Simplification, Explicitation and Normalization

Simplification, Explicitation and Normalization: Corpus-Based Research into English to Italian Translations of Children's Classics

By

Margherita Ippolito

CAMBRIDGE
SCHOLARS
PUBLISHING

Simplification, Explicitation and Normalization:
Corpus-Based Research into English to Italian Translations of Children's Classics,
by Margherita Ippolito

This book first published 2013

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2013 by Margherita Ippolito

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-4568-X, ISBN (13): 978-1-4438-4568-7

To my parents and Francesco

De’ remi facemmo ali al folle volo
(Dante, *Inferno*, Canto XXVI)

TABLE OF CONTENTS

List of Tables ix

List of Graphs xi

Acknowledgements xiii

Introduction 1

Chapter One 3
A Theoretical Framework

Chapter Two 15
The Corpus

Chapter Three 25
Simplification

Chapter Four 33
Explicitation

Chapter Five 45
Normalization

Conclusions 57

Appendix A 61

Appendix B 73

Appendix C 77

Appendix D 93

Appendix E 99

Appendix F 111

Notes... 119

References 123

LIST OF TABLES

Table 2-1. The translational component
Table 2-2. Tokens in the translational component
Table 2-3. Number of writers for each country
Table 2-4. Number of texts translated in different periods
Table 2-5. The non-translational component
Table 2-6. Texts ordered by date of first publication and tokens
Table 2-7. Number of texts and tokens ordered by date of publication in the two subcorpora
Table 2-8. Differences between the two subcorpora
Table 3-1. Statistic results concerning MSTTR, MATTR, MTLD and HD-D values
Table 3-2. Statistic results of content words/tokens
Table 4-1. Gerunds introducing a non-finite construction
Table 4-2. Participles introducing a non-finite construction
Table 4-3. Relative pronouns in the translational subcorpus
Table 4-4. Relative pronouns in the non-translational subcorpus
Table 4-5. *A causa di* and *allo scopo di* in the corpus of children's literature
Table 5-1. Basic colour terms in the two subcorpora
Table 5-2. Concordance lines from the non-translational subcorpus for the shades of blue
Table 5-3. Concordance lines from the translational subcorpus for the shades of blue
Table 5-4. Nuances of the basic colour terms in the two subcorpora
Table 5-5. Observed and expected frequencies of basic colours and nuances in translations and originals
Table 5-6. The Pearson residuals in translations and originals
Table 5-7. Concordance lines from the translational subcorpus for the shades of blue
Table 5-8. Shades of red in the two subcorpora
Table 5-9. Observed and expected frequencies of red and its shades in translations and originals
Table 5-10. The Pearson residuals concerning red and its shades in translations and originals
Table A-1. MSTTRs in the non-translational component

Table A-2. MSTTRs in the translational component
Table A-3. MATTRs in the non-translational component
Table A-4. MATTRs in the translational component
Table A-5. MTLDs in the non-translational component
Table A-6. MTLDs in the translational component
Table A-7. HD-Ds in the non-translational component
Table A-8. HD-Ds in the translational component
Table B-1. Content Words/Tokens in the non- translational component
Table B-2. Content Words/Tokens in the translational component

LIST OF GRAPHS

Graph 3-1. MSTTRs, MATTRs, MTLDs and HD-Ds in original texts and in translations
Graph 3-2. Differences between the mean scores
Graph 3-3. Content Words/Tokens in originals and translations
Graph 5-1. Basic colour terms in the two subcorpora (instances pmw)
Graph 5-2. Nuances of the basic colour terms in the two subcorpora (instances pmw)
Graph 5-3. Association plot for the Pearson residuals
Graph 5-4. Shades of red in the two subcorpora
Graph 5-5. Association plot for the Pearson residuals

ACKNOWLEDGMENTS

I would like to thank Doctor Neil Millar, my MA dissertation supervisor, for his invaluable help and comments on my dissertation, from which the present work derives.

I am also grateful to Professor Rosella Mallardi who has encouraged and supported my research from the beginning.

I would also like to thank Doctor Eileen Mulligan for the linguistic revision of the work.

INTRODUCTION

The application of Corpus Linguistics (CL) tools and methodology to Translation Studies (TS) has opened new perspectives for the study of the translation process and product. CL can provide quantitative data to identify translation tendencies, recurrent patterns in target texts (TTs) and "the principles that govern translational behaviour" (Baker, 1993: 235).

Baker (1993: 235) was the first to encourage the cooperation between CL and TS, as, in her view, "large corpora will provide theorists of translation with a unique opportunity to observe the object of their study and to explore what it is that makes it different from other objects of study".

According to Baker (1999: 283) translations are specific communicative acts that have their own peculiar aims that occur in a well-defined context and that are governed by their own laws. Corpora can provide evidence that the language of translation differs from the language of non-translated texts.

Data collected from a comparable corpus of English translated and non-translated texts, namely the TEC (Translational English Corpus) and a subcorpus of the BNC (British National Corpus), has allowed Baker and her research team to infer the typical features[1] that characterize translations. They are defined as simplification, explicitation, normalization and levelling out and are considered "universal features of translation"[2] (Baker 1996:179-183). However, is the TEC sufficient to prove the existence of "translation universals"? Are these translation features valid in different socio-cultural and literary conditions? Do they dominate the translation process in other countries, apart from the UK? Do they apply to all text-types?

This study aims at testing whether the simplification, normalization and explicitation processes[3] dominate in the field of children's literature in Italy. It is hypothesised that a different translation context implies different translation processes. For instance, factors such as the target text (TT) socio-cultural background, the relations of power existing between the cultures of the languages involved in the translation process, the status accorded to translation itself in a particular country, the position of the TT within the literary canons of the TT culture, and even the specific expectations of the TT audience, can determine the choice of a particular translation process. Consequently, simplification, normalization and explicitation are not universal features, but only some of the possible translation processes that a translator can employ.

The chapter entitled “A Theoretical Framework” introduces the theories that connect translation with the cultural background of a particular speech community. It also places translation in the perspective of the “polysystem theory” (Even-Zohar 1979, 1990), and proposes an analysis of the translation of children’s literature within the Italian “polysystem”. Then, it focuses on “translation universals”—simplification, explicitation, normalization and levelling-out—and on how they have been tested by different scholars.

The chapter “The Corpus” introduces the corpus employed in the present study. It is a comparable monolingual corpus formed by twenty books of translated texts and twenty books of non-translated texts, addressed to 8- to 10-year-old children. This chapter also deals with the levels of comparability between the two corpus subsets.

The following three chapters deal with the translation processes taken into consideration in the present study. They all present a section on the most relevant approaches employed to examine these translation processes, a section on the methodology used in the present research, a paragraph on the results and a discussion of the results.

Simplification is tested by comparing lexical diversity and the information load in the two subsets of the comparable corpus. Explicitation is analysed by comparing the frequencies of gerunds and participles in non-finite constructions, of the relative pronouns *il quale, i quali, la quale* and *le quali,* and of the complex prepositional phrases *a causa di* and *allo scopo di* in the two subcorpora. Finally, normalization is examined by studying the occurrences of basic colour terms, the nuances formed by a basic colour term and a modifier, and the shades of red in the two subcorpora.

In the “Conclusions” the findings are summarized and interpreted, the implications of the findings in the field of TS are discussed, the limitations of the present study are presented and, finally, future research directions concerning the study of children’s literature in translation using the CL approach are suggested.

CHAPTER ONE

A THEORETICAL FRAMEWORK

1.1 Translation, Culture and Systems

The notion that translating is more than a simple conversion of a text from one language into another, is widely recognized in current Translation Studies (TS). Translating, in fact, involves a complex process of intercultural transfer. A source text (ST) is embedded in a specific cultural context, it reflects particular historical and social conditions and expresses values, knowledge and experiences belonging to that speech community. Consequently, the translation process implies the decoding and the recoding of the ST into a new language and a new culture for another group of people with a different worldview. Toury (1985: 18-19) points out that the translation process revolves around the target cultural system:

> Semiotically speaking, it will be clear that it is the *target* or *recipient culture,* or a certain section of it, which serves as the *initiator* of the decision to translate and of the translating process. Translating […] is to large extent conditioned by the goals it is designed to serve, and these goals are set in, and by, the prospective receptor system(s). Consequently, translators operate first and foremost in the interest of the culture *into* which they are translating, and not in the interest of the source text, let alone the source culture.

Bassnett and Lefevere (1990) advocate the “cultural turn” in Translation Studies, maintaining that attention to the text and a comparative study of the ST and of its TTs, must be integrated with careful consideration regarding the cultural contexts from which both of them emerge: “The object of study has been redefined; what is studied is the text embedded in its network of both source and target cultural signs” (1990: 12). A study of the translation activity and the translation process offers an insight into the “textual and extratextual constraints” (Bassnett and Lefevere 1998: 123) that govern the process of manipulation of the ST. The two fundamental translation strategies that guide any manipulation of the ST are defined *source-oriented* and *target-oriented strategies*. They have also been called "adequacy" and "acceptability" (Toury 1980), "retention" and "re-creation" (Holmes 1988), and "foreignization" "domestication" (Venuti 1995). The former tends to evoke the foreign cultural context, while the latter tends to remove it and to set the events into a familiar cultural background.

The choice of texts to be translated and the adoption of a particular translation strategy depend on a complex and dynamic system of interrelationship between cultural and literary forms, norms and genres dominant in a given culture. Even-Zohar (1990: 11) conceives literature as a *polysystem*: "a multiple system, a system of various systems which intersect with each other and partly overlap, using concurrently different options, yet functioning as one structured whole, whose members are interdependent". Translations are part of this *polysystem*. When the home culture and literatures are "young", "weak" and not fully developed, translations are likely to occupy a central and dominant position within the *polysystem*, because they can renew and bring new energy to the home cultural and literary *polysystem*: they can introduce new techniques, new themes and original stylistic features. Thus, in such contexts, translators will try to adhere to the ST as much as possible and the translation strategy that is likely to be adopted is source-oriented: "Under such conditions the chances that the translation will be close to the original in terms of adequacy (in other words, a reproduction of the dominant textual relations of the original) are greater than otherwise" (*Ibid.*: 50). On the other hand, when the home culture and literature is well established, translations occupy a secondary position within the *polysystem*, hence they do not exert any influence over the centre of the *polysystem* and are rather conservative, i.e. they preserve the accepted models and reinforce the existent canons: "translations, by which new ideas, items, characteristics can be introduced into a literature, become a means to preserve traditional taste" (*Ibid.*: 49). Consequently, the most suitable translation strategy chosen to maintain conventional forms and contents is target-oriented: translators will mitigate or omit cultural differences, will conform the stylistic features of the ST to those common in the target culture and language and will produce "non-adequate" (*Ibid.*: 51) translations with no distinctive aspects and that result to be very similar to texts produced in the target language.

1.2 Translation for Children in the *Polysystem*

Translation for children is a system of its own existing within the literary *polysystem*, in relation with all the other systems and subjected to various cultural and social forces. When children are the addressees, the main cultural constraints that direct the translation process often concern pedagogical and moral issues: i.e. the TT has to be suitable for children's reading skills and experiences and has to conform to what the TT society thinks is right and appropriate for children. Shavit (1981: 172), for instance, emphasizes that the two main translators' concerns are: "a) Adjusting the text in order to make it appropriate and useful to the child, in accordance with what society thinks is 'good for the child'; b) Adjusting plot, characterization and language to the child's level of comprehension and his reading abilities". Similarly Klingberg (1986: 10) points out that translators aim at "giving the readers a text that they can understand" and "contribut[ing] to the development of the readers' set of values". When translators think that a particular passage or reference cannot be understood by children, violates a taboo, can shock young readers, or is meaningless in the TT culture, they can decide to change the text, to adapt it or to omit some parts. This target-oriented translation approach is strongly supported by Oittinen who maintains that a translation for children must be faithful to the ST fundamental aims, namely catching young readers' interest and nourishing their imagination (2000). However, a text exhibiting loans, foreign proper names, describing unknown dishes and unfamiliar settings is unlikely to engage the young reader. The Finnish scholar asserts that "translating for children […] refers to translating for a certain audience and respecting this audience through taking the audience's will and abilities into consideration" (2000: 69), that is the reason why "names can be domesticated, the setting localized; genres, historical events, cultural or religious rites or beliefs can be adapted for future readers of texts" (*Ibid*.: 99).

However, translators can also opt for a source-oriented translation strategy. Traces of a foreign language and culture in the TT can give voice to the inherent *otherness* of the text and open a different perspective on the world. Translations can become a bridge which connect children to other cultural dimensions and favour their interest in, and appreciation of, the novelties coming from abroad. According to Klingberg (1986: 10) "another aim of translating children's books is to further the international outlook and understanding of the young readers" and "this aim will lead to the same adherence to the original".

The choice of a target-oriented or source-oriented translation strategy can depend on the position that translations for children have in the literary *polysystem*. If they occupy an important position in the *polysystem*, many books are imported from other countries, translations are made by famous writers, domestic literature will imitate foreign models and translators are more likely to retain significant features of the foreign text (see Even-

Zohar 1990: 46-47). Conversely, if they occupy a marginal position, few books are translated and seldom become bestsellers, translators tend to accommodate the text to the readers' tastes and expectations. For instance, Shavit maintains that translations of children's books into Hebrew occupy a peripheral position in the Israeli literary polysystem, therefore the translator "can permit himself great liberties regarding the text" (1981: 171). S/he can adjust the ST to models existing and appreciated in the target culture, add explanatory comments, delete unconventional passages, simplify themes, structures, characterization and employ the most widespread stylistic norms (*Ibid.*: 172-177).

1.2.1 Translation for Children in the Italian *Polysystem*

In Italy, translation for children plays a central role in the dynamics of the literary system addressed to children. From 1987 the number of translations has gradually grown, accounting for almost half of the total literary production addressed to young readers in 1997 (Zeli 1997). The latest data available shows that since then, there has been a slight decrease in the number of books translated for young readers: in 2009 44% of books for children and teenagers were translations[4]. A new, positive trend that is emerging in the last few years concerns the copyright sold to other countries. In 2001 486 titles were sold to foreign publishing houses while 1,250 were bought. In contrast, between 2009 and 2010 1,607 titles were sold and 1,283 bought[5]. Therefore, nowadays, the Italian literary system dedicated to young people seems not only eager to welcome new ideas and styles coming from abroad, but is also attracting the attention of other countries' literary systems.

This data demonstrates that Italian domestic children's literature was a "weak" system from 1997 to 2001. It needed models coming from abroad to update and develop it: hence "translated literature [was] not only a major channel through which fashionable repertoire [was] brought home, but also a source of reshuffling and supplying alternatives" (Even-Zohar 1990: 48). Ziliotto, a well-known Italian writer, editor and translator of children's literature in the 1990s claimed that: "The few interesting [Italian children's writers] belong to the new generation of Italian authors, influenced by the new trend coming from the North European countries, not concerned with moral or pedagogical goals" (Ziliotto quoted in Zeli 1997). Nowadays, a reversal seems to be happening, as the number of exported books for children is slightly higher than the number of imported ones. It seems likely that the impact of foreign literature has enriched and reshaped Italian domestic literature, so that now it is strong and influential enough to be able to assert itself not only in the Italian literary market but also outside the national borders.

Most translated books for children come from Anglophone countries. The figures provided by the Italian Statistical Institute (ISTAT) show that in 2009 books addressed to young readers, excluding school books, constituted about 40% of the whole, of which 29% were translated from English, 6% from French, 2% from German and 1% from Spanish[6]. The influence of Anglophone literary production on the Italian system is overwhelming if compared with translations from other languages. From this perspective it is not surprising that some Italian writers prefer signing their works with an English pseudonym—e.g. Peter Coolback stands for Gianluca Balocco, Geronimo Stilton hides Elisabetta Dami and Moony Witcher is the pseudonym of Roberta Rizzo—maybe to add an exotic nuance to their work or to conform to the most fashionable trends. Other traces of English language are present in the setting, in characters' names and in the titles of Italian books for children—e.g. "Britannia" is the main setting of *Merlino* (2009) by Luisa Mattia; "Scarlett" is the main character in a novel by Barbara Baraldi (2010) of the same name; "Mrs Butterfly O'Connor" and "Ezra Moore" are characters in the *Arx Mentis* saga (2006-2008) by Moony Witcher; *Il castello di Doom Rock* (2003) and *Le ombre di Halloween* (2007) are, respectively, the titles by Giovanni Del Ponte and Maurizio Giannini, awarded the children's literature prize "Premio Bancarellino" in 2004 and 2007. Further evidence of the strong relationship between the Italian and Anglophone literary systems for children is reflected in the prizes awarded to the best books for children. For instance, in 2011 three out of fourteen prizes of the "Premio Andersen"[7] were awarded to English speaking writers[8] (five Italian writers were awarded), while in 2012 two out of fifteen prizes were awarded to Anglophone writers[9] (seven Italians were awarded).

1.3 The Consequences of the Position of Translated Children's Literature on Translation

It is possible to hypothesize that the interconnections between Anglophone and Italian children's literature and the fundamental role that translations have played and continue to play in creating this network, affect the way of translating and the language of translations. It has already been emphasized that, according to Even-Zohar (1990: 50), when translations occupy a central position in the target language *polysystem*, it is likely that the favourite translation strategy will be source-oriented, as this foregrounds the ST and reflects the status that the foreign literature has in the receiving literary context. Hence, the ST language and culture provide new models and themes as well as different stylistic features, innovatory narrative techniques and languages. Translators, in seeking to adhere to the ST, tend to challenge the structural and grammatical norms of the source language and to stretch the potential of its vocabulary. In other words, the adoption of a source-oriented translation strategy might imply enlargement and innovation of the linguistic system of the TT language as a whole.

1.4 The Third Code

Translation scholars, such as Toury (1991: 50) and Even-Zohar (1979: 77), suggested that the translation process produces language patterns that are typical of translated texts. However, the first scholar who systematically theorized the language of translation as an autonomous and distinctive code was Frawley: "the translation itself […] is essentially a third code which arises out of the bilateral consideration of the matrix and target codes: it is, in a sense, a sub-code of each of the codes involved" (1984: 168). This notion of a "third code" differs radically from another well-known concept in TS, namely "traslationese", i.e. when the TT is characterized by some odd and incorrect expressions that derive from source language interferences. The "third code", on the contrary, is a non-intentional deviation from the standard target language, which is part of the translation process itself. In this view, it is recognized that translation has its own "distinctive nature", as it is "a communicative event which is shaped by its own goals, pressures and context of production" (Baker 1996: 175). Baker clarifies that:

> Given that all language is patterned, and that this patterning is influenced by the purpose for which language is used and the context in which it is used, the patterning of translated text must be different from that of original text production; the nature and pressures of the translation process must leave traces in the language that translators produce. (ivi: 177)

She continues by examining the possible reasons underlying the uniqueness of the language of translations: 1) it is strictly related to a foreign language and text; 2) it derives form a compromise between the target readers and context and the ST audience and context; and 3) it is influenced by the ideas that any social group develops about translation in a determined historical period (*Ibid.*).

Interest in the possibility to investigate the essence of translation language, i.e. those features that belong specifically to translated texts and differentiate them from original texts, has given rise to the application of

CL to TS. In 1995 Baker stated that "access to comparable corpora should allow us to capture patterns which are either restricted to translated text or which occur with a significantly higher or lower frequency in translated texts than they do in originals" (1995: 235). She added that "these patters may be quite local", when limited to one or two languages, or "global", when they occur in any translated text and "may tell us something about the nature of translated text in general and the nature of the process of translation itself" (1995: 236).

Baker and her research team based at the University of Manchester, have collected a large corpus, called *The Translational English Corpus* (TEC), which contained 10 million words in 2003 and can be accessed freely[10]. It consists of translations into English from different languages and covers four text-types: fiction, in-flight magazines, newspaper articles and biographies. Results from TEC are compared with a comparable subcorpus from the British National Corpus (BNC).

The explorations and the studies utilizing this comparable corpus have provided evidence that in fact there are some specific language features that occur regularly and frequently in TTs and that can be grouped around four fundamental tendencies: simplification, explicitation, normalization and levelling out[11].

1.4.1 Simplification

This is "the tendency to simplify the language used in translation" (Baker 1996: 181-182). Laviosa, for instance, finds some evidence regarding simplification in her study of a multi-source-language comparable corpus made up of translated and non-translated newspapers and narrative prose in English, called *The English Comparable Corpus* (ECC). She puts forward the following hypothesises: 1) in translated texts the range of vocabulary will be comparatively narrower (i.e. there will be less variety in the lexis); 2) the proportion of content words to running words will be relatively lower (i.e. the information load is lower) and 3) the average sentence length will be shorter (2002: 60-62).

Laviosa demonstrates the first hypothesis in three different ways. Firstly, she calculates the proportion of high frequency words over low frequency words in both subcorpora. She notices that it is lower in translations. Secondly, she considers the times each of the first 108 most frequent words occurs in each subcorpus—results show that the most frequent words are repeated more often in the translational subcorpus. Finally, she calculates the lemmas present in the head list and finds that there are fewer lemmas in translated texts. These findings are interpreted as support for hypothesis (1). Laviosa explores the ‘information load’, hypothesis (2), by measuring the proportion of lexical words over grammatical words. The results show that the ratio is lower in translations. Finally, she calculates the average sentence length, hypothesis (3), and finds that sentences are shorter in literary translations but not in newspaper articles. At the end of her study, Laviosa maintains that there is enough evidence to show that translated texts are simpler than non-translated ones. At the same time, she reflects upon the weak aspects of her research by saying that many translations come from Romance languages, so relatively few languages are equally represented in her translation subcorpus, and that the corpus is of a limited size (translated texts consist of one million word), so it is risky to draw solid generalizations (1998: 565).

1.4.2 Explicitation

This is the tendency to “spell things out rather than leave them implicit” (Baker 1996: 180). Features of “explicitation” are all those grammatical and lexical elements that are absent in the ST and that render the TT more precise and unambiguous. Olohan and Baker find that the optional *that* after the reporting verbs “say” and “tell” is much more frequent in TEC than in the subcorpus of originals in the BNC (2000). Olohan measures the frequency of *in order* preceding *to*, *for* and *that* and concludes that this optional chunk is markedly more present in TEC than in the BNC comparable subcorpus of original texts (2004). She also counts the instances in which the complementizer *to* is omitted after *help* and finds that it is omitted more frequently in the BNC subcorpus than in the TEC one, so the translations show a higher degree of syntactic explicitation. Mutesayire (2005) studies the frequency of reformulation markers, such as *that is*, *that is to say*, *to be (more) precise* and *namely*, and she observes that they occur significantly more often in translated than in non-translated texts.

1.4.3 Normalization

Also called “standardization”, “conservatism”, or “conventionalization”, normalization is “the tendency to conform to patterns and practices which are typical of the target language, even to the point of exaggerating them” (Baker 1996: 176-7). Olohan, for instance, aims at proving the hypothesis that the language of translation is more conservative, by studying the words ending with the suffix *-ish* (2004). Firstly she notes that this suffix especially modifies numbers and colours and then she underlines that, from a semantic viewpoint, it means “approximately” or “similar to”. She finds that in TEC there are more colours ending with *-ish* than in the BNC comparable subcorpus, but in the latter corpus there is a wider range of colours. She concludes that the higher frequency of colours with *-ish* suffix in the translational component of the corpus, can demonstrate that translators prefer giving a rough idea of colour: they opt for the most common source language basic colour and add *-ish*, because the ST colour may sound strange and unfamiliar in the English language.

1.4.4 Levelling-out

This refers to the fact that translated texts tend to “steer a middle course between any two extremes, converging towards the centre” (Baker 1996: 184), i.e. the texts of the translation subcorpus have more similar and homogenous language features than the texts composing the subcorpus of originals. According to Baker (1996: 184) and Olohan (2004: 100) it is difficult to test this tendency and there are no well-developed studies on this topic yet. Laviosa prefers to call this translation phenomenon “convergence” and studies it in TEC newspaper articles subset. She reports that the variance, “a statistical measure of the variability or dispersion of scores around the average value” (2002: 71), in respect to lexical density, type-token ratio and average sentence length, is lower than in the BNC comparable corpus (*Ibid*.).

1.5 Translation Universals

The regular occurrence of these translation tendencies in TEC provides evidence that they “may prove typical of English translated texts” (Laviosa 1998: 8) and has given rise to the hypothesis that they are also typical of any translated text, independently from the language pair involved in the translation process. Baker puts forward the idea that they are “universal features of translation” that “typically occur in translated text rather than original utterances and which are not the result of interference from specific language systems” (1993: 243). In her view simplification, explicitation, normalization and levelling out are hypotheses of universal features of translation that need to be verified in all languages.

Many studies based on the analysis of parallel corpora have been written with the aim of confirming or disconfirming these hypotheses and new translation hypotheses have also been proposed[12]. For instance, Øverås (1998) focuses on explicitation and the opposite process, implicitation, in a set of texts written in Norwegian and translated into

English and another set of texts written in English and translated into Norwegian. She studies the cases in which lexical and grammatical connectives are added or specified; she points out that explicitation is more common than implicitation in both translational corpora and that the level of explicitation is high in both translational subsets.

Scott (1998) analyses Clarice Lispector's Brazilian novel *A Hora da Estrela* and its English version *The Hour of the Star* by Giovanni Pontiero in parallel and studies how the negative type *nao* is translated in English. She observes that it is rendered by 72 different forms and is omitted 50 times. The final result is that the idea of nothingness emphasized in the redundant use of *nao* is diluted and normalized in the TT.

Zanettin (2000, 2001a, 2001b) investigates simplification in a parallel corpus formed by six novels and one short story by Salman Rushdie in English and Italian. He compares statistics regarding the lexical variety of each subcorpus against a reference corpus: Italian translations are compared with a corpus containing seventeen novels in Italian, while the originals by Rushdie are compared with a corpus of fictional texts extracted from the BNC. He notices that the standardized type/token ratio of the translations is closer to the average of the Italian reference corpus than that of Rushdie's originals to the average of the BNC section considered. Thus, he concludes that translations have a less varied lexis than their STs.

Kenny (2001) studies lexical normalization in a parallel corpus of German literary texts and their English translations and focuses in particular on creative lexis through the analysis of hapax legomena (words occurring only once in the corpus) in the German corpus. After having discarded all those hapax legomena that do not belong to the field of creative lexis, and after a careful comparison of these hapax with their translations in the translational component of the corpus, she finds that translators normalize in 44% of the cases. Normalization appears, thus, to be confirmed as a predominant translation feature.

Pápai (2004) using an English-Hungarian parallel corpus and a comparable corpus of Hungarian originals, examines explicitation devices and finds that translations are characterised by a higher level of explicitness than non-translations.

Xiao, He and Yue (2010) investigate a monolingual comparable corpus of translational Chinese versus non-translational Chinese, counting one million words for each subcorpus. Firstly they compare lexical density, lexical variability, mean sentence and the repetition of the high frequency words in the two parts of their corpus; secondly they study connectives as a device to create explicitation and, in the end, they study the passive construction as a grammatical device which is indicative either of a normalizing or of a “shining through” process in translation (see note 12). They come to the conclusion that simplification and explicitation processes act in Chinese translated texts too and that the use of the passive construction is revealed more as a “shining through” process (see note 8) than a normalizing one.

Interestingly, one of the few studies that contrasts with the hypotheses of translation universals, is on a corpus of eighty children’s books: forty texts originally written in Finnish and forty translated from English into Finnish. Puurtinen (1998) retrieves four different types of non-finite constructions in the two sets of the corpus and finds that the non-finite constructions are markedly more common in translations than in originals. As the non-finite constructions are supposed to raise the level of complexity of a text, because the relation between propositions is more implicit than in finite constructions, she maintains that the data from her corpus of children’s literature does not support the explicitation hypothesis. She also hints that this result may be due to the central position held by translations of children’s literature in the Finnish literary *polysystem*. Translations, in fact, account for 65-70% of the overall literary production addressed to children (1998: 526):

> However, bearing in mind the considerable proportion of translations in Finnish children’s fiction, they may have acquired a more central role and might even be an innovatory force; it is not implausible that translators of children’s books create, perhaps unintentionally, new norms which may gradually make their way into original Finnish children’s literature as well. (Ivi: 529)

CHAPTER TWO

THE CORPUS

2.1 The Comparable Corpus

The corpus that has been used in the present study is a monolingual comparable corpus[13]. It is made up of two monolingual corpora: one translational, i.e. formed of translations from English into Italian, and the other non-translational, i.e. consisting of originals written in Italian. The corpus includes 40 narrative prose works addressed to children with a total of 1,571,625 running words.

2.1.1 The Translational Component

The translational part of the corpus comprises 20 narrative works written for children by famous Anglophone writers and the number of the running words in the overall collection is 977,730 (tables 2-1 and 2-2).

Different cultural backgrounds are represented in the corpus, and although most authors are from England (35%) and the USA (30%), there are some who are from other English-speaking countries (35%), such as Ireland, Scotland, Wales, India and Canada (table 2-3).

The dates of first publication of the books chosen, spans from 1852 (*Uncle's Tom Cabin*) to 1997 (*Harry Potter and the Philosopher's Stone*), however the Italian translations selected for the corpus encompass a period starting from the 1980s to today. In particular, the translations published in the 1980s represent 25% of the corpus, those printed in the 1990s are 5%, while those published from 2000 to 2012 constitute the bulk of the corpus with 75% of the total (table 2-4). Therefore, the language represented reflects the lexico-grammatical features of contemporary Italian language and the typical stylistic tastes of the current Italian literary context addressed to children.

Titles and Dates of Publication	**Authors**	**Translators**	**Publishers**
Piccole donne (2012) *Little Women* (1868)	Alcott	Speckel	New Compton
Peter Pan nei giardini di Kensington (2011) *Peter Pan in Kensington Gardens* (1906)	Barrie	not mentioned	REA Edizioni
Il mago di Oz (1998-2006[14]) *The Wonderful Wizard of Oz* (1900)	Baum	Ceppellini	De Agostini
Il giardino segreto (1990-2011[15]) *The Secret Garden* (1909)	Burnett	Mirandoli	De Agostini
Alice nel paese delle meraviglie (2003) *Alice's Adventures in Wonderland* (1865)	Carroll	Bokos	De Agostini
Ladre di regali (2004) *The Present Takers* (1983)	Chambers	Scarlini	Giunti
Il GGG (1987) *The BFG* (1982)	Dahl	Ziliotto	Salani
Canto di Natale (2005) *A Christmas Carol* (1843-1858)	Dickens	Sala	Giunti
Il libro della giungla (1998-2009[16]) *The Jungle Book* (1894)	Kipling	Pasini	De Agostini
Il richiamo della foresta (2011) *The Call of the Wild* (1903)	London	Dauli	Edizioni di Karta