarily be seen from a polarising perspective. Indeed, what the Hallidayan analysis of scientific discourse distinctly

shows is how the two aspects are intertwined. Terminology is not a linguistic resource developing on its own, irrespective of the meaning-making mechanisms of non-terminological resources. It is, rather, a *systemic* effect of the very same mechanisms that scientific discourse resorts to with higher frequency than other discourses.

Appendix 1: Source Texts used for the study

Source Text used for Task 1

The rate of crack growth depends not only on the chemical environment but also on the magnitude of the applied stress. The development of a complete model for the kinetics of fracture requires an understanding of how stress accelerates the bond-rupture reaction. In the absence of stress, silica reacts very slowly with water. The dissociative reaction we have discussed causes the surface of silica glass to dissolve in water at a rate of 10^{-17} meter per second; stress-free silicon-oxygen surface bonds are so unreactive toward water that they will not even adsorb water vapor. Yet the application of stress can cause cracks to grow at speeds greater than one millimeter per second.

A crack tip focuses stress in much the same way as a needle point focuses an electric field. The closer one approaches the crack tip, the higher the stress becomes, until it reaches several million pounds per square inch within a few atomic dimensions of the tip. Exposed to these large stresses, the atomic structure of silica becomes distorted from its normal bonding configuration. Theoretical calculations show that if a silicate tetrahedron is distorted by pulling on the corner oxygen atoms, the silicon atom in the center is more likely to bond with a water molecule. Furthermore, the chemical interaction with water will decrease the force required to distort the silicon-oxygen bonds further.

Source: Michalske T.A and Bunker B.C. (1987) The fracturing of glass. *Scientific American*, December, 78-85.

Source Text used for Task 2

Phenomenological models have been developed to address the mechanism of stress/environment dependent crack extension. These models assume that the crack-tip bond rupture event is facilitated by reactions with species from the surrounding environment so as to increase // the rate of macroscopic crack extension.This view of the fracture process requires that the rate of chemical reaction at the crack tip increase as the stress on the crack-tip bond increases.For the specific case of silica glass fractured in water, these phenomenological models assume that the rate of hydrolysis of siloxane bonds will increase with applied stress.

In a previous study we showed that the strain-enhanced chemical reactivity of silicate materials can be examined through the use of model compounds containing small siloxane rings. Small siloxane ring structures produce constrained Si-O-Si bonding geometries that store strain energy; the smaller the siloxane ring, the greater the stored strain energy per atomic bond. By measuring the kinetics of hydrolysis for different size siloxane ring structures, we have shown that the rate of siloxane hydrolysis increases approximately 8 orders of magnitude for bond strains equal to about one-sixth the Si-0 bond energy. In this paper, we use siloxane ring hydrolysis results to develop a quantitative chemical-kinetics-based model to predict the rate of crack extension in silica glass in humid conditions.

Source: Michalske T.A and Bunker B.C. (1993) A Chemical Kinetics Model for Glass Fracture. *Journal of the American Ceramic Society* 76, 10, 2613-2618.

Source Text used for Task 3

Cracks in glass grow continuously at controlled rates ranging from less than one-trillionth of an inch per hour to roughly half the speed of sound when glass shatters. The rate of crack propagation depends on the applied stress and the chemical environment.

Cracks develop and grow when the silicon-oxygen bonds are broken by stress. The tip of the crack is roughly the size of the opening created when the ring structure is broken, about 0.5 nanometers. The amount of energy required to cleave the silicon-oxygen bond decreases by a factor of twenty in the presence of water. This indicates that the water molecule fits into the crack tip and converts a silicon-oxygen bond into two silanol (SiOH) groups. Other compounds, such as ammonia, can also facilitate crack rates. Ammonia is actually more reactive than water with strained silicon-oxygen bonds. If the nucleophilic molecule gets larger than 0.5 nanometers, the effect is no longer experienced. Water and ammonia have molecular sizes of about 0.3 nanometers.

Two factors, stress and the chemical environment, affect the durability of glassware. Stress, in glass that is worked, occurs when the glass cools unevenly. Uneven cooling is unavoidable because glass does not conduct heat well and the surface cools more rapidly than the interior. Commercial glass blowers use ovens (lehrs) in which they first heat the object to the annealing temperature (slightly below the softening temperature) and then cool the glassware slowly to remove any internal strains. This process is called annealing.

Source: Banks, R. (nd) Fracturing Glass. In *A Short History of Glassblowing and Glassblowing Techniques*. Available: http://chemistry.boisestate.edu/rbanks/glassblowing/glassblowing_history.htm [Accessed 20 April 2006].

Appendix 2: Segmented Source Texts and explanatory notes

Notes on the criteria for segmentation

The three STs used in the study were segmented in order to identify units for analysis which were relevant in relation to the research question. In other words, the segmentation was meant to contribute to operationalising the hypothesis the study investigates, which is that grammatical metaphor can play a role in increasing the degree of difficulty of a text for translation. Grammatical metaphor is the concept introduced by Halliday to illustrate variation in markedness at the level of expression. A given form can be said to be the unmarked expression of a certain meaning, whereas other forms (entailing grammatical but often also lexical variation) are described as the marked expressions of that same meaning. Marked forms are the result of what Halliday calls grammatical metaphor. In scientific text, grammatical metaphor is most often realised as nominalisation and therefore criteria were sought that could help identify nominal units in the STs used in the study. Alongside such units, however, other constituents had to be identified, so as to allow a comparison with data and findings obtained for the former.

Given that the main hypothesis guiding the study is formulated in Hallidayan terms, it seemed only natural to adopt a Hallidayan framework for the segmentation of the STs. In particular, *'ranked' constituent analysis* (Halliday 1994: 22-24; Eggins 1994: 123-143) was felt to provide a relevant and useful method for ST segmentation. Constituent analysis is aimed at illustrating how the language is structured at the lexico-grammatical level, which is the intermediate level of grammatical coding between that of meaning and that of sounds or letters. It is based on two principles (Eggins 1994: 123). The first says that the lexico-grammar is composed by units of different kinds. The second says that these units are related to each other through "constituency – i.e. that smaller units make up bigger units, and bigger units are made up of smaller units" (Eggins 1994: 123). Being concerned with the grammatical, as opposed to the discourse-semantic, stratum, constituent analysis is limited by two general characteristics (Eggins 1994: 129): "1) it relates items of the same kind to each other (e.g. clauses to clauses, words to words, phrases to phrases, etc.); and 2) it relates items that are adjacent o nearly adjacent, to each other". This means that a unit such as "text" is left out from the grammatical rank scale, in that it is best regarded as a semantic unit rather than a unit of form (cf. also Halliday and Hasan 1976: 2).

The rank scale at the grammatical stratum is as listed in Table 1, reproduced from Eggins (1994: 129).

Table 1 – Rank scale at the lexico grammatical level (see Eggins 1994: 129)

clause – clause complex
group/phrase
word
morpheme

The rank scale in the table defines units for description and analysis. In identifying such units the aim is that of singling out units that carry different types of linguistic patterns,

that is "different structural organizations, or different structural configurations" (Eggins 1994: 129). Note, also, that units are labelled in terms of their class membership. i.e. what they are on their own, rather than in terms of their role in relation to the whole at each rank (cf. Halliday 1994: 25).

In the present context, the focus is especially on clauses and groups/phrases, seen as the units where lexico-grammatical variation across the three STs is predominantly manifested (see 3.4.3 for details). The clause is the highest unit of grammatical analysis. It is the largest grammatical constituent and is made up by the other units at each of the lower ranks. Groups and phrases represent the rank coming immediately below the clause. More specifically, a group is "a combination of words built up on the basis of a particular logical relation" (Halliday 1994: 180). There are three main classes of group: nominal, verbal and adverbial. Other, minor classes are conjunctions and preposition groups. A phrase is "different from a group in that, whereas a group is an expansion of a word, a phrase is a contraction of a clause" (Halliday 1994: 180). Groups and phrases start from opposite ends but "achieve roughly the same status on the rank scale, as units that lie somewhere intermediate between the rank of a clause and that of a word" (Halliday 1994: 180).

Some classes of phrases, for instance prepositional phrases, may contain nominal groups (e.g. segment 59 in ST 1 below: *from its normal bonding configuration*). They are an example of a unit of one rank that is made up of a constituent of the same rank – "[i]t is because the prepositional phrase has this more complex embedded structure that we call it a phrase rather than a group" (Eggins 1994: 132-133). These prepositional phrases are constituents at the first rank of the clause. In some cases, however, prepositional phrases operate *within* a nominal group to postmodify or qualify the head noun of the group. Such prepositional phrases are not constituents at the first rank of the clause but at the next rank down, i.e. that of phrase/group. They function within a unit at this lower rank and not at clause rank. An extreme example of such a prepositional phrase is segment 4 in ST 1: *the development of a complete model for the kinetics of fracture*, a self-contained segment which can be seen as resulting from an "iterated rankshift" (nominal group in prepositional phrase in nominal group in ...), i.e. *the development* [*of* [*a complete model* [*for* [*the kinetics* [*of* [*fracture*]]]]]] (cf. Halliday 1993a: 56, where this same example is discussed).

In the three tables below, groups/phrases are only indicated as segments and numbered progressively in each table. This is because the segmentation has essentially the aim of providing units for analysis, so as to provide a sort of 'reference grid'. In particular, the STs are segmented so as to verify whether findings based on the various source of data used in the study cluster in or around specific units of a ST. A more detailed characterisation of such units is given in Chapter 7 (see 7.2.3).

The other units listed in Table 1 are clause complexes, words and morphemes. The notion of clause complex replaces the folk-linguistic notion of "sentence"; it is placed next to, not above, the clause. This is because the relationship between two clauses in a clause complex is not seen as a constituency relationship but, rather, as a "logical" structure (cf. Halliday 1994: 216). Clause complexes have mainly been used here as one of the bases for measuring the degree of grammatical metaphoricity in the STs (see 3.4.3). As for the other units (words and morphemes), they are not considered in the segmentation as the present investigation does not pursue analysis at this level.

Finally, note that some clauses are not counted separately but are considered as part of a segment at group/phrase rank (in the tables below they are enclosed within square brackets). In constituent analysis these are called "embedded clauses", meaning that they come to function within the structure of a group.

Segmentation of Source Text 1

Clause complex	Clause	Segment	
1	1	1	The rate of crack growth
		2	depends
		3	not only on the chemical environment
		4	but also on the magnitude of the applied stress.
2	2	5	The development of a complete model for the kinetics of fracture
		6	requires
		7	an understanding of
	3	8	how
		9	stress
		10	accelerates
		11	the bond-rupture reaction.
3	4	12	In the absence of stress,
		13	silica
		14	reacts
		15	very slowly
		16	with water
4	5	17	The dissociative reaction [we have discussed]
		18	causes
		19	the surface of silica glass
	6	20	to dissolve
		21	in water
		22	at a rate of 10^{-17} meter per second;
5	7	23	stress-free silicon-oxygen surface bonds
		24	are
		25	so unreactive
		26	toward water
	8	27	that they
		28	will not even adsorb
		29	water vapour
6	9	30	Yet
		31	the application of stress
		32	can cause
		33	cracks
	10	34	to grow
		35	at speeds greater than one millimeter per second.
7	11	36	A crack tip
		37	focuses
		38	stress
	12	39	in much the same way as
		40	a needle point
		41	focuses
		42	an electric field.
7	13	43	The closer
		44	one

		45	approaches
		46	the crack tip,
	14	47	the higher
		48	the stress
		49	becomes,
	15	50	until
		51	it
		52	reaches
		53	several million pounds per square inch
		54	within a few atomic dimensions of the tip.
9	16	55	Exposed
		56	to these large stresses,
	17	57	the atomic structure of silica
		58	becomes distorted
		59	from its normal bonding configuration.
10	18	60	Theoretical calculations
		61	show that
	19	62	if
		63	a silicate tetrahedron
		64	is distorted
	20	65	by pulling
		66	on the corner oxygen atoms
	21	67	the silicon atom in the center
		68	is
		69	more likely to bond
		70	with a water molecule.
11	22	71	Furthermore,
		72	the chemical interaction with water
		73	will decrease
		74	the force [required to distort the silicon-oxygen bonds further].

Segmentation of Source Text 2

Clause complex	Clause	Segment	
1	1	1	Phenomenological models
		2	have been developed
	2	3	to address
		4	the mechanism of stress/environment dependent crack extension.
2	3	5	These models
		6	assume that
	4	7	the crack-tip bond rupture event
		8	is facilitated
		9	by reactions with species from the surrounding environment
	5	10	so as
		11	to increase
		12	the rate of macroscopic crack extension.
3	6	13	This view of the fracture process
		14	requires that
	7	15	the rate of chemical reaction at the crack tip
		16	increase
	8	17	as
		18	the stress on the crack-tip bond
		19	increases.
4	9	20	For the specific case of silica glass [fractured in water],
		21	these phenomenological models
		22	assume that
	10	23	the rate of hydrolysis of siloxane bonds
		24	will increase
		25	with applied stress.
5	11	26	In a previous study,
		27	we
		28	showed that
	12	29	the strain-enhanced chemical reactivity of silicate materials
		30	can be examined
		31	through the use of model compounds [containing small siloxane rings].
6	13	32	Small siloxane ring structures
		33	produce
		34	constrained Si-O-Si bonding geometries [that store strain energy];
7	14	35	the smaller
		36	the siloxane ring,
	15	37	the greater
		38	the stored strain energy per atomic bond.
8	16	39	By measuring
		40	the kinetics of hydrolysis
		41	for different size siloxane ring structures,
	17	42	we

		43	have shown that
	18	44	the rate of siloxane hydrolysis
		45	increases
		46	approximately 8 orders of magnitude
		47	for bond strains equal to about one-sixth the Si-0 bond energy.
9	19	48	In this paper,
		49	we
		50	use
		51	siloxane ring hydrolysis results
	20	52	to develop
		53	a quantitative chemical-kinetics-based model
	21	54	to predict
		55	the rate of crack extension in silica glass in humid conditions.

Segmentation of Source Text 3

Clause complex	Clause	Segment	
1	1	1	Cracks in glass
		2	grow
		3	continuously
		4	at controlled rates
	2	5	ranging
		6	from less than one-trillionth of an inch per hour
		7	to roughly half the speed of sound
	3	8	when
		9	glass
		10	shatters.
2	4	11	The rate of crack propagation
		12	depends
		13	on the applied stress
3		14	and the chemical environment.
	5	15	Cracks
		16	develop and grow
	6	17	when
		18	the silicon-oxygen bonds
		19	are broken
		20	by stress.
4	7	21	The tip of the crack
		22	is
		23	roughly
		24	the size of the opening [created when the ring structure is broken]
		25	about 0.5 nanometers.
5	8	26	The amount of energy [required to cleave the silicon-oxygen bond]
		27	decreases
		28	by a factor of twenty
		29	in the presence of water.
6	9	30	This
		31	indicates that
	10	32	the water molecule
		33	fits
		34	into the crack tip
	11	35	and converts
		36	a silicon-oxygen bond
		37	into two silanol (SiOH) groups.
7	12	38	Other compounds,
		39	such as ammonia,
		40	can also facilitate
		41	crack rates.
8	13	42	Ammonia
		43	is
		44	actually
		45	more reactive

		46	than water
		47	with strained silicon-oxygen bonds.
9	14	48	If
		49	the nucleophilic molecule
		50	gets
		51	larger than
		52	0.5 nanometers,
	15	53	the effect
		54	is [...] experienced
		55	no longer
10	16	56	Water and ammonia
		57	have
		58	molecular sizes of about 0.3 nanometers.
11	17	59	Two factors, stress and the chemical environment,
		60	affect
		61	the durability of glassware.
12	18	62	Stress,
		63	in glass [that is worked],
		64	occurs
	19	65	when
		66	the glass
		67	cools
		68	unevenly.
13	20	69	Uneven cooling
		70	is
		71	unavoidable
	21	72	because
		73	glass
		74	does not conduct
		75	heat
		76	well
	22	77	and the surface
		78	cools
		79	more rapidly
		80	than the interior
14	23	81	Commercial glass blowers
		82	use
		83	ovens (lehrs)
	24	84	in which
		85	they
		86	first
		87	heat
		88	the object
		89	to the annealing temperature
		90	(slightly
		91	below the softening temperature)
	25	92	and then
		93	cool
		94	the glassware
		95	slowly
	26	96	to remove

		97	any internal strains.
15	27	98	This process
		99	is called
		100	annealing.

Appendix 3: Binary and non-binary errors in the target texts

Errors in the texts are highlighted as follows:
- subcategory 1 of binary errors ('inaccurate lexical equivalence') is indicated by grey highlighting, e.g. la reazione della frattura dei legami
- subcategory 2 of binary errors ('inaccurate ST content transfer') is indicated by double underlining, e.g. Il livello di fragilità
- non-binary errors are indicated by dotted underlining, e.g. quando il vetro va in mille pezzi

Note that each error is highlighted at its first occurrence only in each individual text. Other types of errors may be found in the text beyond those included in these categories (e.g. spelling errors or formatting inconsistencies). Such errors are not highlighted.

Translations of Source Text 1

Translator: CB

Il livello di fragilità del vetro non dipende solo da fattori chimici, ma anche dalla misura in cui si esercita la pressione su di esso. Lo sviluppo di un modello completo per l'enegia cinetica riguardante la frattura del vetro richiede un'ampia conoscenza di come la pressione possa accelerare la reazione della frattura dei legami.
Nel caso in cui non venga esercitata pressione, la silice reagisce molto lentamente a contatto con l'acqua. La reazione dissociativa discussa in precedenza porta la superficie di vetro di silice a dissolversi nell'acqua secondo un rapporto di 10 metri al secondo. I legami di superficie silicio-ossigeno, che non sono stati sottoposti a pressione, non presentano nessuna reazione all'acqua e non assorbono nemmeno il vapore acqueo. Tuttavia il fatto di esercitare pressione può portare a fenditure che si formano ad una velocità che supera il millimetro al secondo.
L'estremità della fenditura esercita la pressione in modo molto simile alla punta di una ago che converge verso un campo elettrico. Più ci si avvicina all'estremità della fenditura, più alta è la pressione, fino a raggiungere milioni di libbre per ogni pollice quadrato in alcune dimensioni atomiche dell'estremità. Essendo esposta a tali pressioni la struttura atomica della silicie risulta alterata rispetto alla sua normale configurazione di legame. Secondo alcuni calcoli teorici, se un tetraedro silicato viene alterato spingendo atomi di ossigeno nell'angolo, è possibile che l'atomo di silicio al centro si leghi con una molecola d'acqua. Inoltre l'interazione chimica con l'acqua diminuirà la forza necessaria per alterare ulteriormente i legami silicio-ossigeno.

Translator: MS

La velocità di espansione della frattura dipende non soltanto dall'ambiente chimico ma anche dalla quantità di stress applicato. Lo sviluppo di un modello completo per lo

studio della cinetica legata alla frattura richiede una comprensione di come lo stress acceleri la reazione legame-rottura (degli atomi).
In assenza di stress, la silice reagisce molto lentamente con l'acqua. La reazione dissociativa di cui abbiamo discusso fa sì che la superficie del vetro di silice si dissolva nell' acqua ad una velocità di 10-17 metri al secondo; laddove i legami silicone-ossigeno privi di stress risultano essere talmente non reattivi al contatto con l'acqua da non assorbire neppure il vapore acqueo. Tuttavia l'applicazione di stress può provocare l'aumento della velocità di fratturazione che supera il millimetro al secondo.
Il punto di rottura fa convergere lo stress allo stesso modo in cui la punta di un ago converge verso un campo elettrico. Più il soggetto si avvicina al punto di rottura più aumenta il livello di stress, fino a raggiungere diversi milioni di chili per pollice quadrato all'interno delle poche dimensioni atomiche del punto d'origine. Esposta a stress così notevoli, la struttura atomica della silice risulta distorta rispetto alla sua normale configurazione di legame. Calcoli di tipo teorico dimostrano che se un tetraedro di silice viene distorto con una tensione ai lati degli agli atomi di ossigeno, l'atomo di silicone che è collocato al centro della struttura ha maggiori probabilità di legarsi con la molecola dell'acqua. Inoltre, l'interazione chimica con l'acqua provocherà una diminuzione della forza necessaria per modificare ulteriormente i legami silicone-ossigeno.

Translator: MT

il tasso di crescita della frattura dipende non soltanto dalle condizioni chimiche, ma anche dalla [OMISSION] pressione esercitata. lo sviluppo di un modello completo di cinetica della frattura richiede la comprensione del modo in cui la pressione accelera la reazione di rottura dei legami.
in assenza di pressione, la silice reagisce molto lentamente con l'acqua. la reazione dissociativa, di cui si è parlato, provoca lo scioglimento della superficie del vetro di quarzo in acqua ad una velocità di 10-17 metri al secondo; i legami silicio-ossigeno di superficie, non sottoposti a pressione, non reagiscono con l'acqua e non assorbono il vapore. esercitando una certa pressione, tuttavia, è possibile che le fratture aumentino a velocità superiori a un millimetro al secondo. l'estremità di una frattura focalizza la pressione così come la punta di un ago focalizza un campo elettrico. più ci si avvicina all'estremità della frattura, più grande sarà la pressione, che potrà raggiungere alcuni milioni di libbre per pollice quadrato all'interno delle dimensioni atomiche della frattura (?). esposta a tali pressioni, la struttura atomica della silice cambia la normale configurazione dei legami. calcoli teorici dimostrano che se un tetraedro di silicato viene deformato spostando in un angolo gli atomi di ossigeno, è molto probabile che l'atomo di silicio al centro si leghi con una molecola d'acqua. l'interazione chimica con l'acqua, inoltre, diminuirà la forza richiesta per deformare ulteriormente i legami silicio-ossigeno.

Translator: SC

La rapidità con cui si propagano le incrinature non dipende solo dalle condizioni chimiche ma anche dall'intensità della tensione esercitata. Per elaborare un modello esaustivo della cinetica della frantumazione è necesassario comprendere come le tensioni accelerano la reazione di rottura dei legami. In assenza di tensioni, il silice reagisce molto lentamente con l'acqua. La reazione dissociativa di cui abbiamo parlato fa sì che la superficie di un vetro di silice si sciolga in acqua alla velocità di 10-17 metri al

secondo, i legami di superficie silicio-ossigeno liberi da tensioni reagiscono talmente poco con l'acqua che non ne assorbono nemmeno il vapore, ma il fatto di esercitare tensione può far sì che le incrinature aumentino a velocità superiori ai tre millimetri al secondo. L'estremità dell'incrinatura fa convergere la tensione nello stesso modo in cui la punta di un ago la fa convergere in un campo elettrico: più ci si avvicina all'estremità, più aumenta la tensione, fino a che non raggiunge qualche milione di kilogrammi per metro quadro ad una piccolissima distanza dall'estremità dell'incrinatura.Quando è esposta a tensioni di tale entità, la struttura atomica del silice modifica la propria configurazione atomica normale. Calcoli teorici dimostrano che se un tetraedro di silicato si modifica perché gli atomi di ossigeno vengono spinti in un angolo, è più probabile che l'atomo di silicio che si trova al centro si leghi con una molecola d'acqua. Inoltre l'interazione chimica con l'acqua farà diminuire la forza richiesta per modificare ulteriormente il legami silicio-ossigeno.

Translator: VM

La velocità con cui una crepa si ingrandisce non dipende solo dall'ambiente chimico, ma anche dall'intensità della pressione applicata. Per sviluppare un modello completo per la cinetica di frattura bisogna capire come la pressione accelera la reazione tramite la quale i legami si rompono.
In assenza di pressione, la silice reagisce molto lentamente con l'acqua. Tramite le reazioni [OMISSION] di cui si è parlato, la superfice del vetro di silice si scioglie in acqua alla velocità di 10-17 metri al secondo; i legami di superficie tra silicio e ossigeno sui quali non viene applicata alcuna pressione non reagiscono con l'acqua e non ne assorbono neppure il vapore. Tuttavia, qualora venga applicata pressione, le incrinature si possono ingrandire a velocità maggiori di un millimetro al secondo.
Un'estremità della crepa concentra la pressione nello stesso modo in cui la punta di un ago concentra un campo elettrico: più ci si avvicina all'estremità della crepa, più la pressione aumenta, fino a raggiungere diversi milioni di bar in alcune dimensioni atomiche dell'estremità. Esposta a queste pressioni intense, la struttura atomica della silice devia dalla normale configurazione della sua formazione di legame.
Calcoli teorici dimostrano che se un tetraedro contenuto in un silicato viene deformato estraendo atomi di ossigeno da un angolo, è più probabile che l'atomo di silicio al centro si leghi con una molecola d'acqua. Inoltre, l'interazione chimica con l'acqua provoca una diminuzione della forza necessaria per deformare ulteriormente i legami tra il silicio e l'ossigeno.

Translations of Source Text 2

Translator: CB

I modelli fenomenologici sono stati sviluppati per ricorrere al meccanismo riguardante l'estensione della fenditura che dipende dal rapporto pressione/ambiente. Tali modelli presuppongono che la rottura del legame fenditura-estremità sia facilitata da reazioni [OMISSION] provocate dall'ambiente circostante in modo tale da aumentare il livello macroscopico della fenditura. Questa considerazione sul processo di frattura richiede

che il livello della reazione chimica all'estremità della fenditura sia direttamente proporzionale alla pressione esercitata sul legame fenditura-estremità. Nel caso specifico del vetro di silice fratturato nell'acqua, tali modelli fenomenologici presuppongono che il livello di idrolisi nei legami di silossano aumenterà con la pressione esercitata.
In un precedente studio si è dimostrato che la reattività chimica che aumenta la deformazione dei materiali di silicato possa essere analizzata attraverso l'utilizzo di componenti [OMISSION] che contengono piccoli anelli di silossano. Le strutture dei piccoli anelli di silossano producono geometrie ristrette del legame Si-O-Si che immagazzinano energia di deformazione; quanto più piccolo è l'anello di silossano, tanto più grande è l'energia di deformazione immagazzinata per ogni legame atomico. Misurando la cinetica dell'idrolisi per diverse strutture di anelli di silossano, si è dimostrato che il livello di idrolisi di silossano aumenta di circa 8 ordini di grandezza per deformazioni di legame, pari a circa un sesto dell'energia di legame Si-O. In questa relazione si utilizzano i risultati riguardanti l'idrolisi degli anelli di silossano per sviluppare un modello quantitativo basato sul rapporto chimico-cinetico al fine di prevedere il livello dell'estensione della fenditura nel vetro di silice in condizioni di umidità.

Translator: MS

Alcuni modelli fenomenologici sono stati sviluppati con lo scopo di studiare il meccanismo di dipendenza dell' estensione di una frattura rispetto allo sforzo esercitato su di essa e rispetto all' ambiente. Tali modelli assumono che la rottura del legame chimico nel punto di frattura venga facilitato da reazioni con specie chimiche provenienti dall' ambiente circostante in modo da aumentare la velocità di estensione della frattura macroscopica. Questa interpretazione del processo di frattura vuole che la velocità di reazione chimica nel punto di frattura aumenti proporzionalmente allo sforzo esercitato sul legame chimico presente nel punto di frattura stesso. Nel caso specifico del vetro di silice fratturato nell' acqua questi modelli fenomenologici prevedono che la velocità di idrolisi dei legami di silossano aumentino con l' applicazione di uno sforzo.
In un lavoro pubblicato precedentemente, avevamo dimostrato che la reattività chimica intensificata dalla deformazione di materiali silicati può essere esaminata attraverso l' uso di composti modello contenenti anelli di silossano di piccole dimensioni. Le strutture ad anello di piccole dimensioni del silossano producono vincolate geometrie di legame Si-O-Si che conservano energia di deformazione; più piccolo è l' anello di silossano maggiore è l'energia conservata per legame atomico. Misurando la cinetica dell' idrolisi per le diverse dimensioni delle strutture ad anello del silossano, abbiamo dimostrato che la velocità di idrolisi del silossano aumenta di circa 8 ordini di grandezza per deformazioni di legame pari a circa un sesto dell' energia di legame Si-O. In questo lavoro vengono usati i risultati relativi all' idrolisi dell' anello di silossano per sviluppare un modello quantitiativo ottenuto su basi chimico-cinetiche per prevedere la velocità di propagazione di frattura nel vetro di silice in condizioni umide.

Translator: MT

Sono stati sviluppati modelli fenomenologici per indirizzare il meccanismo di propagazione della cricca dipendente dal rapporto sforzo/ambiente. In base a tali modelli la rottura del legame cricca-apice è facilitato dalle reazioni con specie

provenienti dall'ambiente circostante che permettono di aumentare il tasso di propagazione della cricca macroscopica. Questa analisi del processo di frattura implica che il tasso di reazione chimica in corrispondenza dell'apice della cricca cresce all'aumentare dello sforzo sul legame cricca-apice. Nel caso specifico del vetro di silice fratturato in acqua, secondo tali modelli fenomenologici il tasso di idrolisi dei legami di silossano aumenterà se verrà applicato uno sforzo.
In uno studio precedente abbiamo dimostrato che la reattività chimica aumentata dalla deformazione dei materiali a base di silicato può essere esaminata mediante l'utilizzo di composti modello contenenti piccoli anelli di silossano. La struttura dei piccoli anelli di silossano dà vita a geometrie di legame Si-O-Si forzate che accumulano l'energia di deformazione; più è piccolo l'anello di silossano, maggiore sarà l'energia di deformazione per legame atomico accumulata. Misurando la cinetica dell'idrolisi relativa a strutture di anelli di silossano di diversa grandezza, abbiamo dimostrato che il tasso di idrolisi del silossano aumenta di circa 8 ordini di grandezza per deformazioni di legame pari a circa un sesto dell'energia di legame Si-O. Nel presente studio faremo riferimento ai risultati dell'idrolisi degli anelli di silossano per sviluppare un modello cinetico-chimico quantitativo per predire il tasso di propagazione della cricca nel vetro di silice in condizioni di umidità.

Translator: SC

Sono stati sviluppati dei modelli fenomenologici per spiegare il meccanismo dell'estensione dell'incrinatura che dipende da condizioni ambientali o di sollecitazione. Tali modelli partono dal presupposto secondo cui il fenomeno della rottura dei legami all'estremità dell'incrinatura sia facilitato dalle reazioni con specie chimiche presenti nell'ambiente circostante in modo da aumentare l'estensione dell'incrinatura macroscopica. Questo modo di intendere il processo di fratturazione presuppone che la quantità di reazioni chimiche all'estremità dell'incrinatura diventi maggiore con l'aumentare della sollecitazione sui legami all'estremità dell'incrinatura. Nel caso specifico del vetro di silice fratturato in acqua, questi modelli fenomenologici partono dal presupposto secondo cui il tasso di idrolisi dei legami di silossano aumenti in seguito all'applicazione di una sollecitazione.
In uno studio precedente abbiamo dimostrato che la reattività chimica dei silicati aumentata in seguito a una sollecitazione può essere analizzata attraverso l'impiego di composti base contenenti anelli di silossano di piccole dimensioni. Queste strutture creano legami del tipo Si-O-Si che immagazzinano l'energia prodotta da una sollecitazione: più piccole sono le dimensioni dell'anello di silossano, maggiore è la quantità di energia immagazzinata in ogni legame atomico. Dopo aver misurato la cinetica dell'idrolisi in strutture contenenti anelli di silossano di diverse dimensioni, abbiamo dimostrato che il tasso di idrolisi del silossano aumenta di circa 8 ordini di magnitudine in caso di sollecitazioni sui legami pari a circa un sesto dell'energia di legame presente in strutture Si-O. Nel corso di questo studio, utilizzeremo i risultati ottenuti dall'idrolisi degli anelli di silossano al fine di elaborare un modello quantitativo basato sulla cinetica chimica per prevedere l'estensione dell'incrinatura del vetro di silice in condizioni di umidità.

Translator: VM

Sono stati sviluppati modelli fenomenologici per indirizzare il meccanismo di estensione della cricca dipendente dalla tensione o dall'ambiente. Questi modelli implicano che l'evento di rottura del legame della punta della cricca sia facilitato da reazioni con specie provenienti dall'ambiente circostante così da aumentare il tasso dell'estensione macroscopica della cricca. Questa visione del processo di frattura richiede un aumento del tasso di reazione chimica sulla punta della cricca all'aumentare della tensione sul legame della punta della cricca. Nel caso specifico del vetro di silice fratturato in acqua, questi modelli fenomenologici implicano che il tasso di idrolisi dei legami di silossano aumenti con la tensione applicata.
In uno studio precedente abbiamo mostrato che la reattività chimica aumentata dalla deformazione dei silicati può essere esaminata attraverso l'uso di componenti del modello che contengono piccoli anelli di silossano. Le strutture dei piccoli anelli di silossano producono geometrie di legame Si-O-Si vincolate che immagazzinano energia di deformazione; minore è l'anello di silossano, maggiore è l'energia di deformazione immagazzinata per legame atomico. Misurando la cinetica dell'idrolisi per diverse dimensioni di struttura degli anelli di silossano, abbiamo mostrato che il tasso di idrolisi del silossano aumenta approssimativamente di 8 ordini di magnitudine per deformazioni di legame uguali a circa un sesto dell'energia del legame Si-O. In questo lavoro utilizzeremo i risultati dell'idrolisi dell'anello di silossano per sviluppare un modello basato sulla cinetica chimica quantitativa per prevedere il tasso di estensione della cricca nel vetro di silice in condizioni umide.

Translations of Source Text 3

Translator: CB

Le fenditure nel vetro aumentano in modo costante e a determinati livelli variando da meno di un trilionesimo di pollice (un pollice è pari a cm 2,54) all'ora fino a circa la metà della velocità del suono nel momento in cui il vetro si frantuma. Il livello dell'estensione della fenditura dipende dalla pressione esercitata e dall'ambiente chimico.
Le fenditure si sviluppano e aumentano nel momento in cui i legami tra il silicio e l'ossigeno vengono interrotti a causa della pressione. L'estremità della fenditura corrisponde a circa la dimensione dell'apertura quando la struttura dell'anello viene rotta, all'incirca 0,5 nanometri. La quantità di energia richiesta al fine di spezzare il legame silicio-ossigeno diminuisce di un fattore pari a venti in presenza di acqua. Questo comporta il fatto che una molecola d'acqua si inserisca nell'estremità della fenditura e trasformi il legame silicio-ossigeno in due gruppi di silanolo (SiOH). Anche altri elementi come ad esempio l'ammoniaca possono contribuire alla fenditura. Per quanto riguarda i legami silicio-ossigeno deformati, l'ammoniaca è in realtà più reattiva dell'acqua. Nel caso in cui la molecola nucleofilica si allarghi più di 0,5 nanometri, non si ottiene più tale effetto. L'acqua e l'ammoniaca hanno una dimensione molecolare di circa 0,3 nanometri.
Ci sono due fattori che influenzano la durabilità del vetro: la pressione e l'ambiente chimico. La pressione, nel caso del vetro lavorato, ha luogo quando il vetro si raffredda in modo disuguale. Il raffreddamento disuguale è inevitabile in quanto il vetro non

conduce bene il calore e la superficie si raffredda più rapidamente della parte interna. I soffiatori del vetro commerciale utilizzano i forni (forni di rinvenimento) nei quali per prima cosa riscaldano l'oggetto fino a raggiungere la temperatura di ricottura (leggermente al di sotto del punto di rammolimento) per poi raffreddare lentamente l'oggetto di vetro al fine di rimuovere ogni deformazione interna. Tale processo viene chiamato ricottura.

Translator: MS

Le fratture nel vetro crescono regolarmente ad una velocità ben determinata che può variare da un valore inferiore a un trillionesimo di pollice (1 pollice = 2,54 cm) all' ora a metà della velocità del suono quando il vetro va in mille pezzi. La velocità di propagazione della frattura dipende dal tipo di tensione esercitata su di essa e dall' ambiente chimico circostante.
Le fratture vengono prodotte e aumentano di dimensione quando una tensione spezza i legami silicio-ossigeno. Il punto d'origine della frattura è circa delle dimensioni dell' apertura che si viene a creare quando la struttura ad anello è spezzata (circa 0,5 nanometri). La quantità di energia necessaria per rompere il legame silicio-ossigeno subisce un calo di fattore venti in presenza di acqua. Questo indica che la molecola dell' acqua penetra nel punto d' origine della frattura e trasforma il legame silicio-ossigeno in due gruppi di silanolo (SiOH). Anche altri composti, come ad esempio l' ammoniaca, sono in grado di accelerare la velocità di frattura. L' ammoniaica è in effetti più reattiva rispetto all' acqua con i legami silicio-ossigeno già deformati. Tuttavia, se la molecola nucleofila diventa maggiore di 0,5 nanometri, l' effetto viene a mancare. Acqua ed ammoniaca hanno dimensioni molecolari di circa 0,3 nanometri.
Due fattori, la tensione esercitata e l' ambiente chimico circostante determinano la durata degli oggetti in vetro. La tensione sul vetro lavorato ha i suoi effetti quando il vetro viene raffreddato in modo non uniforme. Un raffreddamento di questo tipo è inevitabile in quanto il vetro non è un buon conduttore di calore e la sua superficie si raffredda più rapidamente rispetto al suo interno. I soffiatori di vetro [OMISSION] usano dei forni (detti forni di rinvenimento) nel quali dapprima riscaldano l' oggetto alla temperatura di ricottura (appena inferiore alla temperatura di rammollimento) per poi lentamente raffreddare l' oggetto in vetro ed eliminarne qualsiasi deformazione interna. Tale processo viene chiamato ricottura.

Translator: MT

Le fratture del vetro crescono in maniera continua a velocità controllate che vanno da meno di un quintilionesimo di pollice all'ora a circa la metà della velocità del suono quando il vetro si frantuma. la velocità di propagazione della frattura dipende dalla tensione applicata e dall'ambiente chimico. le fratture si sviluppano e crescono quando i legami silicio-ossigeno si rompono a causa della tensione. le dimensioni dell'apice della frattura sono simili a quelle dell'apertura che si crea quando la struttura dell'anello viene rotta, ovvero sono pari a circa 0,5 nanometri. la quantità di energia necessaria a scindere il legame silicio-ossigeno diminuisce di un fattore pari a 20 in presenza di acqua. ciò significa che la molecola di acqua si inserisce nell'apice della frattura e trasforma il

legame silicio-ossigeno in due gruppi di silanolo (SiOH). anche altri composti, come l'ammoniaca, possono agevolare la velocità di frattura: essa è in effetti più reattiva dell'acqua con i legami silicio-ossigeno deformati. se la molecola nucleofila raggiunge dimensioni superiori a 0,5 nanometri, tale fenomeno non si verifica più. l'acqua e l'ammoniaca possideono dimensioni molecolari pari a circa 0,3 nanometri. la durabilità del vetro è influenzata da due fattori, la tensione e l'ambiente chimico. la tensione nel vetro lavorato ha luogo quando il vetro stesso si raffredda in modo irregolare. il raffreddamento irregolare è inevitabile poichè il vetro non è un buon conduttore di calore: in questo modo la superficie si raffredda più rapidamente rispetto alla parte interna. i soffiatori [OMISSION] utilizzano forni di ricottura in cui prima riscaldano l'oggetto, fino a raggiungere la temperatura di ricottura (che si trova leggermente al di sotto della temperatura di rammollimento), e in seguito lo raffreddano lentamente per eliminare qualsiasi deformazione interna. questo processo viene denominato ricottura.

Translator: SC

Le incrinature nel vetro si sviluppano continuamente a velocità controllate che vanno da meno di un trilionesimo di millimetro all'ora fino a raggiungere quasi la metà della velocità del suono nel momento in cui il vetro si frantuma. La velocità di propagazione dell'incrinatura dipende dalla sollecitazione applicata e dall'ambiente chimico.
Le incrinature si sviluppano e aumentano quando i legami silicio-ossigeno si rompono in seguito a una sollecitazione. L'estremità dell'incrinatura è più o meno delle dimensioni dell'apertura creatasi in seguito alla rottura della struttura ad anelli, ovvero circa 0,5 nanometri. La quantità di energia necessaria per rompere il legame silicio-ossigeno diminuisce di venti volte in presenza d'acqua. Ciò significa che la molecola d'acqua riesce ad entrare nell'estremità dell'incrinatura e a convertire un legame silicio-ossigeno in due gruppi di silanolo (SiOH). Altri composti, come per esempio l'ammoniaca, possono favorire la propagazione delle incrinature: l'ammoniaca, infatti, è molto più reattiva rispetto all'acqua in presenza di legami silicio-ossigeno. Se la molecola nucleofilica assume dimensioni superiori ai 0,5 nanometri, quest'effetto non verrà più riscontrato. Le dimensioni molecolari dell'acqua e dell'ammoniaca si aggirano intorno ai 0,3 nanometri.
I fattori che possono influenzare la durevolezza del vetro sono due: le sollecitazioni e l'ambiente chimico. Nel caso di vetro lavorato, le sollecitazioni si riscontrano quando questo si raffredda in modo non uniforme; poichè il vetro non è un buon conduttore di calore, ciò è inevitabile, in quanto la superficie si raffredda più rapidamente rispetto all'interno. I soffiatori di vetro [OMISSION] utilizzano forni di rinvenimento nei quali l'oggetto viene prima riscaldato fino a raggiungere la temperatura di ricottura (poco al di sotto della temperatura di rammollimento) e poi raffreddato lentamente in modo da eliminare qualsiasi sollecitazione interna. Tale processo viene chiamato ricottura.

Translator: VM

Le fenditure nel vetro crescono costantemente a ritmi controllati, da meno di un trilionesimo di pollice all'ora a circa la metà della velocità del suono quando il vetro si

frantuma. La velocità di propagazione della frattura dipende dalla tensione applicata e dall'ambiente chimico.
Le fratture si sviluppano e crescono quando i legami silicio-ossigeno vengono rotti dalla tensione. La punta della fenditura ha approssimativamente le dimensioni dell'apertura creata quando la struttura ad anello viene rotta, circa 0,5 nanometri. La quantità di energia richiesta per spezzare il legame silicio-ossigeno diminuisce per un fattore venti in presenza dell'acqua. Questo indica che la molecola d'acqua si inserisce nella punta della fenditura e converte un legame silicio-ossigeno in due gruppi silanolici (SiOH). Altri composti, come l'ammoniaca, possono anche agevolare i ritmi di rottura. L'ammoniaca, in realtà, è più reattiva dell'acqua con legami silicio-ossigeno estesi. Se le dimensioni della molecola nucleofila superano 0,5 nanometri, l'effetto non sarà più lo stesso. Le dimensioni molecolari dell'acqua e dell'ammoniaca sono pari a circa 0,3 nanometri.
Due fattori, la tensione e l'ambiente chimico, influenzano la durabilità degli oggetti in vetro. La tensione, nel vetro lavorato, si ha quando il vetro si raffredda in modo disuguale. Un raffreddamento disuguale è inevitabile perché il vetro non è un buon conduttore di calore e la superficie si raffredda più rapidamente rispetto alla parte interna. I soffiatori di vetro destinato al commercio utilizzano forni di rinvenimento in cui prima riscaldano l'oggetto al punto di ricottura (poco al di sotto del punto di rammollimento) e in seguito raffreddano lentamente l'oggetto in vetro per rimuovere ogni eventuale deformazione interna. Questo processo viene chiamato ricottura.

Appendix 4: Editing process as reconstructed from keystroke logs

Note: the table shows the interim solutions proposed by translators as they drafted the target text. Interim solutions are indicated between angled brackets (<>) and in bold. The rightmost solution within each pair of brackets is the final one. Where the number of interim solutions is particularly high, these are presented as a list (the bottom solution is the final one).

Underlining indicates that a solution was proposed during revising, i.e. after completion of a first draft; dotted underlining indicates that a solution was arrived at a later stage than the first draft and before the final revision stage; highlighting in grey indicates multiple interim solutions typed as such by the translators themselves (i.e. alternative solutions not deleted immediately). Other annotations typed by translators are also highlighted in grey. The segmentation is mapped onto that carried out on the STs (see Appendix 2).

Translations of Source Text 1

Clause	Segment	CB	MS	MT	SC	VM
1	1 Il livello di **<rottura / fragilità>** del vetro	**<Il livello di <rottura / di velocità di frattura> / La velocità di frattura>**	il tasso di crescita della frattura	La rapidità con cui si propagano le **<crepe/incrinature / incrinature>**	**<La possibilità che il vetro si rompa / La rottura / La velocità con cui <un'incrinatura / una crepa> si ingrandisce>**
	2 non dipende	dipende	dipende	Dipende	non dipende
	3 solo da fattori chimici	non soltanto dall'ambiente chimico	non soltanto dalle condizioni chimiche,	**<non solo dall'am / dalle condizioni chimiche>**	solo dall'ambiente chimico,
	4 ma anche dalla misura in cui **<viene esercitata / si esercita>** la pressione su di esso	ma anche dalla **<potenza/grandezza / intensità>** dello **<stress/sforzo / sforzo (stress)>** applicato.	ma anche **<dalla grandezza della pressione / dalla pressione>** esercitata.	ma anche dall'intensità della **<forza applicata / tensione esercitata>**.	ma anche dall'intensità della pressione applicata.
2	5 Lo sviluppo di un modello completo **<per la cinetica del meccanismo / per l'energia cinetica>** riguardante **<la fragilità / la rottura / la frattura del vetro>**	Lo sviluppo di un modello completo della **<cinetica/meccanica / di una meccanica>** della frattura	lo sviluppo di un modello completo di cinetica della frattura	Per elaborare un modello **<esaustivo/completo / esaustivo>** della cinetica della **<frammentazione / frammentazione/frantumazione / frantumazione>**	Per sviluppare un modello completo **<che ci permetta di / per la cinetica>** di frattura <, / [no comma]>
	6 richiede	richiede	Richiede	**<richiede / è necessario>**	bisogna
	7 un'ampia conoscenza	la comprensione	la comprensione	comprendere	capire
3	8 di come	di come	del modo in cui	come	come

	9	la pressione	lo stress	la pressione	le tensioni	la pressione
	10	possa **<accellerare / accelerare>**	acceleri	Accelera	accelerano	**<accelleri / accellera / accelera>**
	11	la reazione di **<rottura / frattura / della frattura dei legami>**	la reazione legame-frattura	la reazione di rottura dei legami.	la reazione di rottura dei legami.	**<la reazione di rottura dei legami / tramite <cui / la quale> i legami si rompono>**.
4	12	Nel caso in cui non venga esercitata pressione	**<In assenza di stress, / se in assenza di stress [repositioned at end of clause]>**	in assenza di pressione,	In assenza di tensioni,	In assenza di pressione,
	13	la silice	la silice	**<il / la>** silice	il silice	la silice
	14	reagisce	**<reagisce / ha una reazione>**	reagisce	reagisce	reagisce
	15	molto lentamente	**<molto lentamente / molto lenta>**	**<in / molto>** lentamente con	molto lentamente	molto lentamente
	16	a contatto con l'acqua	**<con l'acqua / a contatto con l'acqua>**	con l'acqua.	con l'acqua.	con l'acqua.
5	17	La reazione dissociativa discussa in precedenza	La reazione di dissociazione appena spiegata	la reazione dissociativa **<di cui si è parlato / , di cui si è parlato,>**	La reazione dissociativa di cui abbiamo parlato	**<Le <reazioni dissociative / reazioni> di cui si è parlato / Tramite le reazioni di cui si è parlato>**
	18	porta	provoca	provoca	fa sì che	**<<causano / provocano> / [verb deleted]>**
	19	la superficie di vetro di silice	lo scioglimento **<nell'acqua del vetro di silice / della superficie del vetro di silice nell'acqua>**	**<la / lo>** scioglimento della superficie del vetro di quarzo	**<la superficie di un bicchiere in silice / il silice / la superficie di un vetro di silice [3rd version arrived at while translating 7.23]>**	**<<una / la> dissoluzione della superficie del vetro di silice / la superficie del vetro di silice>**
6	20	a dissolversi			si sciolga	si scioglie
	21	in acqua		in acqua	in acqua	in acqua
	22	**<ad un tasso / secondo un rapporto>** di 10 metri al secondo	ad una velocità di 10-17 metri al secondo;	ad **<un / una>** velocità di 10-17 metri al secondo;	alla velocità di 10-17 metri al secondo,	alla velocità di 10-17 metri al secondo;
7	23	I legami di superficie silicio-ossigeno, **<non sottoposti a pressione, / che non sono stati sottoposti a pressione,>**	**<i legami di superficie / i legami silicio- di superficie / i legami di superficie silicio-ossigeno> <non sottoposti / che non siano/vengano esposti / che non vengono esposti>** ad alcun tipo di stress	i legami **<di superficie / <ossigeno-silicio / silicio-ossigeno> di superficie> <privi /, non sottoposti a pressione,>**	i legami di superficie silicio-ossigeno **(?) [question mark eventually deleted]** liberi da tensioni	i legami di superficie silicio ossigeno **<su cui / sui quali>** non viene applicata alcuna pressione
	24	non presentano **<reazioni / nessuna reazione>**	risultano		**<sono / reagiscono>**	non reagiscono
	25		talmente **<privi di reazione / solidi>**	**<con n / non>** reagiscono	talmente poco	
	26	all'acqua	a contatto con l'acqua	con l'acqua	con l'acqua	con l'acqua
8	27	e non assorbono nemmeno	da non assorbire neppure	**<tanto da non assorbire neppure / e non assorbono neppure / e non assorbono>**	che non ne assorbono nemmeno	con l'acqua e **<non assorbo / non ne assorbono / e non assorbono /**
	28					

						e non ne assorbono> neppure
	29	il vapore acqueo	il vapore acqueo.	il vapore.	il vapore <. / ,>	<il vapore acqueo / il vapore>.
9	30	Tuttavia	Tuttavia	<se si esercita una pressione / esercitando una certa pressione,> tuttavia,	ma	Tuttavia <[no comma] / ,>
	31	il fatto di esercitare <la pressione / pressione>	l'applicazione <di uno sforzo/di un qualche tipo di stress / di un qualche tipo di stress>		il fatto <che / di> esercitare <della / [deleted]> tensione	<se la p / se viene applicat / qualora venga applicata pressione,>
	32	può portare a	può <aumentare / pro / far sì che>	è possibile che	può far sì che	le incrinature
	33	<rotture / fenditure>	<<le fratture / la velocità di frattura> superi un / delle fratture si espandano a velocità che <superano / superino / possono superare il>>	le fratture	le incrinature	si possono
10	34	che <si espandono / si formano>		aumentino	Aumentino	ingrandire
	35	a velocità <superiori / che superano / a una velocità che supera> <un / il> millimetro al secondo /	millimetro al secondo.	a velocità <mag / superiori> <di / a> un millimetro al secondo.	<ad una vel / a velocità> superiori ai tre millimetri al secondo.	a velocità maggiori di un millimetro al secondo.
11	36	<L / Un' / L'> estremità della fenditura	<Il / un> punto di frattura	l'estremità di una frattura	L'estremità dell'incrinatura	Un'estremità della crepa
	37	esercita	fa convergere	<concentra / focalizza>	<concentra / fa convergere>	concentra
	38	la pressione	lo stress	la pressione	la tensione	la pressione
12	39	in modo molto simile <a com / alla>	allo stesso modo in cui	così come	nello stesso modo in cui	nello stesso modo in cui
	40	punta di un ago	la punta di un ago	la punta di un ago	la punta di un ago	la punta di un ago
	41	che <indica / converge verso>	fa convergere	<concentra / focalizza>	<concentra / la concentra / fa convergere>	<riunisce / concentra>
	42	un campo elettrico.	un campo elettrico.	un campo elettrico.	in un campo elettrico <. / :>	un campo elettrico <. / :>
13	43	Più	Più	più	più	più
	44	ci si avvicina	<l'una si avvicina/ci si avvicina / questo si avvicina>	ci si avvicina	ci si avvicina	ci si avvicina
	45					
	46	all'estremità della fenditura,	al punto di rottura,	all'estremità della frattura,	all'estremità,	all'estremità della crepa,
14	47	più alta	Più	più grande	più aumenta	più
	48	è	<lo stress/il livello di stress / lo stress>	sarà		la pressione
	49	la pressione,	aumenta	la pressione,	la tensione,	aumenta,
15	50	fino <al momento in cui / a>	fino a	<fino a / che potrà>	fino a che	<finch / fino a>
	51	<raggiunge / si raggiungono / raggiungere>	raggiungere	raggiungere	non raggiunge	raggiungere
	52					
	53	milioni di libbre per ogni pollice quadrato	diversi milioni di chili per centimetro quadrati	alcuni milioni di libbre per pollice quadrato	<i milioni / milioni / qualche milione> di kilogrammi per metro quadro	<diverse / diversi> milioni di <libbre per pollice quadrato /

						bar> <. / ,>
	54	**<entro / in>** alcune dimensioni atomiche dell'estremità.	**<all / comprese>** nelle **<poche / limitate>** dimensioni atomiche **<della punta / del punto di rottura stesso.>**	**<all'in / entro le / all'interno delle > <picco / dimensioni>** atomiche della frattura **(?)**.	ad una piccolissima distanza dall'estremità dell'incrinatura **(?) [question mark eventually deleted]**.	in alcune dimensioni atomiche **<della/ dell'>** estremità.
16	**55**	Essendo esposta a tali pressioni	Esposto	esposta	Quando è esposta	Esposta
	56	a tali pressioni	**<ad / a>** stress così forti	**<a grandi /tali>** pressioni,	a tensioni di tale **<tipo / entità>**,	a queste pressioni intense,
17	**57**	la struttura atomica della silice	la struttura atomica della silice	la struttura atomica della silice	la struttura atomica del silice	la struttura atomica della silice
	58	risulta alterata	viene distorta	cambia	**<trasforma/modifica / modifica>**	devia
	59	rispetto alla sua normale configurazione di legame	dalla **<propria naturale / sua classica>** configurazione di legame.	la normale configurazione atomica dei legami.	la propria configurazione atomica normale.	**<dalla sua normale configurazione di legame / dalla normale configurazione della sua formazione di legame.>**
18	**60**	**<Calcoli teorici dimostrano / Secondo alcuni calcoli teorici>**	Calcoli teorici	Calcoli teorici	Calcoli teorici	Calcoli teorici
	61		hanno dimostrato che	dimostrano che	dimostrano che	**<mostrano / dimostrano>** che
19	**62**	se	**<[no equivalent] / se>**	se	Se	se
	63	un tetraedro silicato	il tetraedro del silicio	**<un tetraedro silicato / un tetraedro di silicato>**	**<un silicato / un tetraedro di silicato>**	un tetredro **<di silice / che f / contenuto in un silicato>**
	64	viene alterato	viene distorto da	viene deformato	**<è / si modifica>**	viene deformato
20	**65**	spingendo	da una tensione	spostando	perché gli atomi di ossigeno vengono spinti in un angolo,	estraendo
	66	gli atomi di ossigeno nell'angolo,	degli atomi di ossigeno agli angoli della struttura, **/ e /**	**<<a / gli atomi> di ossigeno / in un angolo gli atomi di ossigeno,>**		atomi di ossigeno **<dall' / da un>** angolo,
21	**67**	l'atomo di silicio **<che si trova al centro / al centro>**	l'atomo di silicio posto al centro	**<l'atomo di silicio al centro / è molto probabile che l'atomo di silicio al centro>** si leghi	**<l'atomo di silicio / è più probabile che l'atomo di silicio>** che si trova al centro	è più probabile che
	68	è	tende a legarsi			l'atomo di silicio al centro
	69	possibile che si leghi			si leghi	si leghi
	70	con una molecola d'acqua.	con una molecola d'acqua	con una molecola d'acqua.	con una molecola d'acqua.	con una molecola d'acqua.
22	**71**	Inoltre	Inoltre,	**<inoltre, l'interazione chimica con l'acqua / l'interazione chimica con l'acqua, inoltre,>**	Inoltre	Inoltre
	72	l'interazione chimica con l'acqua	l'interazione chimica con l'acqua		**<la reazione / l'interazione>** chimica con l'acqua	l'interazione chimica con l'acqua
	73	diminuirà	diminuirà ulteriormente	diminuirà	farà diminuire	**<diminuisce / provoca una diminuzione>**
	74	la forza necessaria per alterare ulteriormente i legami silicio-ossigeno.	**<la forza / la quantità di forza>** necessaria per alterare i legami silicio-ossigeno.	la forza richiesta per deformare ulteriormente i legami silicio-ossigeno.	la forza richiesta per **<distorcere / modificare>** ulteriormente i legami silicio ossigeno.	**<la / della>** forza necessaria per deformare ulteriormente i legami **<sili / tra il silicio e l'ossigeno>**

Translations of Source Text 2

Clause	Segment	CB	MS	MT	SC	VM
1	1	I modelli relativi a questo fenomeno	**<Modelli / Alcuni modelli>** fenomenologici	**<i modelli fenomenologici sono stati sviluppati / Sono stati sviluppati modelli fenomenologici>**	Sono stati sviluppati dei modelli fenomenologici **(?) [question mark eventually deleted]**	**<I / Sono>** stati sviluppati modelli fenomenologici
	2	sono stati sviluppati	sono stati sviluppati			
2	3	per **<rivolgersi / ricorrere>**	con lo scopo di studiare	per indirizzare	per spiegare	per **<rivolgere / indirizzare>**
	4	al meccanismo riguardante l'estensione della fenditura che dipende dal rapporto pressione/ambiente.	il meccanismo di dipendenza dell'estensione **<della / di una>** frattura - **rispetto <allo sforzo / alla tensione> esercitato/ambiente** - **rispetto allo sforzo esercitato e all'ambiente** - **rispetto allo sforzo esercitato su di essa e all'ambiente**	il meccanismo di **<estensione / propagazione>** della fenditura dipendente **<da tensione/ambiente / dal rapporto <tensione/ambiente / sforzo/ambiente / stress/ambiente / sforzo/ambiente>>**.	il meccanismo dell'estensione dell'incrinatura **(?) [question mark eventually deleted]** che dipende da condizioni **<di stress/ambiente / ambientali o di sollecitazione>**.	il meccanismo di estensione - **dipendente dalla** - **<della fenditura / di rottura / della cricca> dipendente dallo stress e dell'ambiente** - **della cricca <del / nel> legame tensione/ambiente** - **della cricca dipendente dalla tensione o dall'ambiente**
3	5	Tali modelli	Tali modelli	**<tali modelli / In base a tali modelli>**	Tali modelli	Questi modelli
	6	presuppongono che	**<presuppongono / assumono>** che		partono dal presupposto secondo cui	implicano che
4	7	la rottura **<dell'estremità della fenditura / del legame fenditura-estremità>**	la rottura del **<legame / legame chimico>** nel punto di frattura	la rottura del legame **<fenditura-estremità / cricca-estremità / cricca-apice>**	il fenomeno della rottura dei legami all'estremità dell'incrinatura	l'evento di rottura del **<legame / collegato alla / legame della>** punta della **<frattura / cricca>**
	8	sia facilitata	**<viene / venga>** facilitato	è facilitato	sia facilitato	**<è / sia>** facilitato
	9	da reazioni **<con l' / provocate dall'>** ambiente circostante	da reazioni chimiche **<dall' / provenienti dall'>** ambiente chimico circostante	dalle reazioni con specie **<provienti / provenienti>** dall'ambiente circostante	dalle reazioni con specie chimiche presenti nell'ambiente circostante	da reazioni con specie provenienti dall'ambiente circostante
5	10	in modo tale da	in modo da	**<in modo aumentare / che permettono di aumentare>**	in modo da	così da
	11	aumentare	Aumentare		aumentare	aumentare
	12	il livello macroscopico della fenditura.	la velocità di estensione della frattura macroscopica.	il tasso di estensione della **<frattura / cricca>** macroscopica.	**<il tasso di estensione / le dimensioni dell / l'estensione>** dell'incrinatura macroscopica.	il tasso dell'estensione **<di rottura macroscopico / macrosopica della cricca>**.
6	13	Questa considerazione sul processo di frattura	Questa interpretazione del processo di frattura	**<questa / Questa>** analisi del processo di frattura	Questo modo di intendere il processo di fratturazione	Questa visione del processo di frattura
	14	richiede che	**<richiede/necessita / vuole>** che	implica che	presuppone che	richiede
7	15	il livello della reazione chimica all'estremità della fenditura	la velocità di reazione chimica nel punto di frattura	il tasso di reazione chimica in corrispondenza dell' **<estremità / apice>** della **<fenditura / cricca>**	la quantità di reazioni chimiche all'estremità dell'incrinatura	un aumento del tasso di reazione chimica sulla punta della cricca **<. / [deletes full stop]>**

	16	sia direttamente proporzionale	**<sia / aumenti>** proporzionalmente	**<aumenti / cresce all'aumentare>**	**<aumenti a mano a mano che lo stress / diventi maggiore con l'aumentare> <dello stress / della sollecitazione>**	
8	17 18 19	alla pressione esercitata sul legame fenditura-estremità.	allo sforzo esercitato sul legame chimico presente nel punto di frattura stesso.	**<della tensione / dello sforzo>** sul legame **<fenditura-estremità / cricca- estremità / cricca-apice>**.	(?)**[question mark eventually deleted]** sui legami all'estremità dell'incrinatura.	all'aumentare della tensione sul legame della punta della cricca.
9	20	Nel caso specifico del vetro di **<silicie / silice>** fratturato nell'acqua,	**<Pe / Nel>** caso specifico del vetro di silice fratturato nell'acqua	**<nel / Nel>** caso specifico del vetro di silice fratturato in acqua,	Nel caso specifico del vetro di silice fratturato (?)**[question mark eventually deleted]** in acqua,	Nel caso specifico del vetro di silice fratturato in acqua,
	21	tali modelli fenomenologici	questi modelli fenomenologici	**<in base a/secondo / secondo>** tali modelli	questi modelli fenomenologici	questi modelli fenomenologici
	22	presuppongono che	**<teorizzano/prevedono / prevedono>** che		partono dal presupposto secondo cui	**<implicano / presuppongono / implicano>** che
10	23	il livello di idrolisi **<del/nei / nei>** legami di silossano	la velocità di **<frattura / idrolisi>** dei legami di silossano	il tasso di idrolisi dei legami di silossano	il tasso di idrolisi **<nei legami / dei legami di silossano>**	il tasso di idrolisi dei legami di silossano
	24	aumenterà	aumentino	Aumenterà	**<aumentino / aumenti>**	aumenti
	25	con la pressione esercitata.	con l'applicazione di uno sforzo.	se verrà **<esercitata tensione / applicato uno sforzo>**.	in seguito all'applicazione di **<uno stress / una sollecitazione>**.	**<con la / in proporzione alla / con la>** tensione applicata.
11	26	In un precedente studio	In un lavoro pubblicato precedentemente,	**<in / In>** uno studio precedente	In uno studio precedente	In uno studio precedente
	27	si è dimostrato che	avevamo dimostrato che	abbiamo dimostrato che	**<è / abbiamo>** dimostrato che	abbiamo mostrato che
	28					
12	29	la reattività chimica che aumenta la deformazione dei materiali di silicato	la reattività chimica intensificata dalla deformazione di materiali silicati	la reattività chimica **<de / aumentata>** dalla deformazione **<dei materiali di silicato/a base di silicato / dei materiali a base di silicato>**	la reattività chimica dei silicati aumentata in seguito a **<uno stress / una sollecitazione> (?) [question mark eventually deleted]**	la reattività chimica aumentata dalla deformazione dei **<materiali / silicati>**
	30	possa essere analizzata	può essere analizzata	può essere esaminata	può essere analizzata	può essere esaminata
	31	attraverso l'utilizzo di componenti che contengono piccoli anelli di silossano.	attraverso l'uso di composti modello contenenti anelli di silossano di piccole dimensioni.	mediante l'utilizzo di composti modello contenenti piccoli anelli di silossano.	attraverso l'impiego di composti base **<che contengono piccoli anelli di silossano / contenenti anelli di silossano di piccole dimensioni>**.	attraverso l'uso di **<dei / di>** componenti del modello che contengono piccoli anelli di silossano.
13	32	Le strutture dei piccoli anelli di silossano	**<Anelli di piccole dimensioni / Le strutture di silossano di piccole dimensioni / Le strutture ad anello di silossano di piccole dimensioni / le strutture ad anello di piccole dimensioni del silossano>**	**<le struttura/la struttura / La struttura>** dei piccoli anelli di silossano	**<Le strutture contenenti anelli di silossano / Queste strutture>**	Le strutture dei piccoli anelli di silossano
	33	producono	producono	**<danno / dà>** vita a	creano	producono
	34	geometrie ristrette **<di / del>** legame Si-O-Si che immagazzinano energia di	vincolate geometrie di legame Si-O-Si che conservano energia di deformazione;	geometrie di legame Si-O-Si forzate che accumulano l'energia **<della / di>** deformazione **<. / ;>**	legami del tipo Si-O-Si che immagazzinano l'energia prodotta da una sollecitazione:	geometrie di legame Si-O vincolate che immagazzinano energia di deformazione;

		deformazione;				
14	35	**<più / quanto più>** piccolo è	più **<piccola / piccolo>** è	più è piccolo	più piccole sono	**<più è piccolo / minore è>**
	36	l'anello di silossano	l'anello di silossano	l'anello di silossano,	le dimensioni dell'anello	l'anello di silossano,
15	37	tanto più grande è	maggiore è	maggiore sarà	maggiore è	maggiore è
	38	l'energia di deformazione immagazzinata per ogni legame atomico.	l'energia conservata per legame atomico.	l'energia **<della / di>** deformazione per legame atomico accumulata.	la quantità di energia immagazzinata in ogni legame atomico.	l'energia di deformazione immagazzinata per legame atomico.
16	39	Misurando	Misurando	**<misurando / Misurando>**	Dopo aver misurato	Misurando
	40	la cinetica dell'idrolisi	la cinetica dell'idrolisi	la cinetica dell'idrolisi	la cinetica dell'idrolisi	la cinetica dell'idrolisi,
	41	per diverse strutture di anelli di silossano,	**<delle / per le>** diverse dimensioni delle strutture ad anello del silossano,	**<per / relativa a>** strutture di anelli di silossano di diversa grandezza,	in strutture **<di silossano / contenenti anelli di silossano>** di diverse dimensioni,	per diverse dimensioni **<di strutture / di struttura>** degli anelli di silossano
17	42	si è dimostrato che	abbiamo dimostrato che	abbiamo dimostrato che	abbiamo dimostrato che	abbiamo mostrato che
	43					
18	44	il livello di idrolisi di silossano	la velocità di idrolisi del silossano	il tasso di idrolisi del silossano	il tasso di idrolisi del silossano	il tasso di idrolisi del silossano
	45	aumenta	aumenta	aumenta	aumenta	aumenta
	46	di circa 8 ordini di grandezza	di circa 8 ordini di grandezza	di circa 8 ordini di grandezza	di circa 8 **<ordini di magnitudo / ordini (?) / ordini di magnitudine>**	approssimativamente di 8 ordini di magnitudine
	47	per deformazioni di legame **<[no comma] / ,>** pari a circa un sesto dell'energia di legame Si-O.	per deformazioni di legame pari a circa un sesto dell'energia di legame Si-O.	per deformazioni di legame pari a circa un sesto dell'energia di legame Si-O.	in caso di sollecitazioni sui legami pari a circa un sesto dell'energia di legame presente in strutture Si-O.	per deformazioni di legame uguali a circa un sesto dell'energia del legame Si-O.
19	48	In questa relazione	In questo lavoro	**<nel / Nel>** presente studio	Nel corso di questo studio,	In questo lavoro
	49	si utilizzano	vengono utilizzati	**<ci riferiamo / ci riferiremo / faremo riferimento>**	utilizzeremo	utilizzeremo
	50					
	51	**<risultati / i risultati>** riguardanti l'analisi degli anelli di silossano	i risultati **<ottenuti <dall' / sull'> / relativi all'>** idrolisi dell'anello di silossano	ai risultati dell'idrolisi degli anelli di silossano	i risultati ottenuti dall'idrolisi degli anelli di silossano	i risultati dell'idrolisi dell'anello di silossano
20	52	per sviluppare	per sviluppare	per **<sviluppare/formulare / sviluppare>**	al fine di elaborare	per sviluppare
	53	un modello quantitativo basato sul rapporto chimico-cinetico	un modello quantitativo **<chimico-cinetico / ottenuto su basi chimico- cinetiche>**	un modello **<quantitativo basato sulla cinetica e sulla chi / quantitativo chimico-cinetico/chimico-cinetico quantitativo / chimico-cinetico quantitativo / cinetico-chimico quantitativo>**	un modello quantitativo basato sulla cinetica chimica	un modello basato sulla cinetica chimica quantitativa
21	54	al fine di prevedere	per prevedere	per predire	per prevedere	per prevedere
	55	il livello dell'estensione della fenditura nel vetro di silice in	la velocità di **<estensione/propagazione /**	il tasso di **<estensione della fenditura / propagazione della**	**<il tasso di / l'estensione dell'incrinatura>** del vetro di silice	il tasso di estensione della cricca **<in / nel>** vetro di silice in

condizioni di umidità.	propagazione> di frattura del vetro di silice in condizioni umide.	cricca> nel vetro di silice in condizioni di umidità.	in condizioni di umidità.	condizioni umide.

Translations of Source Text 3

Clause	Segment	CB	MS	MT	SC	VM
1	1	Le fenditure nel vetro	Le fratture nel vetro	Le fratture del vetro	Le incrinature nel vetro	Le fenditure nel vetro
	2	**<crescono / aumentano>**	crescono	crescono	si sviluppano	crescono
	3	in modo costante	**<in c / regolarmente>**	in maniera continua	continuamente	costantemente
	4	**<a livelli / a determinati livelli / e a determinati livelli>**	a velocità ad una velocità ben determinata	a velocità controllate	a velocità controllate	a ritmi controllati,
2	5	variando	che **<va / può variare>**	che **<vanno da/oscillano tra / vanno da>**	che vanno	da meno
	6	da meno di un trilionesimo di pollice (un pollice è pari a cm 2,5) all'ora	da **<meno di / un valore inferiore a> <un trilione / un trilionesimo>** di **<pollice / centimetro / pollice (1 pollice = 2,54 cm)>** all'ora	tra meno di un quintilionesimo di pollice all'ora	da meno di un trilionesimo di **<pollice (?) / millimetro>** all'ora	di un trillionesimo di pollice all'ora
	7	fino a circa la metà della velocità del suono	a metà della velocità del suono	**<a/e / a>** circa la metà della velocità **<della lu / del suono>**	fino a raggiungere quasi la metà della velocità del suono	**<ad approssimativamente la / a un valore approssimativo corrispondente a / a circa la>** metà della velocità della luce
3	8	quando	quando	quando	nel momento in cui	quando
	9	il vetro	il vetro	il vetro	il vetro	il vetro
	10	si frantuma.	va in mille pezzi.	si frantuma.	si frantuma.	si frantuma.
4	11	Il livello della pressione esercitata	La velocità della propagazione della frattura	la velocità di propagazione della frattura	La velocità di propagazione dell'incrinatura	La velocità di propagazione della frattura
	12	dipende	dipende	Dipende	Dipende	dipende
	13	dalla pressione esercitata	dal tipo di **<sforzo applicato / tensione esercitata>** su di essa	dalla tensione **<applicata/esercitata / applicata>**	dalla sollecitazione applicata	dalla tensione applicata e dall'ambiente chimico.
	14	e dall'ambiente chimico.	e dall'ambiente chimico circostante.	e dall'ambiente chimico.	e **<dalle condizioni ch / dall'ambiente chimico>**.	
5	15	Le fenditure	Le fratture	le fratture	Le incrinature	Le fratture
	16	si sviluppano e **<cresc / aumentano>**	vengono prodotte e aumentano di dimensione	si sviluppano e crescono	si sviluppano e **<crescono / aumentano>**	si sviluppano e crescono
6	17	nel momento in cui	quando	quando	quando	quando
	18	i legami tra silicio e ossigeno	**<uno sforzo/una tensione / una tensione>**	i legami silicio-ossigeno	i legami silicio-ossigeno	i legami silicio-ossigeno
	19	vengono interrotti	spezza	si rompono	si rompono	vengono rotti

	20	a causa della pressione.	**<i legami / il legame /** i legami**>** silicio-ossigeno	a causa della tensione.	in seguito a una sollecitazione.	dalla tensione.
7	21	L'estremità della fenditura	Il punto d'origine della frattura	**<l'apice/l'estremità / le dimensioni dell'apice/dell'estremità /** le dimensioni dell'apice**>** della frattura	L'estremità dell'incrinatura	La punta della fenditura
	22	corrisponde	è	sono	è	ha
	23	a circa	**<di circa / circa>**	**<quasi uguali/pari /** pari **/** simili**>** a	più o meno	approssimativamente
	24	la dimensione dell'apertura quando la struttura dell'anello viene **<distrutta/rotta /** rotta**>**,	delle dimensioni delle dimensioni dell'apertura che si viene a creare quando la struttura ad anello **<viene / è>** spezzata	quelle dell'apertura **<creata /** che si crea**>** quando **<l'anello / la struttura dell'anello>** viene rotta,	delle dimensioni dell'apertura creatasi in seguito alla rottura della struttura **<anulare/ad anelli(?) /** ad anelli**>**,	le dimensioni dell'apertura creata quando la struttura ad anello viene rotta,
	25	all'incirca 0,5 nanometri.	(circa 0,5 nanometri).	ovvero **<circa /** sono pari a circa**>** 0,5 nanometri.	ovvero circa 0.5 nanometri.	circa 0,5 nanometri.
8	26	La quantità di energia richiesta al fine di spezzare il legame silicio-ossigeno	La quantità di energia necessaria per rompere il legame silicio-ossigeno	la quantità di energia **<necessario /** necessaria**>** a scindere il legame silicio-ossigeno	La quantità di energia necessaria per rompere il legame silicio-ossigeno	La quantità di energia richiesta per spezzare il legame silicio-ossigeno
	27	diminuisce	**<cala di / subisce un calo <pari a /** di**>>**	diminuisce	diminuisce	diminuisce
	28	di un fattore pari a venti	**<un /** [deleted]**>** fattore venti	di un fattore pari a 20	di venti volte	per un fattore venti
	29	in presenza d'acqua.	in presenza di acqua.	in presenza d'acqua.	in presenza d'acqua.	in presenza dell'acqua.
9	30	Questo	Questo	ciò	Ciò	Questo
	31	comporta il fatto che	indica che	**<significa/indica / indica /** significa**>** che	significa che	indica che
10	32	una molecola d'acqua	la molecola dell'acqua	la molecola di acqua	le molecola d'acqua	la molecola d'acqua
	33	si inserisca	**<si / entra / penetra>**	si inserisce	riesce ad entrare	si inserisce
	34	nell'estremità della fenditura	nel punto d'origine della frattura	nell'**<apice/estremità /** apice**>** della frattura	nell'estremità dell'incrinatura	nella punta della fenditura
11	35	e trasformi	e **<converte /** trasforma**>**	e trasforma	e a convertire	e converte
	36	il legame silicio-ossigeno	il legame silicio-ossigeno	il legame silicio-ossigeno	un legame silicio-ossigeno	un legame silicio-ossigeno
	37	in due gruppi di silanolo (SiOH).	in due gruppi di silanolo (SiOH).	in due gruppi di silanolo (SiOH).	in due gruppi di silanolo (SiOH).	in due gruppi silanolici (SiOH).
12	38	Anche altri elementi	**<Altri / Anche altri>** composti,	**<altri / anche altri>** composti,	Altri composti,	Altri composti,
	39	come ad esempio l'ammoniaca	come ad esempio l'ammoniaca,	come l'ammoniaca,	come per esempio l'ammoniaca,	come l'ammoniaca,
	40	possono **<facilitare / contribuire>**	sono in grado di accelerare	possono **<facilitare/agevolare /** agevolare**>**	possono favorire	possono anche agevolare
	41	**<il livello della fenditura /** alla fenditura**>**.	la velocità di frattura.	**<la / le /** la**> <velocità/tasso/frequenza di frattura /** velocità**>** di farttura <. / : / ; / :>	la propagazione delle incrinature <. / :>	i ritmi di rottura.

13	42	Per quanto riguarda i legami silicio-ossigeno deformati, l'ammoniaca	L'ammoniaca	**<<in realtà / in effetti> l'ammoniaca**	l'ammoniaca	L'ammoniaca,
	43	è	è	**è / essa in effetti è>**	, infatti, **<è molto più reattiva / reagisce meglio / è molto più reattiva>**	in realtà, è
	44	in realtà	in effetti			
	45	più reattiva	più reattiva	più reattiva		più **<attiva / reattiva>**
	46	dell'acqua	**<dell' / rispetto all'>** acqua	dell'acqua	rispetto all'acqua	dell'acqua
	47	[see unit 41]	con **<i legami deformati di silicio-ossigeno / i legami silicio-ossigeno già deformati>**.	con i legami silicio-ossigeno deformati.	in presenza di legami silicio-ossigeno.	con i legami-silicio ossigeno **<deformati / estesi>**
14	47	Nel caso in cui	**<Se / Tuttavia, se>**	se	Se	Se
	49	la molecola nucleofilica	la molecola **<nucelofilica / nucleofila>**	la molecola nucleofila	la molecola nucleofilica	**<la molecola nucleofila <cresce / diventa maggiore di> / le dimensioni della molecola nucleofila superano>**
	50	si allarghi	diventa	raggiunge dimensioni superiori a	assume dimensioni superiori ai	
	51		**<più larga / maggiore>**			
	52	più di 0,5 nanometri,	di 0,5 nanometri,	0,5 nanometri,	0,5 nanometri,	0,5 nanometri,
15	53	non si ottiene più	l'effetto	**<l'effetto / tale effetto / tale fenomeno>**	quest'effetto	**<questo effetto non / l'effetto>**
	54		**<smette / viene a mancare>**.	**<non è più / non si verifica più>**.	non verrà più riscontrato.	non sarà più lo stesso.
	55	tale effetto				
16	56	L'acqua e l'ammoniaca	Acqua e ammoniaca	l'acqua e l'ammoniaca	**<l'acqua e l'ammoniaca / le dimensioni molecolari dell'acqua e dell'ammoniaca>** si aggirano intorno ai 0,3 nanometri.	**<L'acqua e l'ammoniaca hanno dimensioni molecolari di / Le dimensione molecolari dell'acqua e dell'ammoniaca sono pari a>** circa 0,3 nanometri.
	57	hanno	hanno	possiedono		
	58	una dimensione molecolare di circa 0,3 nanometri.	dimensioni molecolari di circa 0,3 nanometri.	dimensioni molecolari pari a circa 0,3 nanometri.		
17	59	Ci sono due fattori	Due fattori, **<lo sforzo applicato / la tensione esercitata>** e l'ambiente chimico circostante,	la durabilità del vetro	I fattori che possono influenzare la durevolezza del vetro	Due fattori, la tensione e l'ambiente chimico,
	60	che influenzano	**<influenzano / possono determinare / determinano>**	è influenzata	sono due:	influenzano
	61	la durabilità del vetro <. / ,> la pressione e l'ambiente chimico.	la durata degli oggetti di vetro.	d due fattori, la tensione e l'ambiente chimico.	**<la sollecitazione / le sollecitazioni>** e l'ambiente chimico.	la durabilità **<del / degli>** oggetti di vetro.
18	62	La pressione,	**<Lo sforzo applicato sul vetro lavorato / La tensione sul vetro lavorato>**	la tensione **<, / [no comma]>**	Nel caso di vetro lavorato,	La tensione,
	63	nel caso del vetro lavorato,		nel vetro lavorato **<, / [no comma]>**	**<la sollecitazione / le sollecitazioni>**	nel vetro lavorato,
	64	ha luogo	**<viene esercitata / si / ha i suoi effetti>**	ha luogo	**<si riscontra / si riscontrano>**	si ha
19	65	quando	quando	quando	quando	quando
	66	il vetro	il vetro	il vetro stesso	**<il vetro / questo>**	il vetro
	67	si raffredda	viene raffreddato	si raffredda	si raffredda	si raffredda

	68	in modo disuguale	in modo non uniforme.	in modo irregolare	in modo non uniforme <. / ;>	in modo disuguale.
20	69	Il raffreddamento disuguale	Un raffreddamento di questo tipo	il raffreddamento irregolare	poiché	Un raffreddamento disuguale
	70	è	è	è	il vetro	è
	71	inevitabile	inevitabile	inevitabile	non è buon conduttore calore,	inevitabile
21	72	in quanto	in quanto	**<perché / poiché>**	ciò	perché
	73	il vetro	il vetro	il vetro stesso	è inevitabile,	il vetro
	74	non conduce	non è	**<non conduce / non è un buon conduttore>**	in quanto	non è un buon conduttore di calore
	75	bene	un buon conduttore			
	76	il calore	di calore	di calore: **[colon added later]**		
22	77	e la superficie	e la sua **<superficie esterna / superficie>**	**<e / così che / e ciò fa sì hc/ in questo modo>** la superficie	la superficie	e la superficie
	79	si raffredda	si raffredda	si raffredda	si raffredda	si raffredda
	79	più rapidamente	più rapidamente	più rapidamente	più rapidamente	più rapidamente
	80	della parte interna.	rispetto **<alla / al>** suo interno.	rispetto alla parte interna.	rispetto all'interno.	rispetto alla parte interna.
23	81	I soffiatori **<di / del>** vetro commerciale	I soffiatori di vetro **<per / [deleted]>**	i soffiatori	I soffiatori di vetro	I soffiatori di vetro destinato al commercio
	82	utilizzano	usano	utilizzano	utilizzano	utilizzano
	83	i forni (forni di rinvenimento)	dei forni (detti forni di rinvenimento)	**<forni (i leher) / forni di ricottura>**	forni di rinvenimento	**<forni (/ forni di rinvenimento>**
24	84	nei quali	nei quali	in cui	nei quali	in cui
	85				l'oggetto viene prima riscaldato	
	86	per prima cosa	dapprima	prima		prima
	87	riscaldano	riscaldano	riscaldano		riscaldano
	88	l'oggetto	l'oggetto	l'oggetto, **[comma added later]**		l'oggetto
	89	fino a raggiungere la temperatura di ricottura	alla temperatura di ricottura	fino a raggiungere la temperatura di ricottura	fino a raggiungere la temperatura di ricottura	al punto di ricottura
	90	(leggermente	(appena	(**<leggermente / che si trova leggermente>**	(poco	(poco
	91	al di sotto del punto di rammollimento)	inferiore alla temperatura di rammollimento)	al di sotto della temperatura di **<ammorbidimento / rammollimento>**), **[comma added later]**	al di sotto della temperatura di rammollimento)	**<inferiore / al di sotto>** del punto di rammolimento)
25	92	per poi	**<e / per>** poi	e **<successivamente / in seguito>**	e poi	e **<quindi / poi / in seguito>**
	93	raffreddare	lentamente	**<raffreddano / lo raffreddano>** lentamente	raffreddato lentamente	raffreddano
	94	lentamente	raffreddare			lentamente
	95	l'oggetto di vetro	l'oggetto in vetro			l'oggetto in vetro

26	**96**	al fine di rimuovere	**<per eliminare / <u>ed eliminarne</u>>**	per eliminare	in modo da eliminare	per rimuovere
	97	ogni deformazione interna.	qualsiasi deformazione interna.	qualsiasi deformazione interna.	**<ogni / qualsiasi>** sollecitazione interna.	ogni eventuale deformazione interna.
27	**98**	Tale processo	**<Q / Tale>** processo	**<tale / questo>** processo	Tale processo	Questo processo
	99	viene chiamato	**<si / viene chiamato>**	viene denominato	viene chiamato	viene chiamato
	100	ricottura.	ricottura.	ricottura.	ricottura.	ricottura.

Appendix 5: Comparison tables used for Choice Network Analysis on the TTs

Task 1
Structural level

	CB	MS	MT	SC	VM
1	a	b	a	c	c
2	a	b	b	a	a
3	a	b	b	a	a
4	a	b	c	b	b
5	a	b	b	c	c
6	a	a	a	b	b
7	a	a	a	b	b
8	a	a	b	a	a
9	a	a	a	a	a
10	a	a	a	a	a
11	a	b	a	a	c
12	a	b	c	c	a
13	a	a	a	a	a
14	a	b	a	a	a
15	a	b	a	a	a
16	a	b	a	a	a
17	a	a	b	c	d
18	a	a	a	a	b
19	a	b	b	a	a
20	a	b	b	c	c
21	a	a	a	a	a
22	a	b	b	b	b
23	a	b	a	b	b
24 25	a	b	c	d	c
26	a	b	a	a	a
27 28	a	b	a	a	a
29	a	a	a	a	a
30	a	a	a	a	a
31	a	b	c	a	d
32	a	a	b	a	c
33	a	b	c	c	c
34	a	b	b	b	c
35	a	a	b	b	b
36	a	b	c	a	d
37	a	b	a	b	a
38	a	a	a	a	a
39	a	b	c	a	a
40	a	b	b	b	b
41	a	b	c	d	c
42	a	b	b	a	b
43	a	a	a	a	a

Task 1
Lexical level

	CB	MS	MT	SC	VM
1	a	b	c	d	e
2	a	a	a	a	a
3	a	b	c	c	b
4	a	b	a	c	b
5	a	b	c	d	e
6	a	a	a	b	c
7	a	b	b	b	c
8	a	a	b	a	a
9	a	b	a	a	a
10	a	a	a	a	a
11	a	b	b	b	b
12	a	b	a	c	a
13	a	a	a	a	a
14	a	a	a	a	a
15	a	a	a	a	a
16	a	a	a	a	a
17	a	b	c	a	a
18	a	b	b	c	d
19	a	a	a	a	a
20	a	b	b	b	b
21	a	a	a	a	a
22	b	a	a	a	a
23	a	b	a	c	d
24 25	a	b	c	a	c
26	a	b	a	a	a
27 28	a	a	a	a	a
29	a	a	b	a	b
30	a	a	a	b	a
31	a	b	a	c	d
32	a	b	c	a	c
33	a	b	b	c	c
34	a	b	c	c	d
35	a	a	a	a	b
36	a	b	c	d	e
37	a	b	c	b	d
38	a	b	a	c	a
39	a	a	b	a	a
40	a	a	a	a	a
41	a	a	b	a	c
42	a	b	b	b	a
43	a	a	a	a	a

44					
45	a	b	a	a	a
46	a	a	a	b	a
47	a	b	a	b	b
48	a	a	a	a	a
49	a	a	a	a	a
50	a	a	b	a	a
51					
52	a	a	b	c	c
53	a	a	a	a	b
54	a	b	c	d	a
55	a	b	b	c	b
56	a	b	a	c	d
57	a	a	a	a	a
58	a	a	b	b	b
59	a	b	c	c	d
60	a	a	a	a	a
61	a	b	b	b	b
62	a	a	a	a	a
63	a	b	c	c	d
64	a	a	a	b	a
65	a	b	a	c	a
66	a	b	c	d	e
67	a	a	a	a	a
68					
69	a	b	a	a	a
70	a	a	a	a	a
71	a	a	a	a	a
72	a	a	a	a	a
73	a	a	a	b	c
74	a	a	a	a	a

44					
45	a	a	a	a	a
46	a	b	c	a	d
47	a	b	c	b	b
48	a	b	a	c	a
49	a	b	a	b	b
50	a	a	b	a	a
51					
52	a	a	a	a	a
53	a	b	a	c	d
54	a	b	c	d	a
55	a	a	a	a	a
56	a	b	a	c	d
57	a	a	a	a	a
58	a	b	c	d	e
59	a	b	c	d	e
60	a	a	a	a	a
61	a	b	b	b	b
62	a	a	a	a	a
63	a	b	a	a	a
64	a	b	c	d	c
65	a	b	c	a	d
66	a	b	a	a	a
67	a	a	a	a	a
68					
69	a	b	c	c	c
70	a	a	a	a	a
71	a	a	a	a	a
72	a	a	a	a	a
73	a	a	a	a	a
74	a	a	b	c	c

Task 2
Structural level

	CB	MS	MT	SC	VM
1	a	a	a	b	a
2	a	a	a	a	a
3	a	b	a	a	a
4	a	b	a	c	c
5	a	a	a	a	a
6	a	a	b	c	a
7	a	b	a	c	d
8	a	a	a	a	a
9	a	b	b	b	b
10	a	a	b	a	c
11	a	a	a	a	a
12	a	b	b	b	c
13	a	a	a	a	a
14	a	a	a	a	a
15	a	a	b	a	a
16	a	b	c	a	d
17 18	a	b	c	d	c
19	a	a	b	b	b
20	a	a	a	a	a
21	a	a	a	a	a
22	a	a	b	c	a
23	a	a	a	a	a
24	a	a	a	a	a
25	a	b	c	d	a
26	a	b	a	a	a
27 28	a	a	a	a	a
29	a	b	b	c	c
30	a	a	a	a	a
31	a	a	a	a	b
32	a	b	c	d	a
33	a	a	a	a	a
34	a	a	a	a	a
35	a	b	b	b	b
36	a	a	a	a	a
37	a	a	a	a	a
38	a	a	a	a	a
39	a	a	a	b	a
40	a	a	a	a	a
41	a	b	c	d	e
42 43	a	a	a	a	a
44	a	a	a	a	a
45	a	a	a	a	a
46	a	a	a	a	a
47	a	a	a	b	a
48	a	a	a	b	a

Task 2
Lexical level

	CB	MS	MT	SC	VM
1	a	a	a	a	a
2	a	a	a	a	a
3	a	b	c	d	e
4	a	b	c	d	e
5	a	a	a	a	a
6	a	b	c	d	e
7	a	b	c	d	e
8	a	a	a	a	a
9	a	a	a	a	a
10	a	a	b	a	c
11	a	a	a	a	a
12	a	b	c	d	c
13	a	b	c	d	e
14	a	b	c	d	a
15	a	b	c	d	c
16	a	b	c	d	b
17 18	a	b	c	d	e
19	a	a	b	b	b
20	a	a	a	a	a
21	a	a	a	a	a
22	a	b	c	d	e
23	a	b	c	c	c
24	a	a	a	a	a
25	a	b	b	c	d
26	a	b	a	a	a
27 28	a	a	a	a	a
29	a	a	a	a	a
30	a	a	a	b	b
31	a	b	b	c	a
32	a	a	a	b	a
33	a	a	b	c	a
34	a	b	c	d	b
35	a	a	a	a	b
36	a	a	a	a	a
37	a	b	b	b	b
38	a	b	c	a	a
39	a	a	a	a	a
40	a	a	a	a	a
41	a	b	c	b	b
42 43	a	a	a	a	a
44	a	b	c	c	c
45	a	a	a	a	a
46	a	a	a	b	c
47	a	a	a	b	a
48	a	b	c	c	b

49					
50	a	a	a	a	a
51	a	a	b	c	b
52	a	a	a	a	a
53	a	a	b	a	c
54	a	a	a	a	a
55	a	a	a	a	a

49					
50	a	a	b	a	a
51	a	a	b	c	a
52	a	a	a	b	a
53	a	a	a	a	a
54	a	b	b	b	b
55	a	b	c	d	e

Task 3
Structural level

	CB	MS	MT	SC	VM
1	a	a	a	a	a
2	a	a	a	a	a
3	a	b	a	b	b
4	a	a	a	a	a
5	a	b	b	b	c
6	a	b	a	a	a
7	a	b	b	c	b
8	a	b	b	a	b
9	a	a	a	a	a
10	a	b	a	a	a
11	a	a	a	a	a
12	a	a	a	a	a
13	a	b	a	a	a
14	a	a	a	a	a
15	a	a	a	a	a
16	a	b	b	b	b
17	a	a	a	a	a
18	a	a	a	a	a
19	a	b	a	a	c
20	a	b	c	a	a
21	a	a	b	a	a
22	a	a	b	a	a
23	a	b	c	d	e
24	a	a	b	b	a
25	a	a	a	a	a
26	a	a	b	a	a
27	a	b	a	b	c
28	a	a	a	a	a
29	a	a	a	a	a
30	a	b	b	b	b
31	a	a	a	a	a
32	a	a	a	b	a
33	a	a	a	a	a
34	a	a	a	a	a
35	a	a	a	a	a
36	a	a	a	a	b
37	a	a	a	b	b
38	a	a	b	a	b
39	a	b	a	a	a
40	a	b	b	b	b
41	a	a	b	a	a
42	a	a	a	a	a
43	a	a	a	a	a
44	a	a	a	b	a
45	a	b	a	b	a
46	a	b	a	c	a
47	a	b	c	c	c
48	a	a	a	a	b

Task 3
Lexical level

	CB	MS	MT	SC	VM
1	a	b	b	c	a
2	a	b	b	c	b
3	a	b	c	c	c
4	a	b	c	c	d
5	a	a	a	a	b
6	a	a	a	b	a
7	a	a	a	a	a
8	a	b	b	a	b
9	a	a	a	a	a
10	a	b	a	a	a
11	a	b	b	c	b
12	a	a	a	a	a
13	a	b	b	c	b
14	a	b	b	c	b
15	a	b	c	a	c
16	a	b	b	b	b
17	a	a	a	a	a
18	a	b	c	c	d
19	a	b	b	c	b
20	a	b	c	a	d
21	a	b	b	b	c
22	a	a	b	c	d
23	a	a	a	a	a
24	a	a	a	a	a
25	a	a	b	c	c
26	a	a	b	a	a
27	a	a	a	a	b
28	a	a	a	a	a
29	a	a	b	b	a
30	a	b	c	c	a
31	a	a	a	a	a
32	a	b	a	c	a
33	a	b	c	a	d
34	a	a	a	b	b
35	a	a	a	a	a
36	a	a	a	a	a
37	a	a	a	a	a
38	a	a	a	a	a
39	a	b	c	d	c
40	a	b	b	c	d
41	a	a	b	a	a
42	a	a	a	a	a
43	a	b	b	c	a
44	a	a	a	a	a
45	a	a	a	a	a
46	a	a	a	b	c
47	a	b	b	b	b
48	a	a	a	a	a

49	a	a	b	b	a
50	a	b	b	b	c
51	a	b	b	b	c
52	a	a	a	a	a
53	a	b	a	a	c
54	a	b	a	a	a
55	a	a	a	b	b
56	a	a	a	a	a
57	a	a	a	b	b
58	a	a	b	b	a
59	a	b	c	d	b
60	a	b	c	a	b
61	a	b	a	a	b
62	a	a	a	a	a
63	a	b	b	a	b
64	a	b	a	c	c
65	a	a	a	a	a
66	a	a	b	c	a
67	a	a	a	a	a
68	a	a	a	a	a
69	a	b	a	c	a
70	a	a	a	a	a
71	a	a	a	a	a
72	a	a	a	a	a
73	a	a	a	a	a
74	a	b	b	b	b
75	a	b	b	b	b
76	a	b	b	b	b
77	a	a	b	c	a
78	a	a	a	a	a
79	a	a	a	a	a
80	a	b	b	b	b
81	a	a	b	a	b
82	a	a	a	a	a
83	a	b	c	c	c
84	a	a	a	a	a
85	a	a	a	a	a
86	a	b	b	b	b
87	a	a	a	a	a
88	a	a	a	a	a
89	a	b	a	a	b
90	a	a	b	a	a
91	a	b	a	a	a
92	a	a	b	c	b
93	a	a	b	a	a
94	a	b	c	c	b
95	a	a	a	a	a
96	a	b	a	a	a
97	a	a	a	a	a
98	a	a	a	a	a
99	a	a	a	a	a
100	a	a	a	a	a

49	a	b	c	d	e
50	a	b	c	c	d
51	a	a	a	a	a
52	a	a	a	a	a
53	a	b	c	d	e
54	a	b	a	a	a
55	a	a	a	a	a
56	a	a	b	c	d
57	a	a	a	a	a
58	a	a	a	a	a
59	a	b	b	c	b
60	a	b	a	a	a
61	a	b	a	c	a
62	a	b	b	c	a
63	a	a	a	a	a
64	a	b	a	c	d
65	a	a	a	a	a
66	a	a	a	b	a
67	a	a	a	a	a
68	a	b	c	b	a
69	a	b	c	d	a
70	a	a	a	a	a
71	a	a	a	a	a
72	a	a	b	b	c
73	a	a	a	a	a
74	a	a	a	a	a
75	a	a	a	a	a
76	a	a	a	a	a
77	a	a	a	a	a
78	a	a	a	a	a
79	a	a	a	a	a
80	a	a	a	a	a
81	a	b	c	b	b
82	a	b	a	a	a
83	a	a	b	a	a
84	a	a	b	a	b
85	a	a	a	a	a
86	a	b	c	c	c
87	a	a	a	a	a
88	a	a	a	a	a
89	a	a	a	a	b
90	a	b	a	c	c
91	a	b	b	b	a
92	a	a	b	a	b
93	a	a	a	a	a
94	a	b	c	d	b
95	a	a	a	a	a
96	a	b	b	b	a
97	a	a	a	b	c
98	a	a	a	a	a
99	a	a	a	a	a
100	a	a	a	a	a

References

Ahmad, K. (1999) Specialist Texts and their Quirks. *Tama '98: Proceedings of the 4th TermNet Symposium*. Vienna: TermNet, 141-157.

Ahmad, K. and Rogers, M. (2001) Corpus Linguistics and Terminology Extraction in Wright, S. E. and Budin, G. (eds) *Handbook of Terminology Management*. Vol. 2. Amsterdam/Philadelphia: John Benjamins, 725-760.

Alves, F. (ed.) (2003) *Triangulating Translation. Perspectives in process oriented research*. Amsterdam/Philadelphia: John Benjamins.

Arntz, R. (1988) Steps towards a Translation-oriented Typology of Technical Texts. *Meta* 33, 4, 468-471.

Banks, R. (nd) Fracturing Glass. In *A Short History of Glassblowing and Glassblowing Techniques*. Available: http://chemistry.boisestate.edu/rbanks/glassblowing/glassblowing_history.htm [Accessed 20 April 2006].

Bazerman, C. (1988) *Shaping Written Knowledge. The Genre and Activity of the Experimental Article in Science*. Academic Writing Landmark Publications in Writing Studies. Available: http://aw.colostate.edu/books/bazerman_shaping/ [Accessed 20 April 2006] (originally published in print by University of Wisconsin Press, Madison, Wisconsin).

Beccaria, G. (ed.) (2004) *Dizionario di linguistica*. Nuova edizione. Torino: Einaudi.

Bernardini, S. (1999) Using Think-Aloud protocols to investigate the translation process: Methodological aspects in Williams, J. N. (ed.) *RCEAL Working papers in English and applied linguistics 6*. Cambridge: University of Cambridge, 179-199.

_____. (2001) Think-Aloud Protocols in Translation Research: Achievements, Limits, Future Prospects. *Target* 13, 2, 241-263.

Bernth, A. and Gdaniec, C. (2001) MTranslatability. *Machine Translation* 16, 175-281.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Bonani, I. Masini, A., Morgana, S. e Piotti, M. (2003) *Elementi di linguistica italiana*. Roma, Carocci.

Bowker, L. and Pearson, J. (2002) *Working with Specialized Language. A practical guide to using corpora*. London/New York: Routledge.

Breedveld, H. (2002) Translation processes in time. *Target* 14, 2, 221-240.

Cabré, M.T. and Estopà, R. (2003) On the units of specialised meaning used in professional communication. *Terminology Science and Research* 14, 15-27.

Campbell, S. (1998) *Translation into the Second Language*. Harlow: Longman.

_____. (1999) A Cognitive Approach to Source Text Difficulty in Translation. *Target* 11, 1, 33-63.

_____. (2000) Choice Network Analysis in Translation Research in Olohan, M. (ed.) *Intercultural Faultlines. Research models in Translation Studies 1. Textual and cognitive aspects*. Manchester: St. Jerome, 29-42.

Campbell, S. and Hale, S. (1999) What makes a text difficult to translate? *Refereed Proceedings of the 23rd Annual ALAA Congress*. Available: http://www.cltr.uq.edu.au/alaa/proceed/camphale.html [Accessed: 15 March 2006].

Campbell, S. and Wakim, B. (2007) Methodological questions about translation research. A model to underpin research into the mental processes of translation. *Target* 19, 1, 1-19.

Chesterman, A. (2000) Teaching Strategies for Emancipatory Translation in Schäffner, C. and Adab B. (eds) *Developing Translation Competence*. Amsterdam/Philadelphia: John Benjamins, 77-90.

_____. (2001) Empirical research methods in Translation Studies. *Erikoiskielet ja käännösteoria* (VAKKI-symposiumi XX) 27, 9-22. Available: http://www.helsinki.fi/~chesterm/2000lVakki.html [Accessed: 29 November 2006].

Crystal, D. and Davy, D. (1969) *Investigating English Style*. London: Longman.
Dancette, J. (1992) Modèles sémantique et propositionnel de l'analyse de la fidélité en traduction. *Meta* 37, 3, 440-449.

Dancette, J. (1997) Mapping Meaning and Comprehension in Translation. Theoretical and Experimental Issues in Danks, J. H, Shreve, G. M., Fountain, S. B. and McBeath, M. K. (eds) *Cognitive Processes in Translation and Interpreting*. Thousand Oaks, Cal.: Sage Publications, 77-103.

Dardano, M. (1996) *Manualetto di linguistica italiana*. Bologna: Zanichelli.

Davison, A. and Green, G.M. (eds) (1988) *Linguistic Complexity and Text Comprehension. Readability Issues Reconsidered*. Hillsdale, NJ: Lawrence Erlbaum Associates.

de Groot, A. (1997) "The Cognitive Study of Translation and Interpretation: Three Approaches". In: Danks, J. H., Shreve, G. M., Fountain, S. B. and McBeath, M. K. (eds) *Cognitive Processing in Translation and Interpretation*. Thousand Oaks, Cal.: Sage Publications, 25-56.

Dörner, D. (1976) *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.

Douglas, D. (1999) *Assessing Language for Specific Purposes*. Cambridge: Cambridge University Press.

Dragsted, B. (2005) Segmentation in translation: Differences across levels of expertise and difficulty. *Target* 17, 1, 49-70.

Dudley-Evans, T. and St John, M. (1998) *Developments in English for Special Purposes*. Cambridge: Cambridge University Press.

Eggins, S. (1994) *An Introduction to Systemic-Functional Grammar*. London: Pinter.

Eggins, S. and Martin J.R. (1997) Genres and Registers of Discourse in van Dijk, T. (ed.) *Discourse as Structure and Process* (Discourse Studies: A Multidisciplinary Introduction, vol. 1). London: Sage, 230-256.

Englund Dimitrova, B. (2005) *Expertise and Explicitation in the Translation Process*. Amsterdam/Philadelphia: John Benjamins.

Ericsson, K.A. and Simon, H.A. (1984, 1993[2]) *Protocol Analysis: Verbal Reports as Data*. Cambridge, Mass.: MIT Press.

Fraser, J. (1996) The Translator Investigated: Learning from Translation Process Analysis. *The Translator* 2, 1, 65-79.

Gile, D. (1992) Les fautes de traduction: une analyse pédagogique. *Meta* 37, 2, 251-262.

_____. (1998) Observational Studies and Experimental Studies in the Investigation of Conference Interpreting. *Target* 10, 1, 69-98.

Göpferich, S. (1998) Levels of comprehending LSP texts in Hansen, G. (ed.) *LSP texts and the process of translation*. Copenhagen Working Papers in LSP 1. 1-29.

Gopnik, M. (1972) *Linguistic Structures in Scientific English*. The Hague: Mouton.

Gouadec, D. (1981) Paramètres de l'évaluation des traductions. *Meta* 26, 2, 99-116.

_____. (1989) Comprendre, évaluer, prévenir. Pratique, enseignement et recherche face à l'erreur et à la faute en traduction. *TTR* 2, 2, 35-54.

Gross, A. G. (1990) *The Rhetoric of Science*. Cambridge, Mass.: Harvard University Press.

Hale, S. and Campbell, S. (2002) The Interaction Between Text Difficulty and Translation Accuracy. *Babel* 48, 1, 14-33.

Halliday, M. A. K. (1978) *Language as Social Semiotic*. London: Edward Arnold.

_____. (1985) *An Introduction to Functional Grammar*. London: Edward Arnold.

_____. (1993a) On the Language of Physical Science in Halliday. M. A. K. and Martin, J. R. *Writing Science: Literacy and Discursive Power*. London/Washington D.C.: Falmer, 54-68.

_____. (1993b) Some Grammatical Problems in Scientific English in Halliday. M. A. K. and Martin, J. R. *Writing Science: Literacy and Discursive Power*. London/Washington D.C.: Falmer, 69-85.

_____. (1994) *An Introduction to Functional Grammar*. 2nd edition. London: Edward Arnold.

_____. (1996) On grammar and grammatics in Hasan, R., Cloran, C. and Butt, D. (eds) *Functional Descriptions. Theory in practice*. Amsterdam/Philadelphia: John Benjamins.

_____. (1997) On the Grammar of Scientific English in Taylor Torsello, C. (ed.) *La frase*. Trieste: Scuola Superiore di Lingue Moderne per Interpreti e Traduttori, 21-56.

Halliday, M. A. K. and Martin, J. R. (1993) *Writing Science: Literacy and Discursive Power*. London/Washington D.C.: Falmer.

Halskov Jensen, E. (2002) Problemas de traducción según géneros – una investigación empírica in Hansen, G. (ed.) *Empirical Translation Studies*. Copenhagen Studies in Language 27. Copenhagen: Samfundslitteratur, 125-145.

Hansen, G. (ed.) (1999) *Probing the Process of Translation. Methods and Results*. Copenhagen Studies in Language 24. Copenhagen: Samfundslitteratur.

_____. (ed.) (2002) *Empirical Translation Studies*. Copenhagen Studies in Language 27. Copenhagen: Samfundslitteratur.

_____. (2003) Controlling the Process. Theoretical and methodological reflections on research in Translation processes in Alves, F. (ed.) *Triangulating Translation. Perspectives in process oriented research*. Amsterdam/Philadelphia: John Benjamins, 25-42.

Harris, P. (2002) *Designing and Reporting Experiments in Psychology*. 2nd edition. Buckingham/Philadelphia: Open University Press.

Hatim, B. (2001) *Teaching and Researching Translation*. Harlow: Longman.

Hatim, B. and Mason, I. (1997) *The Translator as Communicator*. London/New York: Routledge.

Hayes, D.P. (1992) The growing inaccessibility of science. *Nature* 356. 739-740.

Hodge, R. and Kress G. (1988) *Social Semiotics*. Cambridge: Polity Press.

Hogben, L. (1969) *The Vocabulary of Science*. London: Heinemann.

Holmes, J. S. (1972) *The Name and Nature of Translation Studies*. Amsterdam: Translation Studies Section, University of Amsterdam, Department of General Literary Studies [Reprinted in Holmes 1988].

_____. (1988) *Translated! Papers on Literary Translation and Translation Studies*. Amsterdam: Rodopi.

_____. (1978) The future of translation theory: A handful of theses. Paper presented at the International Symposium on Achievements in the Theory of Translation, Moscow, 23-30 October 1978 [Printed in Holmes 1988].

House, J. (1977, 1981[2]) *A Model for Translation Quality Assessment*. Tübingen: Narr.

House, J. (1997) *Translation Quality Assessment. A Model Revisited*. Tübingen: Narr.

Huddleston, R.D. (1971) *The Sentence in Written English*. Cambridge: Cambridge University Press.

Hurtado Albir, A. (2001) *Traducción y traductología. Introducción a la traductología*. Madrid: Cátedra.

Jääskeläinen, R. (1990) *Features of Successful Translation Processes. A Think-Aloud Protocol Study*. Licentiate thesis, University of Joensuu.

_____. (1993) Investigating translation strategies in Tirkkonen-Condit, S. (ed.) *Recent trends in empirical translation research*. Joensuu: University of Joensuu Faculty of Arts, 99-120.

_____. (1998) Think-Aloud Protocols in Baker, M. (ed.) *Routledge Encyclopedia of Translation Studies*. London/New York: Routledge, 265-269.

_____. (1999) *Tapping the Process. An Explorative Study of the Cognitive and Affective Factors Involved in Translating*. Joensuu: Joensuun Yliopisto.

_____. (2000) Focus on Methodology in Think-Aloud Studies on Translating in Tirkkonen-Condit, S. and Jääskeläinen, R. (eds) *Tapping and Mapping the Processes of Translating and Interpreting. Outlooks on Empirical Research*. Amsterdam/Philadelphia: John Benjamins, 71-82.

Jääskeläinen, R. and Tirkkonen-Condit, S. (1991) Automatised processes in professional vs. non-professional translation: A think-aloud protocol study in Tirkkonen-Condit, S. (ed.) *Empirical research in translation and intercultural studies*. Tübingen: Narr, 89-109.

Jakobsen, A. L. (1999) Logging target text production with *Translog* in Hansen, G. (ed.) *Probing the Process in Translation. Methods and Results*. Copenhagen Studies in Language 24. Copenhagen: Samfundslitteratur, 9-20.

_____. (2003) Effects of Think Aloud on Translation Speed, Revision and Segmentation in Alves, F. (ed.) *Triangulating Translation. Perspectives in process oriented research*. Amsterdam/Philadelphia: John Benjamins, 69-95.

Jakobsen, A. L. and Schou, L. (1999) Translog documentation in Hansen, G. (ed.) *Probing the Process in Translation. Methods and Results*. Copenhagen Studies in Language 24. Copenhagen, Samfundslitteratur, 151-186.

Killingsworth, M. J. and Gilbertson, M. K. (1992) *Signs, Genres and Communities in Technical Communication*. Amityville: Baywood.

Kittredge, R. and Lehrberger, J. (eds) (1982) *Sublanguage. Studies of Language in Restricted Semantic Domains*. Berlin: De Gruyter.

Klare, G. R. (1963) *The Measurement of Readability*. Ames, Iowa: Iowa State University Press.

Krings, H. P. (1986) *Was in den Köpfen von Übersetzern vorgeht. Eine empirische Untersuchung der Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern*. Tübingen: Narr.

_____. (2001) *Repairing Texts. Empirical Investigations of Machine Translation Post-Editing Process*. Edited G.S. Koby, translated by G. S. Koby, G. M. Shreve, K. Mischerikow and S. Litzer. Kent, Ohio: Kent State University Press.

_____. (2005) Wege ins Labyrinth – Fragestellungen und Methoden der Übersetzungsprozessforschung im Überblick. *Meta* 50, 2, 342-358.

Kupsch-Losereit, S. (1985) The problem of translation error evaluation in Titford, C. and Hieke, A. E. (eds) *Translation in Foreign Language Teaching and Testing*. Tübingen: Narr, 169-179.

Kussmaul, P. (1995a) *Training the Translator*. Amsterdam/Philadelphia: John Benjamins.

_____. (1995b) Comprehension Processes and Translation. A Think-Aloud Protocol (TAP) Study in Snell.Hornby, M., Jettmarovà, Z. and Kaindl, K. (eds) *Translation as Intercultural Communication. Selected papers from the EST Congress – Prague 1995*. Amsterdam/Philadelphia: John Benjamins, 239-248.

_____. (1998) Die Erforschung von Übersetzungsprozessen: Resultate und Desiderate. *Lebende Sprachen* 2, 49-53.

_____. (2005) Translation through Visualization. *Meta* 50, 2. 378-391.

Lassen, I. (2003) *Accessibility and Acceptability in Technical Manuals. A Survey of Style and Grammatical Metaphor*. Amsterdam/Philadelphia: John Benjamins.

Laviosa-Braithwaite, S. (1996) Comparable corpora: Towards a corpus linguistic methodology for the empirical study of translation in Thelen, M. and Lewandowska-Tomaszczyk, B. (eds) *Translation and Meaning Part 3*. Maastricht: Euroterm. 153-163.

Livbjerg, I. and Mees, I. M. (2002) Problem solving at different points in the translation process: quantitative and qualitative data in Hansen, G. (ed.) *Empirical Translation Studies*. Copenhagen Studies in Language 27. Copenhagen: Samfundslitteratur, 147-190.

_____. (2003) Patterns of Dictionary Use in Non-Domain-Specific Translation in Alves. F. (ed.) *Triangulating Translation. Perspectives in process oriented research.* Amsterdam/Philadelphia: John Benjamins, 123-136.

Lorenzo, M. P. (1999) Apuntes para una discusión sobre métodos de estudio del proceso de traducción in Hansen, G. (ed.) *Probing the Process in Translation. Methods and Results*. Copenhagen Studies in Language 24. Copenhagen: Samfundslitteratur, 21-42.

Lörscher, W. (1991) *Translation Performance, Translation Process and Translation Strategies. A Psycholinguistic Investigation*. Tübingen: Narr.

Markus, V. U. (2000) Verstehen und Kreativität beim Übersetzen. Englische und deutsche Muttersprachler im Vergleich. *Lebende Sprachen* 1, 1- 8.

Martínez Melis, N. and Hurtado Albir, A. (2001) Assessment in Translation Studies: Research Needs. *Meta* 46, 2, 272-287.

Michalske T.A and Bunker B.C. (1987) The fracturing of glass. *Scientific American*, December, 78-85.

_____. (1988) La fratturazione del vetro. *Le scienze*, febbraio, 70-79.

_____. (1993) A Chemical Kinetics Model for Glass Fracture. *Journal of the American Ceramic Society* 76, 10. 2613-2618.

Møller, M.H. (2003) Grammatical Metaphor, Controlled Language and Machine Translation. Paper presented at the 8th EAMT Workshop, Dublin, 14-16 May 2003. Available: http://www.eamt.org/archive/dublin/MOELLER.PDF [Accessed 1 July 2006].

Mounin, G. (1963) *Les problèmes théoriques de la traduction*. Paris: Gallimard.

Nord, C. (1991) *Text Analysis in Translation*. Amsterdam: Rodopi.

_____. (1996) El error en la traducción: categorías y evaluación in Hurtado Albir, A. (ed.) *Estudis sobre la traducció*. Castelló: Universitat Jaume I, 107-129.

_____. (1997) Leicht – mittelschwer – (zu) schwer in Fleischmann, E., Kutz, W. und Schmitt, P. A. (Hrsg.) *Translationsdidaktik. Grundfragen der Übersetzungswissenschaft*. Tübingen: Narr, 92-102.

O'Brien, S. (2006) Pauses as Indicators of Cognitive Effort in Post-Editing Machine Translation Output. *Across Languages and Cultures* 7, 1, 1-21.

Olohan, M. (2003) Investigating Domain Conceptualisation and Scene Construal in Trainee Translators. *Meta* 45, 4, 593-608.

Olohan, M. and Baker, M. (2000) Reporting *that* in Translated English: Evidence for Subconscious Processes of Explicitation. *Across Languages and Cultures* 1, 2, 141-158.

Orozco, M. (2000) Instrumentos de medida de la adquisición de la competencia traductora: construcción y validación. Tesis doctoral, Universitat Autónoma de Barcelona.

Orozco, M. and Hurtado Albir, A. (2002) Measuring Translation Competence Acquisition. *Meta* 47, 3, 375-402.

Orozco Jutorán, M. (2003) The Use of Texts in Measuring Instruments Created and Validated to Measure Translation Competence Acquisition in Gerzymisch-Arbogast, H., Haijčová, E. & Sgall, P., Jettmarová, Z., Rothkegel, A., Rothfuß-Bastian, D. (Hrsg.) *Textologie und Translation*. Tübingen: Narr, 219-248.

Pápai, V. (2004) Explicitation. A universal of translated text? in Mauranen, A. and Pekka, K. (eds) *Translation Universals. Do they exist?*. Amsterdam/Philadephia: John Benjamins, 143-164.

Partington, A., Morley, J. and Haarman, L. (2004) *Corpora and Discourse*. Bern: Peter Lang.

Pearson, J. (1998) *Terms in Context*. Amsterdam/Philadelphia: John Benjamins.

Presas, M. (1997) Problembestimmung und Problemlösung als Komponenten der Übersetzungskompetenz in Fleischmann, E., Kitz, W. und Schmitt, P. A. (Hrsg.) *Translationsdidaktik. Grundfragen der Übersetzungswissenschaft*. Tübingen, Narr, 587-592.

Pym A. (1992) Translation error analysis and the interface with language teaching in Dollerup, C. and Loddegaard, A. (eds) *Teaching Translation and Interpreting*. Amsterdam/Philadelphia: John Benjamins, 279-288.

_____. (2002) Redefining translation competence in an electronic age. In defence of a minimalist approach. Available: http://www.tinet.org/~apym/online/training/competence.pdf [Accessed 10 October 2007].

_____. (2004) Text and risk in translation. Available: http://www.tinet.org/~apym/on-line/translation/risk_analysis.pdf [Accessed 10 October 2007].

Pym, A., Shlesinger, M. and Jettmarová, Z. (2006) *The Sociocultural Aspects of Translating and Interpreting*. Amsterdam/Philadelphia: John Benjamins.

Renzi, L., Salvi, G. e Cardinaletti, A. (2001) *Grande grammatica italiana di consultazione*. 3 vv., nuova edizione. Bologna: il Mulino.

Reuther, U. (2003) Two in One – Can it work? Readability and Translatability by means of Controlled Language. Paper presented at the 8th EAMT Workshop, Dublin, 14-16 May 2003. Available: http://www.eamt.org/archive/dublin/REUTHER.PDF [Accessed 1 July 2006].

Rojas, D. and Aikawa, T. (2006) Predicting MT Quality as a Function of the Source Language. *Proceedings of LREC 2006 – 5th International Conference on Language Resources and Evaluation*. Paris: ELRA, 2527-2530.

Sager, J. C. (1983) Quality and Standards – the evaluation of translations in Picken, C. (ed.) *The Translator's Handbook*. London: Aslib, 121-128.

Sager, J. C., Dungworth, D. and McDonald, P. (1980) *English Special Languages*. Wiesbaden: Brandstetter.

Savory, T.H. (1965) *The Language of Science*. London: Andre Deutsch.

Séguinot, C. (1989) The Translation Process: An Experimental Study in Séguinot, C. (ed.) *The Translation Process*. Toronto: H. G. Publications, 21-53.

_____. (1991) A study of student translation strategies in Tirkkonen-Condit, S. (ed.) *Empirical research in translation and intercultural studies*. Tübingen: Narr, 79-88.

Serianni, L. (1988) *Grammatica italiana. Italiano comune e lingua letteraria: suoni, forme, costrutti*. Con la collaborazione di A. Castelvecchi. Torino: UTET.

Simeoni , D. (1998) The Pivotal Status of the Translator's Habitus. *Target* 10, 1, 1-39.

Sinclair, J. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Snell-Hornby, M. (2006) *The Turns of Translation Studies*. Amsterdam/Philadelphia: John Benjamins.

Steiner, E. (2002) Grammatical metaphor in translation – some methods for corpus-based investigations in Hasselgård,H., Johansson, S., Behrens, B. and Fabricius-Hansen, C. (eds) *Information Structure in a Cross-linguistic Perspective*. Amsterdam/New York: Rodopi, 213-228.

Stratman, J. F. and Hamp-Lyons, L. (1994) Reactivity in Concurrent Think-Aloud Protocols. Issues for Research in Smagorinsky, P. (ed.) *Speaking about Writing. Reflections on Research Methodology*, Thousand Oaks, Cal.: Sage Publications, 89-112.

Taverniers, M. (2003) Grammatical metaphor in SFL: A historiography of the introduction and initial study of the concept in Simon-Vandenbergen, A., Taverniers, M. and Ravelli, L. J. (eds) *Grammatical Metaphor. Views from Systemic Functional Linguistics*. Amsterdam/Philadelphia: John Benjamins, 5-33.

Thompson, D. K. (1993) Arguing for Experimental 'Facts' in Science. *Written Communication* 10, 1, 106-128.

Tirkkonen-Condit, S. (1990) Professional vs. Non-Professional Translation; A Think-Aloud Protocol Study in Halliday, M. A. K., Gibbons, J. and Nicholas, H. (eds) *Learning, Keeping and Using Language*. Amsterdam/Philadelphia: John Benjamins, 381-384.

_____. (2000) Foreword: Challenges and Priorities in Process Research in Tirkkonen-Condit, S. and Jääskeläinen, R. (eds) *Tapping and Mapping the Processes of Translating and Interpreting*. Amsterdam/Philadelphia: John Benjamins, vii-ix.

_____. (2005) The Monitor Model Revisited: Evidence from Process Research. *Meta* 50, 2, 405-414.

_____. (2006) Think-Aloud Protocols in Brown, K. (ed.) *Encyclopedia of Language and Linguistics*. 2nd edition. Oxford: Elsevier.

Toury, G. (1991) Experimentation in Translation Studies: Achievements, Prospects and Some Pitfalls in Tirkkonen-Condit, S. (ed.) *Empirical Research in Translation and Intercultural Studies*. Tübingen: Narr, 45-66.

_____. (1995) *Descriptive Translation Studies and Beyond*. Amsterdam/ Philadelphia: John Benjamins.

_____. (2002) What's the Problem with 'Translation Problem'? in Lewandowska-Tomaszczyk, B. and Thelen, M. (eds) *Translation and Meaning Part 6*. Maastricht: Hogeschool Zuyd, Maastricht School of Translation and Interpretation, 57-71.

_____. (2004) Probabilistic explanations in Translation Studies. Universals – or a challenge to the very concept? in Hansen, G., Malmkjaer, K. and Gile. D. (eds) *Claims, Changes and Challenges in Translation Studies*. Amsterdam/Philadeplhia. John Benjamins, 15-25.

Vinay, J.-P. and Darbelnet, J. (1958) *Stylistique comparée du français et de l'anglais*. Paris: Didier.

Waddington, C. (1999) *Estudio comparativo de diferentes métodos de evaluación de traducción general (Inglés-Español)*. Madrid: Universidad Pontificia Comillas.

Williams, J. and Chesterman A. (2002) *The Map. A Beginner's Guide to Doing Research in Translation Studies*. Manchester: St. Jerome.

Wilss, W. (1982) *The Science of Translation. Problems and Methods*. Tübingen: Narr.