

University of St Andrews



Full metadata for this thesis is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

**MEANING AND TRANSLATION: THEORY AND PRACTICE OF
MACHINE TRANSLATION AS EXEMPLIFIED BY APPLICATIVE
AND COGNITIVE GRAMMARS**

CHANTAL GUERECHEAU

UNIVERSITY OF ST ANDREWS

OCTOBER 2003



Th E488

MEANING AND TRANSLATION: A REVIEW OF THEORY AND PRACTICE OF HUMAN AND MACHINE TRANSLATION AS EXEMPLIFIED BY APPLICATIVE AND COGNITIVE GRAMMARS

This dissertation is articulated around machine translation theory and practice on the one hand, and translation theories and issues on the other hand. Machine translation, being based usually on a single natural language analysis, fails to encompass the complex operations performed on source and target texts belonging to source and target languages and cultures. This thesis compares machine and human translation of a large corpus of sentences (*Le Petit Prince*, by Antoine de Saint-Exupéry), analyses how human beings translate, and how a translation programme processes the same text. The survey of the transfer processes - or shifts - displayed in several human translations (from French into English, German and Russian), as well as the analysis of the machine translation outputs (into English and German), show that machine translation should be reconsidered, not through source language analysis, but through the transfer operations performed on a source text to produce a target 'equivalent' text. Translation, seen as a cognitive operation, is here studied within the perspective of Applicative and Cognitive Grammars, formalisms rooted in Montague Grammar. More specifically, the Applicative and Cognitive Grammar, developed by Desclés, aims at determining a genotype language (a hypothetical universal semiotic system underlying all languages), and is developed around the theory of organised intermediate representation levels. Applicative and Cognitive Grammars can be seen as a step towards the setting-up of an interlingua architecture, which represents the next generation of machine translation systems. This research allows for an analysis of the deficiencies of a transfer translation system, as well as a better understanding of the processes of 'meaning transfer' in translation, seen as a semiotic operation.

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date: 14/11/2003 Signature of candidate: _____

I, Chantal Marie Guérécheau, hereby certify that this thesis, which is approximately 85,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date: 16/11/2003 Signature of candidate: _____

I was admitted as a research student in September 1993 and as a second year candidate for the degree of D.Lang in September 1994; the higher study for which this is a record was carried out in the University of St. Andrews between 1998 and 1999.

Date: 14/11/2003 Signature of candidate: _____

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of D.Lang in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 25.11.2003 Signature of supervisor: _____

TYPOLOGY CONVENTIONS

This thesis presents a multilingual corpus of sentences (translations of *Le Petit Prince*). The specific German characters (ä, ö, ü, ß) as well as the Cyrillic characters have been transliterated in order to allow readers outside the community of Germanists or Slavists to access the corpus. The following transliteration principles are used throughout the thesis:

- In German, ae for ä oe ö
 ue ü ss ß

- In Russian, the system retained (the simplified Library of Congress system) is the one used by the University Library of St Andrews:

a	for	a	r	for	p
b		б	s		с
v		в	t		т
g		г	u		у
d		д	f		ф
e		е	kh		х
e		ё	ts		ц
zh		ж	ch		ч
z		з	sh		ш
i		и	shch		щ
i		й	''		ъ
k		к	y		ы
l		л	'		ь
m		м	e		э
n		н	iu		ю
o		о	ia		я
p		п			

ACRONYMS

The following acronyms are used throughout the dissertation in order to avoid unnecessary repetitions:

AI	Artificial Intelligence
ACG	Applicative and Cognitive Grammars
CG	Categorial Grammars
EHT	English Human Translation
EMT	English Machine Translation
FAHQMT	Fully-Automatic High Quality Machine Translation
GHT	German Human Translation
GMT	German Machine Translation
HAMT	Human-Aided Machine Translation
MAHT	Machine-Aided Human Translation
MG	Montague Grammar
MT	Machine Translation
NLP	Natural Language Processing
PTQ	Proper Treatment of Quantification in Ordinary English
SL	Source Language
ST	Source Text
TL	Target Language
TT	Target Text
UT	Unit of Translation

ACKNOWLEDGEMENTS

I want to express my gratitude to the Trustees of the Association for French Language Studies, and of St Leonard's College. The completion of this research would not have been possible without the scholarships they awarded me. My thanks also go Mrs Donovan and her colleagues at the Ecole Supérieure d'Interprètes-Traducteurs in Paris, who initiated this study of meaning in translation. I also wish to thank Dr Clive Sneddon and Dr Roy Dyckhoff for their patience, and the time they gave me. The stimulating conversations we had were extremely helpful in building this bridge between literary translation studies and computational linguistics. Finally, I would like to dedicate this work to the memory of Dr Sándor Hervey, who introduced me to the world of semiotics, and helped me to find the missing link between human and machine translation.

TABLE OF CONTENTS

1	INTRODUCTION	1
2	SEMIOTICS, SEMANTICS AND MEANING	18
2.1	INTRODUCTION	18
2.2	SIGN, MEANING AND SEMIOTICS.....	19
2.2.1	<i>Introduction</i>	19
2.2.2	<i>Saussure's sign theory</i>	22
2.2.3	<i>Peirce's sign theory</i>	23
2.2.4	<i>Meaning and semiotics</i>	24
2.2.5	<i>Interpretation and 'unlimited semiosis'</i>	26
2.2.6	<i>Sign and text</i>	28
2.3	MEANING AND SEMANTICS.....	30
2.3.1	<i>Introduction</i>	30
2.3.2	<i>Meaning and reference</i>	31
2.3.3	<i>Truth and meaning</i>	32
2.3.4	<i>Formal semantics</i>	33
2.3.5	<i>Semantics and grammar</i>	34
2.4	CONCLUSION: SEMIOTICS, SEMANTICS, LANGUAGE AND COGNITION	35
3	TRANSLATION THEORY AND MEANING IN TRANSLATION	37
3.1	INTRODUCTION	37
3.1.1	<i>Definitions of translation</i>	37
3.1.2	<i>Approaches to translation</i>	38
3.1.3	<i>Brief historical survey of translation theories</i>	41
3.2	ISSUES RAISED BY DIVERGENT APPROACHES TO TRANSLATION	43
3.2.1	<i>Equivalence in translation</i>	43
3.2.2	<i>Unit of translation</i>	47
3.2.3	<i>Decoding/encoding approach</i>	49
3.2.4	<i>The two-phase/three-phase models of translation</i>	52
3.3	MEANING IN TRANSLATION.....	54
3.3.1	<i>Meaning in translation</i>	54
3.3.2	<i>Meaning and concept</i>	57
3.3.3	<i>Transfer of meaning</i>	58
3.4	SEMIOTICS, MEANING AND TRANSLATION.....	60
3.4.1	<i>Translation as communication, or translation as sign</i>	60
3.4.2	<i>Semiotic approach to translation</i>	63
3.5	CONCLUSION.....	68
3.5.1	<i>Filters, shifts, and decision processes</i>	68
3.5.2	<i>Textual dimension of translation</i>	71
4	MACHINE TRANSLATION	76
4.1	INTRODUCTION	76
4.1.1	<i>History of machine translation</i>	76
4.1.2	<i>Definition of machine translation</i>	78

4.1.3	Fully Automatic Machine Translation.....	81
4.1.4	Human-aided machine translation	82
4.1.5	Machine-aided human translation.....	82
4.2	NATURAL LANGUAGE PROCESSING	83
4.2.1	Natural and artificial languages	83
4.2.2	Formal languages.....	84
4.2.3	Syntactic parsing.....	88
4.2.4	Text typology and sublanguage.....	90
4.3	TYPES OF TRANSLATION SYSTEMS.....	92
4.3.1	Direct systems.....	93
4.3.2	Transfer systems.....	94
4.3.3	Interlingua.....	95
4.4	ISSUES IN MACHINE TRANSLATION	97
4.4.1	Segmentation and the unit of translation in MT.....	97
4.4.2	Ambiguity.....	98
4.4.3	Anaphora resolution.....	99
4.4.4	Idioms.....	100
4.4.5	Syntax/semantic separation.....	100
4.4.6	Machine translation evaluation.....	103
4.5	MEANING IN MT	105
4.5.1	Introduction.....	105
4.5.2	Meaning and knowledge representation.....	107
4.5.3	Meaning in MT, semantics and Transformational Grammar.....	110
4.5.4	Conclusion.....	112
4.6	CONCLUSION: HOW DOES MACHINE TRANSLATION RELATE TO TRANSLATION?.....	112
5	APPLICATIVE AND COGNITIVE GRAMMARS AND TRANSLATION.....	116
5.1	INTRODUCTION	116
5.2	TRANSLATION AND FORMAL SEMANTICS.....	119
5.3	LAMBDA-CALCULUS AND COMBINATORY LOGIC.....	121
5.4	MONTAGUE GRAMMAR (MG).....	123
5.4.1	Montague's Universal Grammar.....	123
5.4.2	The Proper Treatment of Quantification in Ordinary English (PTQ) and Syntax	125
5.4.3	Montague Grammar and Translation.....	126
5.4.4	Montague Grammar and Machine Translation.....	127
5.5	CATEGORIAL GRAMMARS.....	129
5.6	APPLICATIVE AND COGNITIVE GRAMMARS.....	131
5.6.1	Applicative languages	131
5.6.2	Cognitive Grammars	134
5.7	TRANSLATION, LEVELS OF INTERPRETATION AND APPLICATIVE AND COGNITIVE GRAMMAR	136
5.7.1	Translation and Genotype.....	136
5.7.2	Cognitive archetypes	137
5.7.3	Levels of representation, language and meaning.....	138
5.8	NATURAL LANGUAGE PROCESSING, COGNITIVE SCIENCE AND ARTIFICIAL INTELLIGENCE.....	138
6	HUMAN AND MACHINE TRANSLATION OF LE PETIT PRINCE	142
6.1	INTRODUCTION	142
6.2	TEXT TYPE, TEXT VARIETY, FUNCTION AND TRANSLATION OF <i>LE PETIT PRINCE</i>	145
6.2.1	Fairy-tales morphological analysis.....	147
6.2.2	Double articulation, ambivalence of children's literature	149
6.3	INTERPRETATION OF <i>LE PETIT PRINCE</i>	150
6.4	STYLISTIC FEATURES OF <i>LE PETIT PRINCE</i>	153
6.5	HUMAN TRANSLATION OF <i>LE PETIT PRINCE</i>	157
6.5.1	Drawings in <i>Le Petit Prince</i> and inter-semiotic references (Annex 8.1).....	161
6.5.2	Compensation (Annex 8. 2)	163
6.5.3	Shifts of cohesion and coherence.....	167
6.5.4	Convergence between the target languages syntax and/or lexicon through translation	174
6.5.5	Units of translation.....	175
6.6	MACHINE TRANSLATION OF <i>LE PETIT PRINCE</i>	177
6.7	COMPARISON OF HUMAN AND MACHINE TRANSLATION OF <i>LE PETIT PRINCE</i>	182

6.7.1	<i>Units of translation</i>	182
6.7.2	<i>Anaphora resolution</i>	183
6.7.3	<i>Reflexivation</i>	185
6.8	CONCLUSION: HUMAN TRANSLATION, MACHINE TRANSLATION, AND THE TRANSFER/INTERLINGUA DEBATE.....	186
7	CONCLUSION	191
8	ANNEXES	197
8.1	DRAWINGS IN <i>LE PETIT PRINCE</i>	197
8.2	COMPENSATION	200
8.3	MACHINE TRANSLATION ERROR ANALYSIS –ENGLISH	219
8.4	MACHINE TRANSLATION ERROR ANALYSIS -GERMAN.....	225
8.5	HUMAN AND MACHINE IDENTICAL TRANSLATIONS (IN TERMS OF SYNTAX) - ENGLISH	229
8.6	HUMAN AND MACHINE IDENTICAL TRANSLATIONS (IN TERMS OF SYNTAX) - GERMAN	243
8.7	BARCELONA MACHINE TRANSLATION TECHNOLOGY (BY COURTESY OF GLOBALINK).....	255
8.8	NUMBER OF SENTENCES IN THE HUMAN TRANSLATIONS (INDICATING A TENDENCY TO TRANSLATE SENTENCE-BY-SENTENCE).....	260
9	BIBLIOGRAPHY	262

1 INTRODUCTION

At least 6000 languages are currently acknowledged in the world, but most linguists agree on the existence of hundreds (or even thousands) more. Despite this diversity, the faculty of language is common to all human beings, and communication between different peoples seems to have existed as long as the human species. Whether for political or economic reasons, linguistic barriers had to be overcome, and translators (or rather interpreters at earlier times) held a special position. In Ancient Egypt, they were given the title of 'Prince', and in the Middle Ages, some translators paid a very high price for supposed "translation errors". E. Dolet, who introduced the terms 'traduction' (translation) and 'traducteur' (translator) in the French language, was burnt at the stake with his books in Paris in 1546, having been accused of 'mistranslating' one of Plato's works. The misinterpretation was actually an alleged addition to a Platonic text, which cast doubt on the existence of eternal life¹. Fortunately, no translator would be burnt today, but translation errors can still have terrible consequences. The recent release of classified documents led some historians to suspect that a misunderstanding (and misinterpretation) between the Americans and the Japanese might have been at the origin of the bombing of Hiroshima and Nagasaki. It is now clear that the Japanese would not have been able to go on fighting for very long, when they were presented with the American ultimatum. Their request for a delay to give an answer was translated literally, and the American officials in charge of the crisis seemed to be unaware of the Japanese cultural and social norms. A less literal and more explicit translation (making clear that the Japanese were asking for a decent retreat), and a better knowledge of the cultural and social customs might have changed history. Nowadays, the main threat to translators is said to be the development of machine (or automatic) translation, the computers being supposed to replace human translators in the future.

¹ Baker 2001: 410; 416.

In a world now dominated by ever increasing information exchanges, as the worldwide Internet network shows, multilingual translation is even more needed, and there are not enough translators to cope with the world demand for translation. Governments and non-governmental agencies have launched many research programmes in Artificial Intelligence in the hope of devising translation software, that would be able to process the enormous amount of translations needed every day. TAUM, EUROTRA, SYSTRAN are amongst the most well-known, the programmes being now widely available and used daily to translate huge amounts of texts every day, but others such as the CETA project (Grenoble, France), META, SUSY, ARIANE, KBMT89, or the Rosetta project (Philips research Laboratory, Eindhoven), are instances of several attempts at reproducing mechanically the process of translation.

Fascinated by languages and the faculty of language (as a means of communication), human beings have continuously explored this field of knowledge, following different paths: the quest for a universal "language of the thought"²; the inventory, comparison and grouping of languages into main "families"; attempts to decipher "languages" in other species (such as ants and bees); the teaching of language to apes; and the creation of artificial languages³.

"Artificial languages are products of human design, based on and derived from natural languages, which do not have the flexibility and multifunctionality of natural languages. Such languages are usually functionally restricted to the conveyance of information and therefore exclude connotative, emotive, aesthetic and other meanings" (Sager, 1993: 258; 259)

The last century saw the development of a new autonomous field of research (translation studies) as well as the emergence of numerous research teams involved in natural language processing (NLP) and machine translation (MT). Natural Language Processing (NLP) is defined as the processing by a computer of sentences, texts, etc. in natural language. The term machine translation will be used in this dissertation in the sense of 'automatic translation of a text by a computer', even though in the larger sense of the expression, 'machine translation' encompasses

² Eco, U.1994.

³ Esperanto is certainly the most common one, even though many other formal languages are developed by computer scientists.

machine-aided human translation, and human-aided machine translation. The first translation software products to be devised (called direct systems) were based on the matching of linguistic units between source and target texts (on a word-by-word basis), with some limited syntactic reordering. This principle was initially inspired by the belief that translating was no more than the recoding of sentences (natural languages being seen as similar to artificial languages, or codes, such as the ones used to crypt secret messages). More powerful software programmes were then devised, but they were still based on a purely linguistic approach to translation, and limited to the sentence analysis, as put forward by J.C. Catford (1965: 1). A purely linguistic approach to translation tends to impose systematic transformation rules (active/passive; verb/substantive; etc), and is usually used to build translation programmes. The linguistic approach was later broadened to include the text-linguistic level of register analysis, the discourse analysis (thematic structure, coherence⁴, and cohesion⁵), and the pragmatic analysis (speech acts, language and text functions), but computational linguistics faced many difficulties in implementing these results in machine translation.

The so-called second generation translation software (transfer systems), was expected to provide 'real' translation, close enough to human translation. The transfer approach views translation as a three-stage approach: the analysis of the input into a source-language syntactic structure representation, the transfer of that representation into the corresponding target-language structure, and the synthesis of the output from that structure⁶. Unfortunately, machine translation researchers quickly faced the problem of ambiguity, and it became obvious that only knowledge of the manner in which the natural world was organised could lead to a satisfactory interpretation of the texts submitted to the computers.⁷ The last few decades have been marked by the focus on understanding and semantics, as well as an emphasis on syntactic and lexical

⁴ Coherence is defined in general terms as the agreement of the text with its situation. (Shuttleworth, 1997: 19). A more precise definition is given by S. Blum-Kulka : "Coherence can be viewed as a covert potential meaning relationship among parts of a text, made overt by the reader or listener through processes of interpretation." (Blum-Kulka, in Venuti, 2000: 298)

⁵ "Cohesion will be considered as an overt relationship holding between parts of the text, expressed by language specific markers" (Blum-Kulka, in Venuti, 2000: 298)

⁶ Somers, in Baker (ed.), 1998: 145.

⁷ Mc Donald, R. 1979: 92.

disambiguation in natural language processing. At the same time, despite the growing awareness among translation scholars (as well as linguists and nowadays computer linguists) of the textual dimension of translation, most current translation software products are unable to integrate the text as a whole as meaning generator. Syntactic parsers, on which machine translation software relies heavily, process texts sentence after sentence, and are unable to cope with anaphora, while a human translator does not seem to hesitate when facing a pronoun, even if the antecedent is located two sentences above. Most of the translation software currently commercialized is based on a direct or a transfer system, but more and more computational linguists advocate the so-called third-generation systems (interlingua), based on an intermediate artificial (pivot) language, marking a renewed interest for the long sought after 'universal' language. Such systems will become operational only when the cognitive operations involved in the translation process will be fully understood.

Translation studies have been enriched by the results of research led in linguistics, sociology, philosophy, psychology, and the more we know about languages, the more complicated translation appears. Although translation benefited from linguistics, it is no longer wholly included in linguistics, but is rather located at the crossroads of different disciplines, translation dealing with the use of languages in context, in a situation of communication. Linguistics is concerned with the study of language, while translation deals with the interpretation of the content of texts (as actualisation of language in context). Translators and scholars do not see translation any longer as a simple transfer of words from one language into another, or as "saying the same thing in another language". But most definitions remain vague in their phrasing. Many dictionaries simply provide lists of synonyms referring to the notion of 'transfer' - *render, rephrase, reword, transmit, reexpress, transmute, transmogrify, interpret, convert, transform, transpose, express, transfer, turn* - 'from one language into another' (mentioning an

'equivalence' between the source and the target entities)⁸, but usually fail to state what is being transferred.

Translation deals with a 'transfer' (of 'meaning', 'content', 'message') between a source text and a target text, involving a transmitter (the writer of the source text), a receptor/transmitter (the translator) and an end-receptor or several receptors (the reader(s) of the target text). In most contemporary translation theories, the source and target texts are said to be in an 'equivalence' relationship, i.e. there is a relationship between the functions of the ST and the TT in their respective source and target cultures and systems.

Until the 60s, research in translation had been heavily imprinted by the structuralist view of language. Works of R. Jakobson⁹ and G. Mounin¹⁰, for instance, exemplify this linguistic approach to translation. Although useful in many respects, the results fail to account for the multidimensional aspect of translation: operations performed on texts belonging to two different languages. Similarly, the generative approach to linguistics fails to provide a complete and reliable account of translation, despite its beneficial contribution to the development of machine translation. Chomsky himself acknowledged the limitations of his research on universals in languages, as far as translation is concerned¹¹.

The gradual shift of emphasis from interlingual relationships to inter-textual relationships (attempts at specifying an intertextual invariant for translation) is considered by G. Toury (1990) as one of the main achievements of the theory of translation in the last decades. C. Nord (1991) defines translation as 'intercultural communication', hence emphasizing the importance of the texts (messages) as well as the role of the translator who produces a 'communicative instrument (target text) for the target culture', or a 'target-culture document of a source-culture communication'.

⁸ Picken, 1989: 12.

⁹ Jakobson, R. 1959.

¹⁰ Mounin, 1963.

¹¹ "The existence of deep-seated formal universals...does not, for example, imply that there must be some reasonable procedure for translating between languages" (Chomsky, 1965: 30)

Although the linguistic and textual structures (of the source text) and their relationships with the system and norms of the SL must be thoroughly understood in the translation process, the decisions which the translator has to make to produce the target text must also be accounted for. Both source and target texts aim at conveying a message, and are therefore determined by the communicative situation in which they are produced. At the analysis stage, the translator determines the situation surrounding the production of the source text, and he must similarly take into account the social background, world-knowledge and communicative needs of the recipients, as well as the situation in which the TT is received (function of the text, or *skopos*)¹². In this communicative situation, the translator plays an essential role. D. Robinson (1991) argues for a complex physicalist approach to translation, determined by the idiosomatic (personal and unique) response of the translator to the ST, as well as by his idiosomatic creation of the TT. He advocates a more people-centred theory of translation, which should focus on what translators (and readers) do, and on how they interact in the activities surrounding translation. The target text produced by the translator is built from linguistic elements **and** non-linguistic ones, the latter depending on the world-knowledge of the translator. The essential role played by the translator (thinking human being) points to the cognitive dimension of translation, as underlined by the observation and study of conference interpreting.

The interpretive approach to translation as developed by the Paris School¹³ breaks free from the structuralist and the generativist views, by emphasizing the difference (considered as fundamental) between the linguistic meaning and meaning actualised in context (*sense*). The interpretive approach, also sometimes referred to as the 'interpretative approach' (which is different from Newmark's interpretative translation), was developed in the late 1960s on the basis of research in conference interpreting. It was later extended to the written translation of non-literary or 'pragmatic' texts, and more recently also used as a tool for the study of literary texts, the linguistic form being seen as part of the overall sense. This theory is based on the distinction between linguistic meaning and non-verbal sense, and views the transfer of meaning

¹² Vermeer, 1987; 1994. Nord, 1991; 1997.

¹³ Ecole Supérieure d'Interprètes-Traducteurs (E.S.I.T.)

as the core issue in translation. The Paris School describes the translation process as consisting of three stages (interpretation or exegesis of discourse, de-verbalization, and reformulation), and puts emphasis on the mental and cognitive processes involved in translation. Proponents of this approach stress that the main difficulty lies in the fact that this meaning is simultaneously carried by the written words and somehow independent from them. Meaning is therefore different from the linguistic units (words) involved. Words have linguistic meaning, while the object of translation is the "sense", carried by the text as a whole. This sense is built by the reader throughout the reading, and is the combination of linguistic meaning and extra-linguistic knowledge. Central to this "sense building" are the cognitive operations taking place in the mind of the translator, even though these operations are not really defined, the mind being seen as a 'black box'. D. Seleskovitch and M. Lederer (1984) explain the relative failure of machine translation by the fact that early research in machine translation focused on the establishment of strictly linguistic correspondences between language pairs (linguistic approach to translation). The first translation programmes, known as direct systems, were developed on this principle, and even the second-generation software products (transfer systems) are still imprinted by this linguistic correspondence principle, even though it is clear nowadays that true translation is far from being a simple replacement of a source string of words by a target string of words (rearranged to comply with the syntactic requirements of the target language). Research done over the past fifty years has shown that translation is a multidisciplinary communicative process dealing with texts (source and target) said to be functionally equivalent. The source text and the target text are characterized by a set of features (pragmatic, semantic, syntactic, and stylistic elements) which are different, being respectively linked to the source and the target cultures. Functional equivalence aims at identifying and isolating the ST elements which must be preserved or adapted in translation in order to create a 'function-in-culture' of the TT similar to the 'function-in-culture' of the ST (Nord, 1991). The source and target texts are composed of words (linguistic units), but the establishment of linguistic correspondences alone does not account for the complex translation process. House (1977) stresses that the translation process is concerned not only with language, but also with the language user (through geographical origin,

social class and the time) and the language use (medium, participation, social role relationship, attitude, and province)

The interpretive theory of translation also puts emphasis on the translator as the central agent in a communicative operation. This approach to translation is based on the concept of 'deverbalization', which maintains that translation is the transfer of the meaning of a message. This transfer is supposed to be operated on a cognitive level without any words (but rather in terms of 'ideas' or 'concepts'), although very little research has been done on the formalization of this 'meaning transfer'. The main methodological point (developed from research on interpreting) lies in the need for the translator to free him/herself from the imprint of the source language in order to render a good 'equivalent' message in the target language. The interpretive approach sees translation as the re-expression in a different language (through the use of different lexical and syntactic means) of the 'thing' referred to. Theoretically, this approach (based on numerous observations, particularly of oral translation – interpretation) seems to give a better account of the processes involved, but raises probably more questions than it provides answers. It is quite obvious that good translators do not match lexical and syntactic features on a one-to-one basis when 'transferring' 'meaning' from one language to another, but the way this transfer is done is mainly unknown. The interpretive approach to translation tends to rely on the cognitive competence of the translator, who is supposed to transfer unconsciously the meaning of the message conveyed, making use of both source and target linguistic systems. This approach insists on the fact that no systematic transformation operations should be needed, since there are many possible interpretations and rephrasing. Up to now, this approach has excluded any fully automatic machine translation.

Translation studies also benefited from advances in cognitive sciences, aiming at describing and analysing human understanding and conceptualising (Edelman and Tononi, 2000). The translator is now reckoned to be an essential element in the translation process, and more emphasis has been put over the last few years on the study of the cognitive operations performed by the translator, for example through the Think-Aloud Protocols (Lörscher, 1991b).

The translator – far from decoding-encoding in a predictable way, following regular patterns – plays an active role as creator of meaning (Seleskovitch, 1975). G. Steiner (1998) pinpoints the main deficiencies of most translation theories that treat ‘meaning’ as more or less transcendental.

« The translator must actualize the implicit ‘sense’, the denotative, connotative, illative, intentional, associative range of significations which are implicit in the original, but which it leaves undeclared or only partly declared simply because the native auditor or reader has an immediate understanding of them ». (Steiner, G., 1998: 291).

Translators are supposed to extract and transfer ‘naturally’, or ‘spontaneously’, the meaning contained in the source text, and they are expected to know intuitively what this meaning is. Meaning being traditionally the object of study of philosophy, semantics and logic, there is a gap between meaning as defined by philosophers and meaning as defined by translators. The former approach it through notions such as ‘reference’, ‘denotation’, ‘propositional semantics’, ‘truth conditional’, ‘conceptual structures’ and so on (Frege, 1892; Church, 1951; Carnap, 1956; Quine, 1959; Keenan, 1975; Cresswell, 1985). The latter focus their discourse on denotative and connotative meaning, phatic meaning, sentence meaning and text meaning, meaning as function, and so on (Catford, 1965; Delisle, 1980; Seleskovitch and Lederer, 1984; Larose, 1989). To claim that translators transfer meaning is one thing, to define meaning and describe the processes involved in its transfer from a source text to a target text is an entirely different (and immensely complicated) task. Linguistics failed to provide an uncontroversial theory of language that would explain translation, as well as allow for the formalisation of the process of translation. But, one can follow Steiner’s view and accept that such an endeavour is impossible, on the grounds that the study of language is not a science (as physics and mathematics are) and cannot be described in terms of laws and effects.

« A ‘complete translation’, i.e. a definitive insight into and generalisation of the way in which any human being relates word to object would require a complete access to him on the part of his interlocutor. The latter would have to experience a ‘total mental change’. This is both logically and substantively a meaningless notion. It could never be shown to have taken place. All

discourse, all interpretation of discourse works at a word-for-word and sentence-for-sentence level. There is no privileged access to underlying totality. » (Steiner, G., 1998: 309-310).

Early computational linguists believed that they could easily formalise translation operations, since it was then assumed that translation was no more than a complex coding operation. Most of them realised (fast enough) that they were dealing with a manifestation of human intelligence, and understood that their software was unable to reproduce the cognitive processes at stake in translation, for lack of a complete understanding of the complex operations performed during the act of translation. There is no doubt that we have only a very partial and fragmented knowledge of the mind and its functioning. Nevertheless, studies in physiology, psychology and cognitive sciences give new insights into some functions of the brain (Langacker, 1991a, 1991b; Edelman and Tononi, 2000). Observations of interpreters at work confirmed (or even originated) some hypotheses on memory, prediction, understanding and cognition (Seleskovitch, 1975; Larose, 1989; Lörcher, 1991a). Similarly, the present study of a written translation provides some interesting information on the surface (or even supposed 'deep' structures) but aims also at tracking the cognitive operations performed in the act of translation. According to Steiner (1998: 309), "language is, at vital points of usage and understanding, an idiolect", hence irreducible to definitive formalisation. On the other hand, if one assumes that there is a transfer of 'meaning' during the translation process, a comparison of the lexical, morphological and syntactic features of the source text and its target texts will get us closer to this shadow 'element' that seems located in the mind and outside of it, partly individual but also shared. A precise 'scientific' description of the 'meaning' of the text might be impossible. Still, this 'meaning' expresses cognitive operations performed at different levels in the brain. If translation consists of "mentally dissociating a concept from its written form in order to match signs drawn from a different linguistic system with it" (J. Delisle, 1988), an interpretive analysis of the process of translation is necessary. This analysis must take into account the 'meaning' of the messages, the act of communication, the cognitive operations involved in the act of translation, as well as the functions of the source and target texts in their respective cultural and linguistic system.

Translation lies at the crossroads of linguistics, philosophy, psychology and sociology, whereas artificial intelligence (AI) attempts to reproduce the activity of the human brain. Artificial intelligence is defined by Wilks (1979: 27) as "the use of computational methods for the simulation of distinctively human intellectual behaviour, by means of complex knowledge structures and their manipulation". Although Bar-Hillel argued in 1962 that machine translation was not only practically but also theoretically impossible¹⁴, the output of some software refutes this claim. Translations produced by computers are far from perfect (some are even not up to the standard of a bad word-for-word translation), but the progress displayed in AI may give hope for some noticeable improvement in the future. In order to overcome all the difficulties raised by natural language processing and machine translation, it is necessary to tackle the problems from a multi-disciplinary approach, and to understand better the link between language and cognition¹⁵, through the analysis of translation, as a manifestation of human behaviour. It is necessary to study and analyse the way human translators operate, in order to build translation software producing a target text as close as possible to the ideal target text a human translator would conceive. In other words, one should explore the possible formalisation of the cognitive operations performed in the human mind (hence getting closer to proper "artificial intelligence"). Assuming that translating is transferring the "sense" from a source text in a source language to a target text in a target language, and assuming this "sense" is an invariant element (by approximation, as it will be discussed in the subsequent chapters, meaning being too complex to be reduced to a mathematical constant), one possible way towards a formalisation of translation processes lies in the study and comparison of source and target texts. For practical reasons, it is necessary to select the source text according to some well-

¹⁴ "Expert human translators use their background knowledge, mostly subconsciously, in order to resolve syntactic and semantical ambiguities which machines will either have to leave unresolved, or resolve by some 'mechanical' rule which will ever so often result in a wrong translation." (Bar-Hillel, 1962)

¹⁵ Cognition is defined as "the functional study of intelligence (human, animal, and artificial) as we apprehend it in phylo/ontogenetic manifestations (perception, language, reasoning, memory, strategy and planning, learning, categorization...), in stable as well as in developing behaviour. The ultimate aim of the cognitive sciences is to understand both formal information processes and the possibilities for their implementation in brains and in machines." (Desclés, 1989: 121)

defined limiting criteria, since translation is concerned with texts before being concerned with languages. The purpose is not to infer any kind of translation of 'languages', but to study the (mainly unconscious) operations governing the transformation of a source text into a target text, even though the source and target texts are in two different languages. Moreover, a text being the expression of intertwined language and culture, it is prone to innumerable interpretations and rephrasings. One selection requirement for the text is to be extensively translated into as many languages as possible, in order to study the translation *process* itself - not the transfer of the text from one language into another single one. Results of a comparative study on a single pair of languages might be biased by similar lexicon and syntactic structures between two historically and/or geographically close languages. It is also necessary to select a text that can (materially) be parsed by translation software. An extract from a book written by M. Proust, for example, would be beyond the limits of even the most powerful translation software currently available. Most translation software programmes currently available process a text through a syntactic parsing (determining the syntactic category of each term of each sentence), and grouping the terms according to their function within the sentence. A too long sentence creates too many possible parse trees, and renders a mechanical translation (following the current systems' architectures) impossible. This text should contain a limited range in terms of syntax and lexicon, but as wide as reasonably possible within the previous limits. The text must also belong to a sublanguage (defined as a subset of a natural language), but this sublanguage must contain more lexical and syntactic variety than sublanguages such as weather forecast reports, or commercial letters. These kinds of texts are said to be "translatable" by computers, while the system simply matches equivalent fixed sentences according to some selection criteria. The principle is closer to an expanded dictionary than to a proper translation system, the programmes being mere advanced bilingual dictionaries, the syntax being practically non-existent.

Le Petit Prince by Antoine de Saint-Exupéry fulfills the selection criteria for this study: translated in many languages (more than one hundred so far), a highly characterised text type (children's literature), and a text belonging to a sublanguage (fairy tale containing a limited

number of words and syntactic structures). In its most general accepted sense, the term 'sublanguage' is taken to mean 'a special language of a particular domain'. From the point of view of MT system designers, the term 'sublanguage' refers to 'the grammar, lexicon, etc. of a particular text-type in a particular domain. *The Little Prince* is characterised by a limited lexical and syntactic range, as developed in chapter 6. Last but not least, children's literature is characterised by very defined norms, both in the source and the target cultures, and the source texts and target texts fulfil specific functions. Moreover, it is a particularly interesting object of study of 'meaning transfer', since belonging to children's literature, it is prone to several interpretations, while at the same time containing features that would not be acceptable in 'our world' (animals speak, objects appear at will, and so on). On the other hand, the very characteristics of children's literature raise many difficulties to a translation software, as underlined by McDonald (1979: 93)

« On occasion, it has been argued that children's stories provide suitable texts for experimental work because of their shortness and because of the simplicity of the language. Actually, the simplicity of the language is largely suppositious; perhaps there are fewer involved sentences, but all of the other difficulties remain. In addition, children's stories may introduce semantic and situational combinations, which would be quite unsuitable for other types of translation. If children fly and spiders talk and wizards convert frogs into princes and vice versa, the morphology and syntax are probably not significantly altered, but the semantic apparatus, depending on how it is conceived, may have to be considerably different from the semantic apparatus necessary for a suitable translation of a newspaper editorial or a text in biology."

The comparison of the source and respective target translations will allow a discussion of the equivalence principle (between the source text and the target text), as well as a possible redefinition (or confirmation) of the concept of 'unit of translation'. Moreover, the study of the shifts in translation, and more particularly the determination of any systematic translation strategy between a source and a target language would contribute to a better understanding of the cognitive processes displayed in translation. The automatic translation of the corpus of

sentences aims at determining the limits of the software selected. The next stage (comparison between machine translation and human translation) should be helpful mainly to computational linguists, faced with traditional difficulties such as anaphora resolution, or ambiguities. Even though machine translation, at its present state of development, does not pretend to mimic human translation (machine translation transforms natural languages into artificial languages and apply formal operations on the strings of symbols so determined), a better understanding of the transfer processes performed by human translators should be of some use (even if limited) to the design of the next generations of translation software.

The first translation software, as well as most of the software currently available, is devised to conform to the principles governing computer sciences. Computers were originally conceived as information processors, following a slightly reductionist analysis of human behaviour. The purpose of such an analysis was manifold¹⁶:

- to set up symbolic representations organised in systems;
- to define operations on these representations within systems;
- to identify the physical organs in which the representations are implanted; and
- to determine architectures such that the operations called for by models of observed behaviour may be executed.

Following the imprint of the Turing machine model, and of machine functionalism, human thought is traditionally described as a system that manipulates symbols, while more recent studies in psychology emphasise the processing and storing of information (Vignaux, 1991; Wagner, 1998; Edelman and Tononi, 2000). The main issue faced by cognitivists is the 'location' and 'organisation' of this symbolic representation, or rather representations, assuming that cognitive representations are situated at different levels (Desclés, 1989, 1994a).

¹⁶ Desclés, 1998: 121.

According to the neo-connectionist hypothesis (Rumelhart et al., 1986; Desclés, 1989), direct relations exist between external representations and neural models, even though there is a gap between symbolic representations and their 'distributed' sub-symbolic coding. An answer is provided by the 'intermediate representations' hypothesis postulating the existence of intermediate symbolic representations between external representations and the representations directly compatible with the observations of neuro-sciences. This hypothesis sees cognition as a succession of "representation changes" between levels of representation, and is called the "principle of generalised compilation" (Desclés, 1989, 1991). A compilation process consists of "a set of programs, arranged in a hierarchy, that automatically ensures the translation of external expressions accessible to human users into internal representations directly compatible with the electronic structures of information processing machines (computers)" (Desclés, 1998: 125). New more powerful formal languages such as applicative or cognitive grammars, have been developed, and they integrate the latest results of research in cognitive science and computer linguistics. The Applicative and Cognitive Grammar, as developed by J.-P. Desclés (1990) from the works of S. K. Shaumyan (1971, 1977, 1987) and R. Langacker (1987, 1991a, 1991b) and rooted in Montague Grammar (Thomason, 1974), aims at integrating the 'principle of generalized compilation' and linguistic analysis. Natural languages (called *phenotypes*) encompass semiotic variants, which express (in theory) semiotic invariants (operations, categories, relations) that constitute the *genotype* language. The genotype language is a formal applicative language, supposed to be 'universal' on two levels. On the linguistic level, its purpose is to characterize language invariants and to establish universal formulations of grammatical categories; on the mathematical level, it aims at defining ideally the genotype language, which is expressed through natural languages. The genotype language is built by establishing a morphism between the genotype and the various phenotypes. The ultimate goal of the Applicative and Cognitive Grammar is to study the laws that govern all semiotic systems, whether natural (natural languages) or artificial (artificial languages).

It is now commonplace to see artificial intelligence (AI) as encompassing all systems that manipulate symbols¹⁷, and the fundamental tools of AI, such as formalised versions of natural languages, mathematics and modes of logic all fall into the arbitrary symbolic category of sign¹⁸. One of the current theses in AI defines an 'intelligent' computer as one that processes a special type of signs, that is symbols. Humans using symbols in every instance of intelligent behaviour, AI aims at studying (and reproducing) this symbol manipulation¹⁹. Such a theory definitely places AI as a subdiscipline of semiotics, or in other words, AI can be defined as applied semiotics.

"AI studies the functioning of a type of sign called symbols in a constructed or artificial system that is interpretable in cognitive terms." (Meunier, J.G., 1989: 55).

Machine translation has often been rejected by some linguists and translation scholars as unfeasible, due to the fact that translation is very much a creative activity, that cannot be described in a succession of determined operations. On the other hand, if one accepts the assumptions of applicative and cognitive grammar, the source and target languages involved in translation are semiotic variants of the theoretical genotype language. Similarly, different between target texts in a same target language (as produced by different translators) could be interpreted as semiotic variants. The genotype language could be the interlingua language necessary for the next generation of machine translation software. What the interpretive approach to translation calls the 'transfer of meaning' is - in theory - expressible in this genotype language.

The division between computational linguists and translation scholars about machine translation (and the process of translation in general) is rooted in the classical division between semantics (as the science concerned with linguistic meaning) and pragmatics. Machine translation, in its

¹⁷ The three different but related usages of the concept of symbol as described by Eco, 1988, will be here considered: symbols associated with the universe of ideas and thought, symbols identified with functions in non-linguistic communication, and symbols identified with conventional marks in natural or artificial languages.

¹⁸ Hilton, J.1993.

¹⁹ "Trying to understand the nature of symbols and symbolic behavior is an approach to the nature of mind" (Newell, 1986).

present state of development, relies heavily on syntactic parsing and analysis, while human translators use the context, their understanding of the source and target cultural and linguistic systems, as well as their background knowledge to disambiguate the sentences they translate. Moreover, computational linguists deal with artificial languages, while human translators deal with natural languages. Semiotics, as the science concerned with all sign systems, bridges the gap between these two communities, in the sense that semiotics establishes links between all sign systems, whether natural or artificial. Translation (for semioticians) can be carried out not only between two natural languages, but also between a natural language and an artificial one, or between artificial languages. In theory, translation can therefore be viewed as a special case of the automation of human creative activities, and a formal (mathematical) description of the translator's activity can be applied to humans as well as machines. Semiotics does not limit translation to linguistic translation, and therefore allows a larger concept of translation as a transfer process between different systems.

Translation is a highly complex cognitive operation; language itself from which translation gets its *'raison d'être'* is the expression of cognition and communication. Machine translation failed to be the magic wand computer scientists described 40 years ago. Today, more and more researchers and scholars involved in natural language processing seem to reckon that the apparent failure of MT might be due to a wrong hypothesis on which (unfortunately) most translation programmes are based. The redefined aims of machine translation are not to produce a translation similar to a human translation, but to produce a 'good enough' translation, i.e. a text that is understandable, and is reasonably accurate (in terms of a percentage of accuracy, defined beforehand by the user). Computer sciences deal with the (formal) treatment of information, but do not necessarily mimic the functioning of the mind, which deals with the treatment of meaning. Translation is the ideal object of study of the transfer of meaning. A better (even if limited) understanding of meaning analysis and transfer as performed by human translators should give us a new view on the way the mind operates, and on the possible future developments of artificial intelligence, through the use of applicative and cognitive grammar in the particular area of natural language processing.

2 SEMIOTICS, SEMANTICS AND MEANING

2.1 Introduction

The main difficulty faced by translation scholars is that their object of study deals with the interpretation of texts (as messages), themselves encoded in sentences. The view of translation as a communicative process (transfer of messages), or as the equivalence of two messages between two different languages, necessarily leads to the issue of meaning (a 'sense' is produced and perceived as soon as there is communication), hence to the issue of the sign (association of a signified to a signifier).

Meaning, as determined in semantics, is located at the sentence level, and is computed from predictable relations linking each element of a sentence. 'Sense', as the 'entity' extracted and reexpressed by translators, is built on pragmatic²⁰ relations, which are not predictable, since they depend on variable discourse situations.

« Semantics, the tradition holds, is concerned with relations between words and things, while pragmatics is concerned with relations among words, things, and the speakers of a language.(...) Semantics *per se* is concerned with grammatically determined meaning properties of expressions, but pragmatics is concerned with such aspects of meaning as are not determined solely by the grammar. » (Taylor, K. 1998: 82, 84)

The word 'meaning' has many different uses, and raises several problems, since it is not only a word *in* the language, but also a word *about* language (Sless, 1986: 90), and meaning is traditionally seen as *the* theoretical issue in the philosophy of language. There has been a long-standing debate on whether meanings are located in objects, or whether they are products of

²⁰ « Syntax [is] the study of syntactical relations of signs to one another in abstraction from the relations of signs to objects or to interpreters...semantics deals with the relation of signs to designata and so to

communication. Proponents of the first approach argue that meaning is a property of the object. For them, the meaning of the symbol is not dependent on its interpretation. The proponents of the opposite approach argue that meaning is the product or result of communication. For Frege (1892), meanings cannot be in our heads, since if it were the case, there would be no communication. But, he also asserted that they could not be in the world, since we can express thoughts that are totally independent from the material world surrounding us. He therefore concluded that meanings must be abstract objects, neither psychological, nor physical in nature. For Wittgenstein (1953), words do not have meaning *per se*, but can be used to perform some tasks, to express a message in a certain context, and meaning should be understood in terms of usage.

Meaning, as the 'entity' said to be transferred in the process of translation, can neither be determined, nor formalised, without being first defined with the frameworks of semiotics (science studying the production, classification and interpretation of signs as meaning generators), and of semantics (defined broadly as the 'study of meaning', but mainly concerned with word meaning, and more recently the meaning of utterances, or pragmatics). Formal semantics had been articulated for centuries around the issue of meaning and truth, while the relation between meaning and thought was already a topic of controversy in ancient Greece.

2.2 Sign, meaning and semiotics

2.2.1 Introduction

Semiotics is concerned with the study of all sign systems (language²¹, religion, literature, myth...) as *meaning* producers, and with the transmission of meaning across diverse cultures.

objects which they may or do denote... 'pragmatics' is designated the science of the relation of signs to their interpreters. » (Morris, 1971, pp.28, 35, 43)

²¹ Hjelmslev argued that languages cannot be described as pure sign systems: « By the aim usually attributed to them they are first and foremost sign systems; but by their internal structure they are first and foremost something different, namely systems of *figurae* that can be used to construct signs. The

Semiotics is not the science of *one* meaning, on the contrary it attempts to establish a typology of *all* signifying systems, and the rules underlying these systems, through the study of the links between signifiers and signifieds. It differs from semantics, which is concerned with signifieds only. Semiotics studies syntactic, semantic and/or pragmatic properties of the sign. Any semiotic description of a sign therefore includes (Hatim and Mason, 1990: 116):

- the syntactic relations between the sign and other signs belonging to the same syntactic set;
- the semantic relations, between the sign and the entities to which it refers in the real world;
and
- the pragmatic relations, between the sign and its users (senders or receivers).

Semiotics gives a new dimension to the study of meaning (and consequently to translation studies) since it considers the relationships at the heart of the process of meaning²², as well as the essential role of the interpretant in the creation and transfer of meaning. Meaning is first and foremost at the heart of semiotics, since it is the very attribute of the sign, understood as word, phrase, sentence, text, indication, symptom, syndrome, signal or symbol.²³

The issue of the sign (signifier) and its relation to the object or concept it refers to (signified), as well as the association of the theory of language with the theory of signs, date back to the Stoics (3rd century b.c.).

« In the Stoics' semiotics, the theory of language becomes rightfully associated with the theory of signs. In order to have signs, propositions must be formulated, and the propositions must be organized according to a logical syntax which is reflected and made possible by the linguistic syntax. Signs emerge insofar as they are rationally expressible through the elements of language. Language is articulated inasmuch as it expresses meaningful events. » (Eco, 1984 : 32)

definition of a language as a sign system has thus shown itself, on closer analysis, to be unsatisfactory. It concerns only the external functions of a language, its relation to the non-linguistic factors that surround it, but not its proper, internal functions (Hjelmslev, 1943 : 47).

²² "In a general semiotic sense, meaning can be regarded as the relationship between a sign and something outside itself" (Uspenskij, 1977:171).

²³ Sebeok, T.A. 1986.

The logical theory of the syllogism (conditional or disjunctive proposition aiming at building the proof) is articulated around a series of signs without proper meaning, similar to a deductive system based on terms (initial elements) following strict rules. The introduction of the sign theory allowed the proof of the validity of the terms. The terms are true because they are *signs*, i.e. necessary relations between the word and the object it designates, through the *lekton*. The sign is therefore an induction, similar to the syllogistic induction, but instead of being a term wholly located within the linguistic formalism, it links the discourse to the outside world, the words to the objects. 'Translating' or 'interpreting' meant, for the Stoics, finding the sum of a regulated system. Within the Stoic tradition, the sign is unique.

Leibniz²⁴ (1646-1716) deeply modified this Stoic tradition by considering that a sign is fundamentally polysemic:

- each term acquires its meaning according to its usage in each of the various thought domains;
- each term's meaning is modified by its place in the multidimensional network, which covers all the domains.

Semiotics is nowadays articulated around two divergent traditions:

- the European tradition, founded by Ferdinand de Saussure (Swiss linguist, 1857-1913), and its followers (Prieto, Hjelmslev, Barthes). Saussure's theory is also called 'semiology' and refers to 'a science which studies the role of signs as part of social life';
- the American tradition, developed from the works of Charles Sanders Peirce (American philosopher, 1839-1914), and its followers (Morris, Searle, Richards, Sebeok). Peirce's theory is called 'semiotics' and refers to 'the formal doctrine of signs', following Locke's tradition.

These approaches are also sometimes divided between structural semiotics (which analyses sign systems independently of the communication process, disregarding the contexts of production

²⁴ Couturat, L. 1903/1961.

and reception, as well as the interpretation and the use of text), and interpretative semiotics (articulated around the process of unlimited semiosis).

Today, the term ‘semiotics’ encompasses both approaches, as well as the approaches of Roman Jakobson and Umberto Eco, which aim at bridging the gap between the European and the American traditions.

2.2.2 Saussure’s sign theory

Saussure defined a sign as being composed of a *signifier* (or ‘form’) and as a *signified* (or ‘concept’).

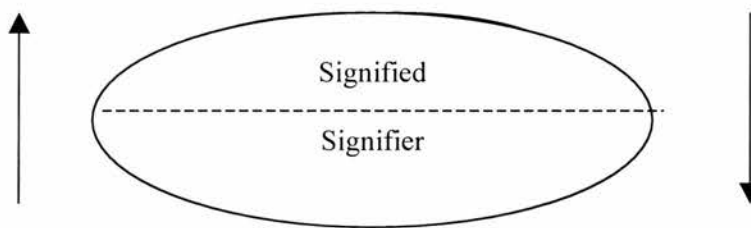


Figure 2-1: Saussure's sign model

Saussure’s signified is not a thing, but rather the ‘notion’ of a thing, the ‘concept’ in the mind (Chandler, 2002: 19). For Saussure, signs only make sense as part of a formal, generalized and abstract system, and there is no direct reference to reality outside the sign. Saussure’s conception of meaning is structural and referential, a sign making sense only in relation to other signs (Chandler, 2002: 22). Signs are essentially arbitrary, and they are determined by their functional differences. The link between the signified and the signifier is arbitrary and conventional, since it depends on conventions imposed by each society and culture.

2.2.3 Peirce's sign theory

Peirce's theory encompasses not only linguistic signs, but also non-linguistic ones. His approach advocates that we start with non-linguistic signs, and then identify the status of language in them. For Peirce, anything can be a sign, as long as it is perceptible, or imaginable. Peirce's signs include physical objects, but more importantly qualities, as well as thoughts, laws, and habits. Peirce saw signs in everything composing the Universe, but he insisted on the necessity for a sign to enter into a relation to its 'object', to be interpreted, and so produce a new sign, its 'interpretant'²⁵.

Peirce's approach is different from Saussure's approach in the sense that it is a triadic model between a *representamen* (the form which the sign takes, and which is not necessarily material), an *interpretant* (the sense made of the sign) and an *object* (to which the sign refers).

"A Sign, or Representamen, is a first which stands in such a genuine triadic relation to a Second, called its Object, as to be capable of determining a Third, called its Interpretant, to assume the same triadic relation to its Object in which it stands itself to the same Object." (*Collected Papers of Charles Sanders Peirce*, 2.274, 1902).

The interpretant in Peirce's theory can be seen as the meaning of the sign (different from Saussure's 'signified'). The interpretant is not an entity but a correlation, a sign interpretative of another sign. The interpretant notion implies that interpretation is a generative process of signification.

« A sign, or representamen, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, create in the mind of that person an equivalent sign, or perhaps a more developed sign. The sign which it creates I call the interpretant of the first sign (*Collected Papers of Charles Sanders Peirce*, 2.228, c. 1897).

The representamen is similar to Saussure's signifier, and the interpretant is similar to Saussure's signified, but the interpretant differs from the signified in the way that it is itself a sign in the mind of the interpreter. Peirce grouped signs into different categories, symbols, icons and

²⁵ Goriée, 1994: 50.

indexes. A *symbol* is a sign, which refers to the object that it denotes by virtue of a law, usually an association of general ideas, which operate to cause the symbol to be interpreted as referring to that object. An *icon* is a sign, which represents its object mainly by its similarity, and the *index* refers to a genuine relation between the sign and the object, which does not depend purely on the interpreting mind.²⁶

The central issue in Peirce's theory is the *interpretant*²⁷, which he classified in three different types: the immediate interpretant (interpretant as it is revealed in the right understanding of the sign itself, and is ordinarily called the 'meaning of the sign'), the dynamical interpretant (the actual effect which the sign, as a sign, really determines), and the final interpretant (the manner in which the sign tends to represent itself to be related to its object²⁸, each of them being determined by the previous interpretant.

2.2.4 Meaning and semiotics²⁹

In Saussure's approach, the meaning of a sign is strictly bound by convention, and it arises from the syntagmatic and paradigmatic differences between signifiers. Syntagmatic relations are possibilities of combination, and they refer to other signifiers co-present in the text. Paradigmatic relations involve differentiation, by referring to signifiers, which are not present in the text.

²⁶ Chandler, 2002: 38-41.

²⁷ Goriée, 1994: 56-57.

²⁸ Peirce, 1906.

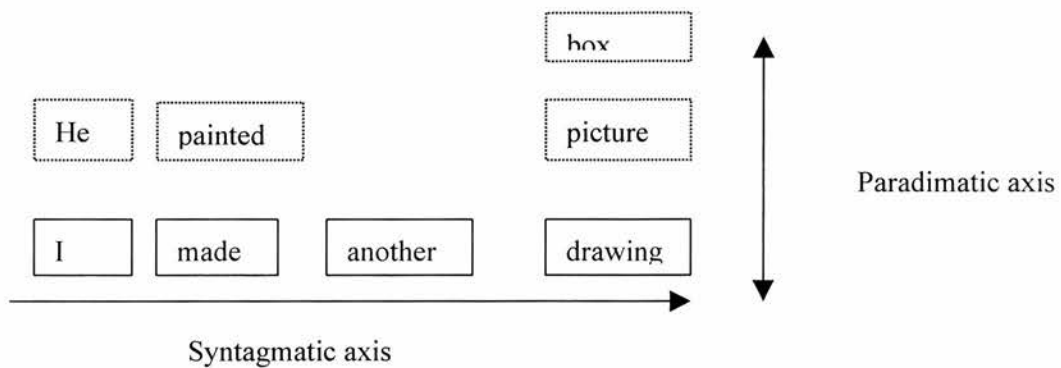


Figure 2-2: Syntagmatic and paradigmatic differentiation

For Saussure, meaning arises from the differences (syntagmatic, concerning positioning; and paradigmatic, concerning substitution) between signifiers. Syntagms and paradigms³⁰ provide a structural context within which a sign makes sense, and both the paradigmatic and the syntagmatic relations determine the ‘value’ of such a sign. The choice of one signifier (e.g. a particular word) rather than another from the same paradigm set (e.g. adjectives) shapes the preferred meaning of a text.³¹

The meaning of a sign is for Peirce abstract and virtual; it consists of a certain inward activity performed by an abstract interpreter and composed of mental operations, and for this reason it is pragmatic, or rather operational; it is of the character of a sign process: it develops gradually and becomes ever more complete.³²

Within Peirce's approach, meaning is a process, rather than an entity, and only the study of the process can give us any valuable insight into meaning. The meaning of any linguistic sign is defined as its translation into some further, alternative sign. This fundamental principle led Jakobson (1959) to define translation in terms of interpretation:

²⁹ "'Sign' and 'meaning' are inextricable; to identify something as a sign is in the next breath to interrogate its meaning, for it is in the nature of signs (or so it would seem) to have meaning." (Sless, 1986: 88)

³⁰ "A paradigm is a set of associated signifiers or signifieds which are all members of some defining category, but in which each is significantly different. Verbs and nouns are grammatical paradigms." (Chandler, 2002: 80).

³¹ Chandler, 2002: 80.

³² Sebeok, T.A.1986.

- intralingual translation is the interpretation of verbal signs by means of other signs of the same language;
- interlingual translation is the interpretation of verbal signs by means of some other language;
- intersemiotic translation is the interpretation of verbal signs by means of signs of nonverbal sign systems.

In that respect, translation as meaning-generator cannot be fully grasped outside semiotics. As Peirce expressed it, the meaning of any sign lies in its translatability, and sign translation equals sign comprehension (or interpretation).

2.2.5 Interpretation and 'unlimited semiosis'

By interpretation, Eco (1984) means the concept elaborated by Pierce, according to which every interpretant (either a sign, or the content of the sign), also increases our understanding of it. The sign, by being a correlation (an 'association' between the signifier and the signified) is potentially greater than the sum of its parts. The interpretant can therefore assume even-wider meaning. Peirce called this sign interpretation 'semiosis', i.e. the triadic process between sign, object, and interpretant.

Even though for some philosophers, meaning is considered as a property of the object, we will follow here Peirce's (as well as Wittgenstein's, 1953, and Ricoeur's, 1976) views, according to which meaning is constantly created through the process of interpretation. For D.Sless (1986), some philosophers view meaning as a property of the objects, because meanings are somehow objectified (our knowledge and expectation being externalised and imposed on the objects around us) and therefore fixed as the common property of the group. But traditionally, the interpreting view is the one that dominated philosophy and semiotics. The meaning of a sign (for Peirce) is its translation into another equivalent or perhaps more developed sign.

"Meaning can best be formulated as consisting of a triadic relation of interpretation, whereby one term affects a mind in such a way that it is taken to stand in some respect for another term. Meaningful is something through which something else can be known (*aliquid stat pro aliquo*)."
(Sebeok, T.A., 1986: 510).

Peirce's pragmatism (that he rephrased later as "pragmaticism", to avoid any confusion with William James' "pragmatism"³³) consists in "a method for ascertaining the real meaning of any concept, doctrine, proposition, word, or other sign"³⁴.

This meaning-generation process, result of the interpretation process (each interpretant using the code to 'extract' the message) accounts for the variety witnessed in translations. It is common knowledge that two translators hardly ever produce the same TT in a TL from the same ST. This observation confirms Peirce's focus on the *interpretant* as a sign triggered in the mind by another sign.

"A sign, or *representamen*, is something which stands for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. The sign which it creates I call the *interpretant* of the first sign. The sign stands for something, its *object*. It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the ground of the representamen." (Collected Papers of Charles Sanders Peirce, 2.228, 1897)

Barthes (1983) developed a theory of second-order systems accounting for the different meanings one can determine (denotative and connotative). The denotative meaning is created by the association of the signifier and the signified. This process creates a new sign, which acquires additional meaning (connotative). Additional connotative values are acquired through this process, that can be renewed more or less indefinitely³⁵.

³³ Williams James (1842-1910) was a friend of Peirce, and his doctrine was more general than Peirce's.

³⁴ Collected Papers of Charles Sanders Peirce, 1905: 5-6.

³⁵ Hatim & Mason, 1990:112.

2.2.6 Sign and text

Y. Lotman (1977a, 1990) and P. Ricoeur (1976), amongst others, showed that human translation - as transfer of messages - entails an accretion of meaning (a "surplus of meaning") since there is no absolute coincidence of codes between the sender and the receiver of a message. The sender encodes the message, the receiver decodes it, but their codes - even though they overlap each other - are not exactly identical, hence producing different target texts.

In contradistinction, the transmission of a text through an artificial language entails a total adequacy between the transmitted and the received message. In such a case, all cultural context is removed, since an artificial language models not language as such, but only one of its functions, that is the ability to transmit a message adequately. On the other hand, a system which fulfils the entire range of semiotic possibilities transmits messages, but also serves as a generator of new messages.

T = text; C= code

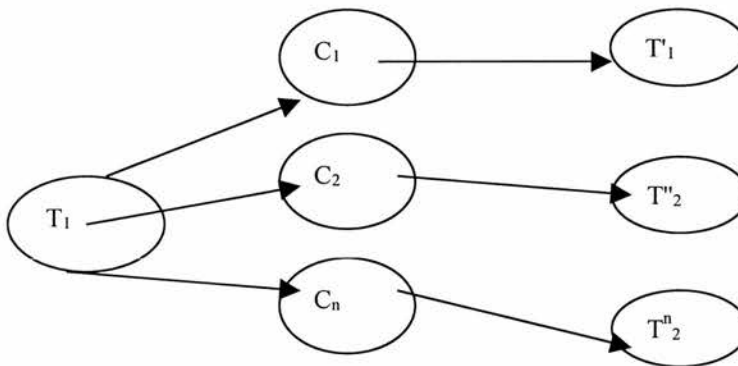


Figure 2-3: Artistic translation (from Lotman, 1990: 15)

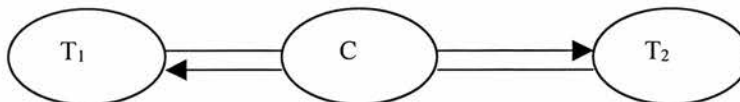


Figure 2-4: Transmission of a text using an artificial language (from Lotman, 1990: 14)

Part of the misunderstanding between translation scholars on the one hand, and computational linguists on the other hand, is due to this different approach to language. Whether one considers the informational point of view ('informational' being taken in the narrow sense), or the creative point of view, language is represented as a machine for transmitting invariant messages (the code being unique), or as a meaning-generator.

For Eco (1984: 24), texts generate, or are capable of generating, multiple (and ultimately infinite) readings and interpretations.

The theory of sign as meaning-bearer and meaning-generator was given a new light by Lotman who attempted to describe the functioning of human consciousness³⁶ based on sign analysis at two different levels, on the sign level, or on the text level:

- at the sign level, the basic bearer of meaning is the sign, the meaning of the text comes second, and it is inferred from the meaning of the signs;
- at the text level, the basic bearer of meaning is the text.

For Barthes (1957), Derrida (1980/1985) and Kristeva (1981), signification is to be located exclusively in the text. Meaning is produced and becomes productive only in texts.

Texts do not only generate new meanings, they are also condensers of memory. The text's memory is defined by Lotman (1990) as the sum of the contexts in which a given text acquires interpretation and which are in a way incorporated in it. The text then acquires semiotic life through the relationship created between this text meaning-space and the cultural memory (tradition) already formed in the consciousness of the audience/readership. The sum of a text's memory is what machine translation try to compute and integrate into their software.

³⁶ "Within one consciousness there are as it were two consciousnesses. The one operates as a discrete system of coding, forms texts, which come together like linear chains of linked segments. In this system, the basic bearer of meaning is the segment (the sign), while the chain of segments (the text) is secondary, its meaning being derived from the meaning of the signs. In the second system, the text is primary, being the bearer of the basic meaning. This text is not discrete but continuous. Its meaning is organized neither

2.3 Meaning and semantics

2.3.1 Introduction

Meaning has traditionally been the object of study of semantics, and has been the concern of philosophers. In traditional philosophy, meaning can be viewed as reference, as logical form, as context and use, as culture, or as conceptual structure. 'Meaning', or 'sense', as understood in English may be translated by 'signification', or 'sens' in French, 'Bedeutung' or 'Sinn' in German, even though the respective differences between the concepts do not correspond from one language to another. Depending on the context, 'meaning' will be either 'sens', or 'signification' (respectively 'Sinn', or 'Bedeutung' – even though 'sens' in French is not necessarily 'Sinn' in German!).

There is no unique definition of meaning, since one should mention several meanings rather than one single meaning, depending on whether one considers a word, a sentence, or a text. Lyons (1981a) lists six different kinds of meaning: the lexical meaning (or word meaning, which is more precisely the object of study of lexicology), the sentence meaning (whose meaning is said to depend upon the meaning of its constituent lexemes, while at the same time the meaning of some, if not all, lexemes, depends upon the meaning of the sentences in which they occur), the grammatical meaning (which is indicated by the grammatical structure of sentences)³⁷, the utterance meaning (which does not really belong to linguistic semantics, but rather to pragmatics), the descriptive meaning of statements (which can be identified with the

in a linear nor in a temporal sequence, but is 'washed-over' the n-dimension semantic space of the given text. In texts of this type, the text is the bearer of the meaning." (Lotman, 1990: 36).

³⁷ Grammatical meaning is the core element in machine translation, while human translation deals more with pragmatics.

proposition that is asserted in statements), and the social meaning (which has to do with the use of language to establish and maintain social roles and social relations).

Propositions, as defined by logic, have a definite truth-value (they are either true or false). This connection between truth and descriptive meaning is at the heart of truth-conditional semantics. Semantics is precisely concerned with the study of linguistic meaning.

2.3.2 Meaning and reference

“Meanings are ideas or concepts which can be transferred from the mind of the speaker to the mind of the hearer by embodying them, as it were, in the forms of one language or another.”³⁸

Such a definition, as shown by Lyons, is raising more problems than providing answers to the question “what is meaning?”, since the term “concept” in itself is very vague. If the meaning of a term is a concept, then meaning is a kind of mental entity. Such a notion was opposed by Frege³⁹ who argued that meanings are public properties. He then identified concepts with abstract entities rather than mental entities. He suggested a distinction between two kinds of meaning, the *reference* and the *sense*. The *reference* is the actual object or real world entity picked out by a linguistic expression. The *sense* is the meaning of an expression, by virtue of how it is said, or the form of the proposition. Reference is the process by which a signifier relates to a signified. According to Frege, a name designates or denotes its reference and expresses its sense. A sense is said in turn to determine a reference. A sense is ‘a way of being given a reference’ or a ‘mode of presentation of a reference’.

³⁸ Lyons, 1981a: 136.

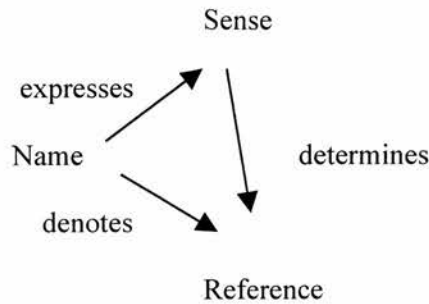


Figure 2-5: Sense and reference

The object itself that is referred to is the extension, the intension being all the information other than the extension. There is variation in meaning despite constancy of reference, because there is a split between extension and intension, with the latter fixing the former. Meaning variation, despite a constant extension, is possible because meaning itself is the selection of properties, or pieces, of a projected world (Frawley, 1992).

2.3.3 Truth and meaning

Meaning can be viewed either as a relation between language and the world (in terms of truth), or as a relation between language and its users (in terms of understanding). A distinction must be made between the truth-value of a proposition and the truth-conditions of a sentence. The truth or falsity of a proposition is its truth-value. On the other hand, most sentences do not have a truth-value, but they may have truth-conditions, i.e. a precisely specifiable account of the conditions which determine the truth-value of the propositions conveyed by sentences when they are used to make statements⁴⁰. The ‘meaning’ of a sentence or an expression is the conditions under which its saying is true, but by reference to a piece of a mentally projected

³⁹ He was puzzled, for example, by statements such as ‘the morning star’ and the ‘evening star’, which refer to the same entity, even though in logic they would be two statements differing in cognitive content $([a=b]/[a=a])$.

⁴⁰ Lyons, 1981a: 171.

world. The content of an expression is determined by differentiating its status as statement (a form conveying true or false information), from its status as sentence (a form with grammatical structure).

2.3.4 Formal semantics

Formal semantics was originally concerned with the semantic analysis of formal systems (formal languages artificially constructed by logicians and computer scientists) and is not committed to any particular theory of syntax, hence its broad range of applications. The term is now also used for the semantic analysis of natural languages, with some restrictions though, since in formal semantics sentence meaning is quite often reduced to propositional content⁴¹. Many linguists have therefore challenged formal semantics, on the grounds that the propositional content of a sentence is insufficient to account for all the different kinds of meanings that can be attributed to sentences. On the other hand, some of the tenets of formal semantics (like the compositionality principle) seem to be confirmed by the results of studies on natural languages. The 'compositionality principle' was originally developed by G. Frege (1892, 1953, 1971), and states that the meaning of a composite expression is a function of the meanings of its component expressions and of the way they are syntactically combined. For example, a competent native speaker is able to interpret indefinitely many composite expressions of that language, although no one actually learns the sense of every composite expression. One must therefore assume that a native speaker is able to determine the sense of composite expressions on the basis of the sense of the lexemes of these expressions. This determination is said to be possible because the sense of the composite expression is a function⁴² of the senses of the lexemes. Moreover, semanticists assume that the sense of such a composite expression is also a function of its grammatical structure, whatever particular theory

⁴¹ In the propositional calculus, expressions are mapped onto the domain {True, False}.

⁴² A 'function' (used in its mathematical sense) refers to a rule formula, or operation, which assigns a single value to each member of the set of entities in its domain. It thus establishes either a many-one or one-to-one correspondence between the members of the domain, D, and the set of values, V. It maps D either into or on to V. (Lyons, 1981b:145)

of syntax is selected as the set of rules. Formal semantics is concerned with the interpretation process between syntax and semantics.

2.3.5 Semantics and grammar

One important issue relevant to translation studies, and more particularly to machine translation is the traditional separation between semantics on the one hand and syntax on the other. According to A. Wierzbicka (1988), the practical division between lexical semantics, grammatical semantics, and illocutionary semantics (supported by Morris' approach to signs - semantics, syntax and pragmatics), makes sense when one considers artificial languages, but is confusing and misleading, when one considers natural languages. In the case of natural languages, syntactic, morphological and illocutionary devices are all carriers of meaning, and cannot be considered separately.

Semantics has long been imprinted by the Katz-Fodor⁴³ theory within the framework of Chomskyan generative grammar. In this approach, every sentence has two distinct levels of syntactic structure, the deep structure, and the surface structure, generated by rules of a different kind. Sentence-meaning is supposed to be connected with the deep structure. And two sentences that have the same deep structure are assumed to have the same meaning. In the Katz-Fodor theory, projection rules (rules of the semantic component) aim at distinguishing meaningful from meaningless sentences, while at the same time assigning a semantic representation of every semantically well-formed sentence. Logicians and philosophers have challenged this concept of semantic representation, on the grounds that this semantic representation actually uses a formal language, whose vocabulary units need interpretation.

Moreover, variations in surface structures were seen as irrelevant, as long as these surface structures had the same deep structures. The works of Wierzbicka (1988), Bolinger (1965) and Langacker (1987, 1991a, 1991b), amongst others, aim at demonstrating that these surface differences signal differences in meaning, linked to the human interpretation of the world.

⁴³ Katz, J.J., Fodor, J.A., 1963.

"Every grammatical construction encodes a certain meaning, which can be revealed and rigorously stated, so that the meanings of different constructions can be compared in a precise and illuminating fashion, both within one language and across language boundaries." (Wierzbicka, 1988: 3)

The differences in the meanings encoded in the languages can be (in theory) explicitly determined, and allow for the establishment of 'universal' meanings, as opposed to those meanings which are restricted to one particular language. In reference to the applicative and cognitive grammar formalism, languages as phenotypes present discrepancies in the way they express meaning. Whether it is possible to establish a one-to-one relationship between grammatical constructions (an 'equivalence' between structures), or whether some constructions are unique to some languages remains to be demonstrated. In this respect, the analysis of human translation can also shed new light on the convergence/differences of meaning expressed through different languages.

2.4 Conclusion: semiotics, semantics, language and cognition

Human languages are sign systems, and at the same time means of communication, and linguistic signs differ in some respects from other signs, in the sense that they belong to conceptual intelligence, which seems to be uniquely human. This difference led some linguists and philosophers to reject the description of languages as pure sign systems. For Hjelmslev (1943: 47):

"By the aim usually attributed to them, they are first and foremost sign systems; but by their internal structure, they are first and foremost something different, namely, systems of *figurae* that can be used to construct signs. The definition of a language as a sign system has thus shown itself, on closer analysis, to be unsatisfactory. It concerns only the external functions of a language, its relation to the nonlinguistic factors that surround it, but not its proper, internal functions."

For some other cognition scientists and semioticians, languages are first and foremost sign systems, and an 'intelligent' computer is - in theory - able to process natural languages, since it is devised to process a specific type of signs, symbols. According to Meunier (1989), the term 'process' is to be understood not in a physical or physiological sense, but rather as a *functional* process, in a way similar to the semiosis theory, as sign generator.

The diverging views on language and artificial intelligence can be reconciled if one considers natural language as a sign system, but takes into consideration Peirce's approach to sign as meaning-generator. Natural languages are bearers and creators of meanings. The analysis of the process through which meaning is encompassed in a text, simultaneously by the lexicon and the grammar, as external expressions of internal cognitive processes present among and across natural languages is the new challenge of semantics and cognitive sciences. Translation provides the material necessary for such a research. A comparative analysis of parallel texts (source text, several target texts in several target languages) provides the basis for a better understanding of how meaning is expressed in a source text and 'transferred' to a target text, said to be in an 'equivalence' relationship.

3 TRANSLATION THEORY AND MEANING IN TRANSLATION

"Le langage est source de malentendus." (*Le Petit Prince*, A. de Saint-Exupéry)

3.1 Introduction

3.1.1 Definitions of translation

In all of the major European languages, the verbs describing the process of translation point etymologically to a crossing over from source language to target language. In Latin ('*transfere*') and English ('*translate*') it is a carrying over, in the romance languages, it is a leading over, and in the Germanic languages it is a setting over ('*übersetzen*').

First introduced by R. Estienne in 1539, '*traduire*' is defined in the Robert Dictionary (1985) as 'to express in a different language what has been said in a source language, keeping a semantic and expressive equivalence between the two messages'.

"A translation should not only be linguistically correct and make sense with regard to any particular subject area. It has to stand in some kind of equivalence relation to the original.(...)
Equivalence in translation is not an isolated, quasi-objective quality, it is a **functional** concept that can be attributed to a particular **translational situation**." (Neubert, 1994: 413).

Most scholars concerned with translation developed theories based on one or the other feature of the operation depending on their approach, and some of the early theories in the 1960s were obviously tainted with the purely linguistic imprint.

"Translation is an operation performed on languages: a process of substituting a text in one language for a text in another. Clearly, then, any theory of translation must draw upon a theory of language - a general linguistic theory." (Catford, 1965: 1)

G. Mounin⁴⁴ sees translation as a contact of languages, an expression of bilingualism and concludes that languages are – rightly – not translatable, but fails to notice that translation is – first and foremost – concerned with texts, not languages. Translation studies slowly freed themselves from the purely linguistic approach⁴⁵ to take in the multidimensional process at stake (cognitive, communicative, text-linguistic, functionalist) and became an autonomous field of research at the intersection of linguistics, comparative literary studies, psychology, sociology, logic, and semiotics. Translation is seen today as a complex synthetic process in which the translator dynamically matches semantic, syntactic, textual, and pragmatic fields to create a unitary whole, the target text. Being an intersection of situation, translator competence, source text and target text-to-be, translation is variable⁴⁶. But one of its main characteristics is that it deals with operations **on messages, through languages** (not **on** languages). These operations are determined by several parameters, such as intentionality, situationality, coherence (sense constancy), cohesion⁴⁷ (in terms of lexis and grammar, the surface components depend upon each other in establishing and maintaining text continuity), and intertextuality.

3.1.2 Approaches to translation

Recent research in linguistics, psychology, logic and philosophy, as well as the development of the Internet and the increasing needs for multilingual translation marked a renewal of interest in translation. Translation is now acknowledged as a worldwide means of communication, and researchers tend to focus their studies more on the processes involved in translation than on the debate about un/translatability. The term *translatibility* is used to discuss the extent to which it is possible to translate either individual words and phrases or entire texts from one language to

⁴⁴ Mounin, G. 1963.

⁴⁵ "Linguistics alone won't help us. First, because translating is not merely and not even primarily a linguistic process. Secondly, because linguistics has not yet formulated the right questions to tackle our problems. So let's look somewhere else." (Vermeer, 1987: 29).

⁴⁶ Neubert and Shreve, 1992.

⁴⁷ For De Beaugrande and Dressler (1981), cohesion is the purely formal connexity of a text, which materializes in the occurrence of particular lexical-grammatical means.

another, considering the unique configurations of grammar, vocabulary and metaphor in each language. The untranslability discussion usually puts forward one of the following arguments:

- the Sapir-Whorf hypothesis (Whorf, 1941), according to which different realities engender different languages and different languages engender different realities;
- the Quine's indeterminacy hypothesis (1960), according to which the subjectivity of both author and translator severely limits an accurate transfer of meaning;
- the literary view (Nabokov, 1941), according to which a text is seen as a 'gestalt' embedded in a cultural or historic matrix. Translatability is therefore limited by the coherence between the linguistic elements of the source text and its cultural matrix.

Many models and theories have been developed in order to analyse and define the intellectual processes involved in translation, which seems to be an activity as old as language itself. Although there is no proof of the earliest stages of translation, one may assume that the first human tribes getting into contact with each other communicated through some form of translation. Interpreters existed in ancient Egypt, and the importance of their rank was acknowledged by their title ('Prince') transmitted from father to son. Written translation also existed then, as proved by the discovery of texts of treaties between Hittites and Egyptians in two different languages, and confirmed by the inscriptions on the Rosetta stone (hieroglyphic, demotic Egyptian, Greek). Nowadays, hardly anybody would question the possibility of translation, which is essentially a pragmatic operation, but because of this very characteristic, translation resists all attempts at unified theorisation. Two main research directions have often been opposed in translation studies: one emphasizing the communicative features of translation as a form of language in use; the other one focusing on the systemic nature of the linguistic relationships that exist in translation.

Different models (conceptual constructs) of translation coexist, although they (individually) fail to account for all the processes involved in translation. Scholars (usually trained in another field – linguistics, philosophy, literature...) can easily argue over the predominance of one or the other of the following models:

- the linguistic approach, which focuses upon a comparison of the two languages – case relations, relative order of words, phrases and clauses, reference markers;
- the philological approach, which focuses upon the comparison of the source and target texts through the study of the register of language, the extent of figurative expressions, the shifts in evident intention and the phonological features;
- the text-linguistic approach, which maintains that an original text and its translation are different for two reasons ; first because the sentences of the ST and TT are determined by the linguistic rules of two different systems; secondly, because translation is reconstructed as a new semantic and pragmatic totality (i.e. the target text) in the target language community;
- the communicative approach, which considers the communicative event in the source language and the corresponding communicative event in the target language;
- the socio-cultural approach, which is concerned with the cross-cultural communication involved in the process of translation;
- the psycho-linguistic approach, which focuses upon the cognitive factors and language processing strategies;
- the functional approach⁴⁸, which focuses on the intended function of the target text or any of its parts;
- or the computational one, which is concerned with the analysis of the meaning components that are packed together into words and systems of words in the lexicon of the language.

⁴⁸ "Translation is the production of a functional target text maintaining a relationship with a given source text that is specified according to the intended or demanded function of the target text (translation skopos, or purpose)." (Nord, C. 1991: 28).

This rather confusing plethora of approaches stresses the complexity of such a ‘natural’ activity performed by human beings. Most early computational linguists seem to have overlooked the complexity of translation, as well as its cognitive dimension, and they blindly applied the results of research on language (and thus from a non-translation perspective) on the process of translation. The results of the first translation software were, not surprisingly for translators, very disappointing. Translation expresses intertwined cognitive, psychological and socio-linguistic operations, which have – so far – resisted modelisation.

3.1.3 Brief historical survey of translation theories

In the XVIIth century, John Dryden put forward three main different approaches to translation, the word-for-word technique (verbal version), the paraphrase technique, following Cicero’s principle of meaning equivalence (‘non verbum de verbo, sed sensum exprimere de sensu’), and the imitation technique, or free translation. He thought that imitation and verbal version should be avoided, and he was already setting the roots of the meaning issue in translation, by being in favour of the second approach. The first theoretical essay on translation in English (“Essay on the Principles of Translation”), published by Alexander Frasey Tytler in 1791, was articulated around three principles:

- the style of the target text ought to reproduce completely the idea of the original;
- the style of the target text should be as close as possible to the style of the source text;
and
- the target text should be read as the source text.

The XVIIIth century was marked by the “Belles Infidèles” trend in France, and in the XIXth century, the notion of fidelity became a dogma, sustaining a literal approach to translation. The first linguistic theories of translation were published in the middle of the XXth century, laying

the grounds for a systematisation of the translation operations, and for translation studies as an autonomous field of research.

The development of translation studies in the last century is linked to the development of literary, linguistic and cognitive research. During the period stretching from 1900 to 1930, translation was mainly seen as an interpretation, which necessarily transforms the foreign text. During the following thirty years, the main approach to translation was a linguistic one, and the main object of study was the un/translatability issue. The next decade (1960-70s) was dominated by the concept of equivalence⁴⁹, and translating was seen as a process of communicating the foreign text by establishing a relationship between the source and the target texts, the target text being embedded in a target linguistic and textual system as much as in a target culture. In the 1980s, translation studies emerged as a separate discipline, overlapping with linguistics, literary criticism, and philosophy. The decade was dominated by the development of the Skopos theory⁵⁰ (Vermeer, 1978), and of the polysystem theory⁵¹ (Even-Zohar, 1978; 1990). The end of the last century was marked by the expansion of the theories previously developed, as well as by the application of research led in linguistics (pragmatics⁵², critical discourse analysis, computerized corpora) and in literary and cultural theory. More attention was given to the study of the 'mental processes', in particular through the "think-aloud protocols (TAP)", in order to discover and explicitate the thought processes taking place when someone is translating a text (the 'black box' of translation), translators being asked to verbalize their thinking during or immediately after the translation process (Lörscher, 1991b). The purpose of the TAPs was to gain a better understanding of the psychological and linguistic mechanisms involved in the activity of translating, but the usefulness of TAPs was limited by the fact the subjects can verbalize only that which is conscious. The results of the experiments were therefore necessarily limited, and could provide only an incomplete account of the

⁴⁹ In 1963, G. Mounin argues that equivalence is based on "universals" of language and culture.

⁵⁰ The skopos theory is oriented towards the function which the target text is to perform in the target culture for the target readers.

⁵¹ The term polysystem refers to the aggregate of literary forms that exist in any given culture.

⁵² Pragmatics based theories assume a communicative intention and a relation of equivalence, based on textual analysis.

processing involved in translation, considered as a cognitive task. Moreover, there is a controversy about the results of such inferences, on the grounds that the data might give access not to the actual mental processes, but rather to the intermediate products of these processes (sometimes called mental contents).

3.2 Issues raised by divergent approaches to translation

3.2.1 Equivalence in translation

"Equivalence is usually defined as the relationship between a source text (ST) and a target text (TT) that allows the TT to be considered as a translation of the ST in the first place. But [...] this definition is problematic because of its circularity: equivalence is supposed to define translation, and translation in turn defines equivalence." (Baker, 2001: 77).

"Equivalence is the aim of translation in that translation is seen as striving towards equivalence, or at least the particular kind of equivalence which suits the occasion. At the same time, equivalence is the precondition of translation in that only a target text which displays the required amount of equivalence, of the right kind, is recognized as a valid translation." (Hermans, 1999: 48)

Translation operations are performed on **texts** belonging to different linguistic systems, with the purpose of producing an '**equivalent**' text or **message** in a different language, but this target text is supposed to trigger a similar reaction on a readership whose language and culture are different. Translation is the expression of a relation⁵³ of **equivalence** between the source text and the target text (text being taken here as any statement, whether a word, a sentence or a whole paragraph). But this equivalence depends necessarily on the nature of the texts, on their target audience/readership, on the links between the two cultures involved, on the circumstances surrounding the production of the source and target texts. Some linguistic correspondences will necessarily appear (even more so, when the two languages considered are morphologically and

⁵³ This relation is supposed to be determined by some 'coding rules'.

lexically close), but linguistic correspondences cannot account for semantic and stylistic equivalence.

The strongest view of the equivalence principle is the 'invariance model' of translation as expressed by Toury (1988: 24): "Translation potential may thus be rewritten as the capacity of substituting TTs for STs under some invariance conditions." Toury considers that translation deals with replacement (substitution) from a source to a target system, being first and foremost inter-textual (involving pairs of texts), and is thus interlingual (involving pairs of languages) only by implication. This process is irreversible, and the invariant relationships are unidirectional. Within these text-oriented theories based on the equivalence principle, the invariant features are established on the level of textual functions (Catford, 1965 ; Even-Zohar, 1971), and tend to be source-oriented.

This notion of equivalence is a central issue and an object of disagreement between translation theorists. It has been interpreted in very different ways and is one of the most ambiguous concepts in translation studies⁵⁴. Many models are based on this 'equivalence principle', according to which the TT must convey the message of the ST while triggering a similar reaction on the target readership. On the one hand, some linguists consider, that there is no possible equivalence in translation, each linguistic system (language) being characterized by a specific analysis of the outer world (analysis which is different from other languages' analysis) and different linguistic structures defining 'different' worlds in different languages (Sapir/Whorf hypothesis). For other linguists, who still acknowledge linguistic differences, equivalence is possible, but it remains to be defined.

⁵⁴ "On the level of interlingual translation, there is ordinarily no full equivalence between code-units, while messages may serve as adequate interpretations of alien code-units or messages.(...) Translation involves two equivalent messages in two different codes. Equivalence in difference is the cardinal point of language and the pivotal concern of linguistics." (Jakobson, 1959: 233).

One possible approach to this 'equivalence in difference' consists in considering the Saussurian opposition 'language/parole', as well as the opposition between linguistics and translation studies. Translation is concerned with the transfer of one message belonging to one linguistic (but also socio-cultural) system into another message belonging to another linguistic system. Comparative studies, centred mainly on a search for similar or equivalent sentence structures in different languages, fail to account for the true dimension of translation, which deals with the interpretation of actualised/contextualised messages. In that respect, despite its interest and use, linguistics alone is insufficient to build a theory of translation. Moreover, the purely linguistic approach associated with the notion of equivalence led to the development of translation theories based not on equivalence, but rather on linguistic correspondences. The abundance of 'translation methodologies' shows, if need be, how attractive such a view might be, but does not prove that these works actually deal with translation (a contextually defined communicative process). They aim at cataloguing linguistic correspondences, and might be useful in teaching foreign languages, but the blind belief of too many linguists in correspondences between languages out of context probably contributed to the early failures of machine translation. Apparent linguistic correspondences (linguistic meaning as opposed to sense) do not account for the highly complex cognitive and sociolinguistic operations at stake in translation.

Nida (1982) defined five fields involved in the transfer from one cultural world to another, and essential to the principle of equivalence: ecology, technology, sociology, religion and language. They express differences in cultures (is there anything like a 'desert' for the people of Amazonia?) which must be taken into account in the attempt to render an 'equivalent message'. From a linguistic viewpoint, structures may be different in languages and the extent of vocabulary may vary considerably from one language (and culture) to the other. For example, the French '*tu*' and '*vous*' which correspond to one form 'you' in modern English, or the dozens of words used by Norwegian peoples to describe 'ice', that have no exact correspondences in most other languages (in terms of single words). Nevertheless, a good translator will be able to express the content of the message, by using all means of the target language and culture at his disposal. A different syntactic structure in the TT may render a word in the ST, or a linguistic

expression in the TL render a syntactic structure in the ST. Subtle cultural hints in the ST may be omitted if the target language does not have the same patterns (for example, the difference between uncle from the mother's side and uncle from the father's side – which is not necessarily marked in all languages) without much loss for the reader of the translation.

The equivalence paradigm has been challenged by some scholars, for example Robinson (1991), who sees equivalence as **only one** (among many) goals of the translator's activity. He questions Nida's theory of dynamic equivalence⁵⁵, as well as the view of structural equivalence (mainly supported by logical theorists of translation), and advocates a **trope** of translation, i.e. an 'interpretive tool that encourages us to **think** of two texts (or two images) **in terms of** equivalence'. According to Robinson, the 'equivalence principle' misled many translation scholars. For him, there is no systematisable structure of correspondence between two texts in the abstract, but rather an idiosomatic (individual/unique) response. Still, this idiosomatic response being largely determined by our culture (or 'society's dominant ideology'⁵⁶, in Robinson's words), the 'equivalence' sought by translators is determined by their shared cultural and ideological patterns. Although Robinson reckons that a 'structure' of equivalence does exist (equivalence being governed by ideological rules programmed into us by our culture), he claims that it rests on two wrong assumptions: first, that the structure of equivalence is natural and universal; secondly, that it is inevitable and immutable. He calls equivalence an 'interpretive fiction' which is only a tool for the translator.

For Schogt (1988), equivalence should be a goal, although some translations (very literal and very free ones) obviously do not follow that principle. For Barnstone (1994: 99), full equivalence is impossible, and relative equivalence remains an ideal, but not necessarily a goal.

⁵⁵ Nida (1964) distinguishes between 'formal equivalence' (aiming at achieving equivalence of form between ST and TT) and dynamic equivalence (aiming at achieving equivalence of effect on the target language reader).

⁵⁶ Ideology being defined as the tacit assumptions, beliefs and value systems, which are shared collectively by social groups.

He suggests the following view: although the equivalence of two texts should be one of the tenets of translation (transfer of the maximal semantic load), real equivalence may be a secondary goal in literal translation, and might even be totally irrelevant in the case of free translation. Lawendowski (1978) warns us against the risk of turning the equivalence principle into a dogma, leading to a hunt for correspondences. Holmes (1988: 53) also denounces the term 'equivalence', on the grounds that "the languages and cultures to be bridged, however close they may sometimes seem, are too far apart and too disparately structured for true equivalence to be possible", and prefers 'counterparts' or 'matching'.

3.2.2 Unit of translation

"The unit of translation is the stretch of source text on which the translator focuses attention in order to represent it as a whole in the target language." (Lörscher, 1993: 209)

Translation scholars usually disagree about the range and scope of the unit of translation (UT). Catford (1965: 8) considers 5 levels of hierarchy: the sentence, the clause, the group, the word, and the morpheme. Newmark (1988: 54) ranks ranges from the complete text, to the paragraph, sentence, clause, word group, word, morpheme. Even though he acknowledges the necessity to take into account the full text, he discards the text as a possible UT ("Ideally, the UT is one word (...) never the text.")

The term 'unit of translation' is defined, in the Dictionary of Translation Studies (1997), as 'the linguistic level at which the source text is recodified in the target language'. Considering the controversies about the linguistic approach to translation and about the decoding-encoding frame, the definition given by Lörscher (1993: 209) would be closer to the actual entity grasped by translators in their daily activity: 'the stretch of source text on which the translator focuses attention in order to represent it as a whole in the target language'. For Vinay and Darbelnet (1958: 16, 37), the unit of translation is " the smallest segment of utterance where the cohesion of signs is such that they cannot be translated separately", or the "lexicological unit where lexical elements converge in the expression of a single element of thought."

The evolution of translation theories led to opposing views on the segmentation of the source text and on the delimitation of "units of translation". The unit of translation is also defined sometimes as a translation *atom*, expressing the fact that it is *the smallest segment* that must be translated as a whole. Several studies⁵⁷ have shown that most experienced translators tend to translate units of "meaning", usually at the level of clauses or sentences, more than at a word-level. Recent theories emphasize the necessity for the translator to consider the largest linguistic unit (the text), while computational approaches to translation insist on keeping the unit of translation as small as possible, on the grounds that translation is viewed as compositional⁵⁸. Within this approach, the translation of a larger unit is seen as the translation of the smaller elements and a combination of these translated elements.

A possible reconciliation between the two extreme views (UT = word; UT = text) is suggested by Zhu (1999), who separates the notion of UT and the 'equivalence principle', and advocates the sentence as the ideal UT, on the grounds that this UT is necessarily formal. The 'equivalence principle' (between ST and TT) holds at the text level, while the UT should be processible in short term memory and should be syntactically comparable to its SL counterpart. The diverging views concerning the UT (more particularly as far as the size of the UT is concerned) are ascribable to the differences in approaches to translation. Views advocating for a smaller UT are essentially formalistic (a unit being a part of a higher unit), and see translation as "merely a transcoding process" determining the translation units and selecting the so-called 'optimal equivalent' (Snell-Hornby, 1988: 16). If the UT is seen as an 'independent and integrated meaning entity' (Zhu, 1999), then the UT is necessarily the text as a whole, but it cannot be processed as such cognitively. One should not use the term 'unit of translation' when considering the whole text, but rather the term 'unit of meaning'. According to Zhu, despite the text's status as a language unit⁵⁹, which verifies the textual integrity of its constituent units, it is unqualified to be the UT for text translation. On the other hand, the sentence is the smallest

⁵⁷ Lörcher, W., 1991a, 1993. Toury, G. 1986.

⁵⁸ The translation of a complex expression is a function of the translation of its parts and the way they are combined.

⁵⁹ A language unit has to be viewed in terms of its textual potential before it can function as a UT.

complete textual unit⁶⁰. The sentence might therefore be the right key functional translation unit. This theory backs observations of translators, who “translate sentence by sentence, and (...) will consciously be looking at the larger units only when it is necessary to do so.” (Newmark, 1988: 65)

3.2.3 Decoding/encoding approach

The decoding-encoding view of translation seen as a complex mathematical operation has always been quite popular, although (as every model of translation) highly controversial. For Wittgenstein, “translating from one language into another is a mathematical task, and the translation of a lyrical poem, for example, into a foreign language is quite analogous to a mathematical problem.”⁶¹ Some translation scholars developed models based on that view, like the decoding-encoding one⁶², according to which the translator converts one set of codes and structures to a second set by applying translation procedures to the source text.

Within the encoding/decoding model, the source and target texts are linked by the referent (supposed common to the ST and TT), the ideational content, and some coding rules. Most machine translation programs, built on this principle and developed from the transformational generative theory, correspond to the following pattern.

⁶⁰ The sentence represents a complete syntactic form with an independent information structure that as a speech act has been found to be ‘identifiable across languages’. (Lambrecht, 1994 : 34-35)

⁶¹ Wittgenstein, L., *Zettel*, 698, 1967: 121.

⁶² For Jakobson, the translator ‘recodes’ and transmits a message received from another source.

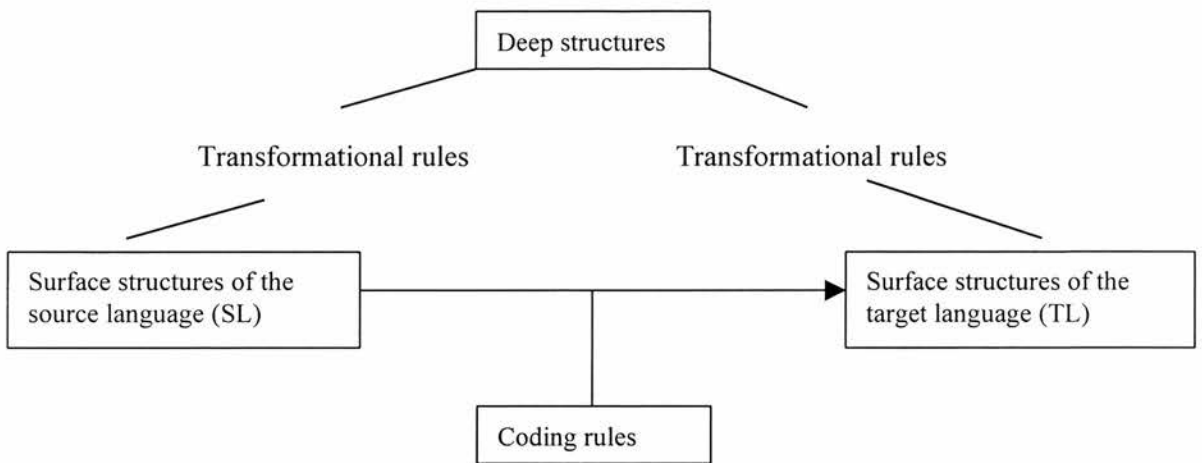


Figure 3-1: Coding-encoding model

According to this theory, every language expresses the transformation (at the level of surface structures) of the primary features of reality (content) located at the level of deep structures, and translating equates to reaching the deep structures of the languages in order to extract the 'ideational content', or in other word the meaning of the word, sentence or text. Within this (controversial) decoding-encoding pattern, the translator is assumed to convert one set of codes and structures to a second set by applying the following translation procedures to the source text:

1. Decoding of the surface structure of the source text;
2. Retrieval of the ideational content of the source text from the surface expression;
3. Identification of the plan and development of the source text from the ideational configuration;
4. Restructuration of the expression, development and plan of the text according to the target language standards;
5. Encoding of the modified expression, development and plan in the linguistic structures of the target language.

The decoding-encoding framework is derived from the theory of communication, but multilingual communication is more complex than the monolingual act of communication, calling for two alternating processes of encoding and decoding, as opposed to a single one in the case of monolingual communication. The translator analyses the message (formulated in the SL code) on the basis of his/her SL communicative competence, and then reconstructs the message in the TL through a succession of interrelated stages, keeping in mind the conventions of the text type in case. The reconstructed text is then sent to the ultimate receiver who decodes the TL text, on the basis of his/her communicative competence.

Catford (1965) acknowledges the operational usefulness of such an approach for machine translation, but he stresses that the 'transcoding view' is rather useless for the deeper understanding of the translation process. Proponents of the interpretive approach (Delisle, 1988; Admiral, 1994; Lederer, 1994) have also denounced this view, which they consider too restrictive, leaving aside cultural referential and cognitive elements. The transcoding view considers the linguistic (syntactic, lexical and morphological) features of the sentences, but is accused of missing the extra-lingual dimension of translation.

The communicative approach to translation also questions the decoding-encoding model, referring to the correlation between private non-linguistic 'meanings' (called 'mental contents by some authors) with public, linguistic symbols. The question of whether 'meaning' can be transferred (transmitted from the sender to the receiver) is a highly disputed one. John Locke⁶³ asserts that communication does not involve any transfer of meaning as such, but the transfer of (auditory or visual) stimuli that 'activate' already present meaning. Meaning itself, hence seen as a 'mental content' (located 'in the mind'), is therefore not transmittable. The most common answer to Locke's argument suggests that meaning is shared thanks to an accurate decoding of the signs transmitted from sender to receiver (apprehension of the 'concept' that underlies the linguistic symbols). "Within the terms of the decoding-encoding model, 'understanding a word'

⁶³ In Hacker, 1986 : 258.

is ‘understanding an associative relationship between the word and its meaning’⁶⁴ In communication, at least five elements are involved: the message transmitted, the system of symbols with which the message is processed and sent out (also called ‘code’), the sender of the message, the receiver of the message, and the channel used for the transmission.

3.2.4 The two-phase/three-phase models of translation

In the attempts made at formalising translation, two main models have emerged. The two-phase model considers translation as a process consisting of chronologically sequential phases⁶⁵: analysis (‘comprehension phase’, also miscalled ‘decoding’) and synthesis (‘reconstruction’, ‘restructuring’, ‘reverbalization phase’). The three-phase model considers another phase between analysis and synthesis: transfer of the ‘meaning’ of the received message into TL on the basis of an equivalent relationship.

The two-phase model based on the decoding-encoding principle, has been mainly used in computer sciences at the beginning of machine translation. This decoding-encoding view is based on the assumption that the translator converts one set of codes and structures to a second set by applying some kind of ‘mathematical operations’ on the ST. The phrasing ‘decoding-encoding’, although very useful, tends to be misleading since it refers to a different concept. In a way, strings of signs (parts of sentences, sentences, texts) are codes, each language using a different code. The words, ‘*book*’, ‘*livre*’, ‘*Buch*’, ‘*kniga*’ are arbitrarily codified in English, French, German and Russian. The spelling of each of these words depends on a socio-cultural and historical coding. But research in cognitive linguistics (Langacker, 1991b; Edelman and Tononi, 2000) tend to show that messages are not necessarily ‘coded’ in a way similar to a mathematical code, and translation seems less and less likely to be a code-switching operation on a sign-for-sign basis.

⁶⁴ W. Husson, 1994: 62.

⁶⁵ For Hervey and Higgins (1992), though, the operations do not occur successively but rather simultaneously.

On the other hand, the three-phase model appears to account in a better way for the actual process of translation. Haas (1968: 86) argues for the three-phase model in the following terms: “When we translate, we seem to establish a relation of three distinct entities, each separately apprehended: the two expressions seen on paper or heard in the air, and the meaning in the translator’s mind. The meaning, presumably, we ‘retain’ and translate, we ‘transfer’ it from one expression to the other”. The ‘transfer phase’, though, remains the greatest unknown territory in translation studies. Authors would mention the ‘translation strategies’, or the ‘cognitive operations’ taking place in the mind of the translator, but no theory is able to describe accurately and precisely the operations performed in the ‘black box’ (the mind). Translation scholars tend to rely on the bilingual competence of the translator to process the ST and produce an ‘equivalent’ TT. The three-phase model tends to be the dominant one in translation, and computational linguists have launched several programs, with the aim of formalizing an ‘*interlingua*’ in MT.

Some authors refute both the two-phase and the three-phase models on the grounds that they are unsatisfactory. Nord (1991), in particular, developed a looping model of translation, centred on the text and the translator, who is seen as the main figure in the process of intercultural communication. Her theory, contrary to most other theories focused on ST as the starting point in translation, claims that the process of translation is initiated by the analysis of the TT situation and function (or ‘skopos’), and followed by the analysis of the ST situation and function (then followed by the ‘classical’ transfer and synthesis scheme). The main interest of her theory lies in her analysis of translation as a forward-backward process. Hervey and Higgins (1992: 15) have a similar approach to translation (although they argue for the two-phase model) when they see it as ‘simultaneous’. Their use of the term ‘simultaneous’ might express a concern for clarity and simplicity in a book aimed at teaching translation methodology, rather than their true view of translation, which seems closer to the ‘looping process’: “One may not even realize that one has imperfectly understood the ST until one comes up against a problem in formulating or evaluating a TT. In such a case, one may need to go back to square one, so as to reinterpret and reconstruct the ST in the light of one’s new understanding of it. In this way, ST

interpretation and TT formulation go hand in hand. Nevertheless, for the purposes of discussion, it is useful to think of them as different, mutually, separable processes". The three-phase models (analysis-transfer-restructuring), such as those developed by Nida , Shveitser (1988) and Malone (1988), are criticized by other translation scholars, on the grounds that these models are marked by information theory and computer modelling approaches, but have little to do with what goes in the minds of actual translators. Some authors (Toury, 1982; Sebeok, 1986) even see translation as a four-stage process: decomposition of the initial entity into 'features', selection of the relevant features to be retained, transfer of these relevant features and (re)composition of a resultant entity around the transferred features. In this approach, the phases are not clear-cut, but rather form a non-interrupted continuum. Research in cognitive linguistics – particularly observations on 'simultaneous' interpreting – proves that language processing (and translation is one form of language processing) is a complex set of intertwined operations (Seleskovitch, 1975; Lörscher, 1991b)

The main issue in translation studies is certainly the 'transfer phase' question, and the formalization of the 'meaning' carried over from ST to TT, in a 'deverbalised' way (, Seleskovitch, Lederer, 1984), or conceptualised way (Langacker, 1987, 1991a, 1991b).

3.3 Meaning in translation⁶⁶

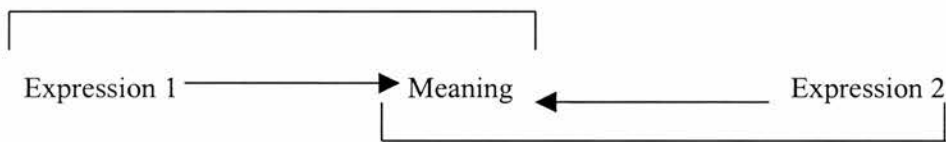
3.3.1 Meaning in translation

Whether they are text-oriented, socio-cultural oriented, linguistic or interpretive, the most salient theories of translation focus on the transfer of 'something' (linguistic, cultural, or textual elements). The interpretive approach is centred on the 'transfer of meaning', although 'meaning' itself remains largely undefined. For some translation scholars, 'meaning' belongs to linguistics since it is partly carried by linguistic elements (such as stylistic or phonetic features), for others

⁶⁶ "It is clearly necessary for translation theory to draw upon a theory of meaning."(Catford, 1965: 35).

it belongs to semantics or pragmatics. The division between syntax, semantics and pragmatics (Morris, 1938) was particularly attractive to computer linguists, but such a division tends to corner the issue of ‘meaning’, which has long been considered as belonging to philosophy and logic, rather than linguistics. While translation studies were traditionally included in linguistics, meaning – although central to translation⁶⁷ – tended to be skipped, overlooked, or assumed as being ‘intuitively grasped by bilingual speakers’ and belonging to cognition. For most translation scholars, meaning is defined as the ‘transferred’ element⁶⁸ from the ST to the TT, although rarely delimited, or precisely determined.

For Haas (1968: 87), “Translation is supposed to be possible on account of a twofold relation of an entity, called ‘meaning’; two expressions are viewed as ‘vehicles’ of the same meaning”.



Although, as Haas stresses, the crucial issue remains the determination of such a ‘theory of meaning’, relating an expression and its meaning, that is establishing a correspondence between two entities, one linguistic, the other extralingual⁶⁹.

He further discusses the indirect and direct theories of translation, in the light of meanings seen not as entities or objects (corresponding to expressions) but rather as the uses of expressions.

The meaning of a word is a collection, an organized recollection of many individual uses of it:

⁶⁷ Saint Jerome’s dictum in which he said that he followed Cicero in not translating « verbum e verbo » but « sensum de sensu » is taken as one of the principles of translation, and understood as the necessity to focus on meaning (main tenet of the interpretive approach). But Vermeer suggests that St Jerome’s dictum was wrongly interpreted, and translating ‘sensus e sensu’ would actually mean to translate words or phrases according to their grammatical form and meaning in a given text (as opposed to ‘verbum e verbo’ – to translate morphematically). (Vermeer, H.J, 1994: 7).

⁶⁸ « Translation is the transfer of meaning from one set of language signs to another. »(Lawendowski, 1978: 264)

⁶⁹ « When we translate, we seem to establish a relation of three distinct entities, each separately apprehended: the two expressions seen on paper or heard in the air, and the meaning in the translator’s

in verbal and non-verbal contexts, and in positions in which it contrasts with other words. Meaning can therefore be seen as an 'acquired property'.

According to Ogden and Richards (1923), 'meanings' (references') are linked to external things ('referents'), and the 'dualist' theory of linguistic signs can be expressed in three terms:

$$\text{Expression} \rightarrow (\text{'reference'} \rightarrow \text{'referent'})$$

Within this theory, the triadic scheme for translation is expressed by five terms:

$$\text{Expression1} \rightarrow [\text{Reference1} \rightarrow \text{Referent} \rightarrow \text{Reference2}] \rightarrow \text{Expression 2}$$

Within the direct view, 'reference' does not involve any mental fact, and expresses a dyadic relation:

$$\text{Expression} \rightarrow \text{Referent}$$

Translation is therefore seen as a relational scheme of three physical terms:

$$\text{Expression 1} \rightarrow \text{Referent} \leftarrow \text{Expression 2}$$

Haas then discards the dualist theory of meaning in the light of Wittgenstein's work.

For Kaiser-Cook (1994), meaning is subjective, culture-specific, and is a process of conceptual restructuring conforming to the linguistic conventions of the target culture.

Meaning is traditionally defined not as an entity, but as a relation, a property attributed to something:

"Meaning can best be formulated as consisting of a triadic relation of interpretation, whereby one term affects a mind in such a way that it is taken to stand in some respect for another term."⁷⁰

Still, as Haas expressed it, 'it is when we think of translation that we are most liable to become confused about meaning – and tempted to locate them in extralingual entities'. The principle of equivalence by translation leads us to believe that we extract something (that we call 'meaning'), although, according to Haas, this common sense **is not** a separate entity related to the expressions, but rather a **correspondence between their functions**.

mind. The meaning, presumably, we 'retain', we 'transfer' it from one expression to the other. » (Haas, W., 1968: 86).

The issue of meaning in translation is also raised by Quine (1959, 1960) who underlines that meaning being subjective and culture-based, the subjectivity of both author and translator severely limits accurate transfer of meaning. Empirical meaning (as defined by Quine) is supposed to be what the sentences in one language and their firm translations in a completely alien language have in common, but through his theory of the indeterminacy of translation, Quine explains that interlinguistic synonymy of words and phrases can only be considered within the terms of some particular system of analytical hypotheses.

3.3.2 Meaning and concept

The different approaches to meaning are rooted in philosophy and express differences in philosophical traditions. Meaning emerged in the Stoic tradition as triadic relation consisting of *semainon* (that which signifies), *semainimenon* (that which is signified), and the object referred to in signification. Katz (following Plato's direct view) argues that actual linguistic forms derive from a world of pure linguistic forms. He considers meanings as abstract objects that exist independently of the minds that perceive them. On the other hand, Jackendoff⁷¹ and Lakoff⁷² (following Aristotle's indirect view) advocate a mental approach of meaning. They consider semantic properties as invariant because the constituents of our mentally projected worlds of reference are derived from the same mental processes. The mentally projected world is seen as made of types (semantic, visual and so on), which form the predicates of the code in which our mental representations are constructed, processed, validated and developed.

⁷⁰ Sebeok, 1986: 510.

⁷¹ « Meaning in natural language is an information structure that is mentally encoded by human beings ». (Jackendoff, 1988: 81).

⁷² « Meaning is an imaginative projection, using mechanisms of schematization, categorization, metaphor and metonymy to move from what we experience in a structured way with our bodies to abstract cognitive models ». (Lakoff, 1988: 121).

From a philosophical viewpoint, the definition of meaning as concept (therefore mental entities) raises much controversy. For example, Frege considered meanings as public property (since the same meaning can be grasped by different persons), abstract entities, neither psychological, nor philosophical.

In order to reconcile these contradicting views, a useful typology of meaning was set up. Depending on the approach, one may consider the lexical meaning (literal as opposed to figurative meaning, polysemy, homonymy, synonymy), the sentence meaning (depending upon the meaning of its constituent lexemes, and similarly determining the meaning of some of these lexemes), the grammatical meaning (defined by the syntactical structure of sentences), the utterance meaning (as determined by pragmatics), the descriptive meaning of statements (truth-conditional semantics), the social meaning (use of language in social relations, also called sometimes expressive meaning). The most developed area of research concerns linguistic semantics (study of literal, decontextualised), grammatical meaning), which is considered as a branch of semantics (study of linguistic meaning in particular), or a branch of semiotics (study of meaning in general). Translation – as explained in the above developed paragraphs – being a far more complex process than a linguistic operation, it seems sensible to consider the issue of the translation of meaning from a semiotic perspective rather than from a purely semantic one.

3.3.3 Transfer of meaning

If one considers the transfer operations between the source text and the target text, there seems to be one ‘entity’ (whether it is called ‘ideational content’, ‘semantic representation’ or ‘virtual translation’, defined as a mental model of the elements and relations which exist in the mental space between real source and not-yet-realized target text), which corresponds to what is commonly called ‘**meaning**’ (transferred from the ST to the TT). The **virtual translation** is defined by Neubert and Shreve (1992) as a composite of the possible relations between a source text and a range of potential target texts. This virtual translation/mental representation includes

the propositional content and the illocutionary force of the messages underlying the source text. Neubert and Shreve identify seven parameters determining the textual character of the virtual translation: intentionality, acceptability, situationality, informativity, coherence, cohesion, and intertextuality. In theory, once we understand how the mental representation of the virtual translation is constructed, and how it is embodied in the target text, it should be possible to conceive a translation programme that would reproduce the cognitive processes involved in translation. The virtual translation hypothesis support the approach to cognition as developed by applicative and cognitive grammars. In this framework, different levels of interpretation are supposed to be successively derived. The virtual translation can be said to be expressed in the 'universal' language supposed to encompass all cognitive processes. Applicative and cognitive grammars are presented in chapter 5.

Many scholars developed theories about meaning extraction from the deep structures, while others focused their attention on the surface structures as expression of part of meaning. For Nida (1982), translating meaning is possible, but it implies translating the total significance of a message in terms of both its lexical or propositional content and its rhetorical significance. The translator has to pay attention to intent (what the translated text is presumed to accomplish in terms of impact on the receptor), coherence (the relation of the text to the real world of objects and events), intertextuality (the relation of the text to other texts), and impact and appeal on the receptors. Meanings are not entities or objects corresponding to expressions, they are the uses of expressions, hence the necessity for the translator to look for the total significance of the text. By doing this, one may be tempted to assume that meaning is 'located' in extralingual entities somehow related to the different linguistic expressions. This view has been thoroughly discussed not only by translation scholars and linguists, but also by philosophers, whose insights into languages are highly valuable for an integrated theory of translation.

3.4 Semiotics, meaning and translation

3.4.1 Translation as communication, or translation as sign

Translation deals with written texts, which can also be described as exchanges between the writer and some implied readers (Hatim and Mason, 1990). As such, translation not only expresses an interaction between signs within texts (context), but also the communicative interaction between producer and receiver of the message. Translation is therefore definitely a semiotic process, even though semiotics and translation studies have somehow excluded each other for a long time.⁷³

The semiotic framework allows for a more elaborate communicative theory of translation as demonstrated by some semioticians, more particularly, E. Blühdorn.

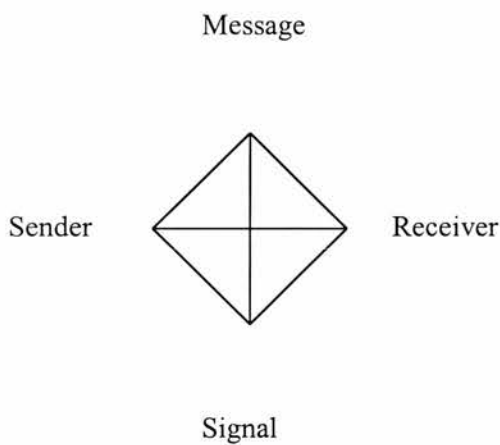


Figure 3-2: Model of the prototypical communicative situation (Blühdorn, 1998: 36)

The message is incorporated in the sign.

⁷³ "...Both translation studies and semiotic studies address, albeit from different methodological vantage points, aspects of communication, and both are concerned with the use, interpretation and manipulation of messages or texts - that is of signs." (Gorlée, D.L., 1994:11)

"By a sign, we usually understand something meaningful, a set of visual, auditory or other stimuli that represent a certain message. Strictly speaking, then, the sign already incorporates the message, for the message is the meaning, and the meaning is part of the sign." (Blühdorn, 1998: 36)

If we consider translation as a communicative process, the written sentences represent the signal (succession of words coded according to some rules, the syntax) while the whole text to be translated conveys a message between the sender and the receiver. But the distinction between signal and message has too often been blurred, creating misconceptions about the translation process, and errors in the development of machine translation. The sentences represent a signal encoded thanks to a natural language, but the message does not equal the signal. The signal is used by the sender who wishes to send a message, but it has to be decoded by the receiver (who constructs in his/her mind a representation of the message intended by the sender). The signal does not constitute any meaning, only the message does⁷⁴. Hence, the role played by the carrier of the message, the signal, and its form. The syntax (the rules governing the structure of the signal) is essential in the analysis of the signal, but is only the first level of analysis leading to the semantic content of the message. Syntax is linked to the signal, semantics is linked to the message⁷⁵, and the difficulty for machine translation consists in bridging the gap between these two separate entities/levels. This is why, as developed by Blühdorn (1998), access to semantic questions open itself through grammar and/or through pragmatics, and this is why the analysis presented in chapter 6 is of some interest in translation studies, since it aims at determining elements in the signals (source and target texts) that can be used as a basis for a semantic interpretation, although these elements by themselves are in some respect independent from the message (or 'meaning').

⁷⁴ "In relation to the sender, the message has the function to stimulate him to produce signs; in relation to the receiver, it has the function to constitute a problem to be solved through cognitive effort; and in relation to the signal, the message constitutes its meaning." (Blühdorn, 1998: 39).

⁷⁵ "Grammar is viewed as a discipline that should be interested exclusively in abstract and formal structures, neither considering meaning nor communication." (Blühdorn, 1998: 43).

Blühndorn's approach is derived from Bühler's theory⁷⁶, who developed in the thirties a semiotic approach to communication. The sign links the sender (producer/emitter), the receiver (who perceives and interprets) and the objects/states of affairs that are communicated. Within this approach, the sign is simultaneously an expression (*Ausdruckfunktion*), a signal (*Appellfunktion*), and a symbol (*Darstellungfunktion*).

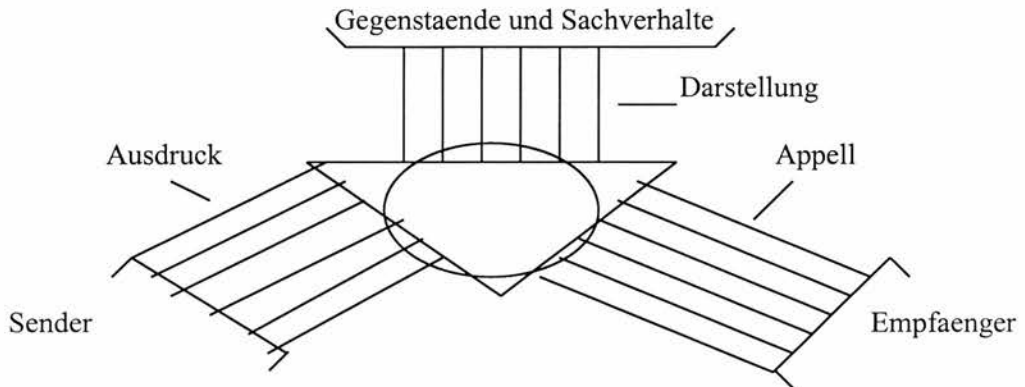


Figure 3-3: Bühler's Organon model (from Blühndorn, H., 1998: 29)

Gideon Toury (1980) underlined the importance of the relationship between the codes, and he considered that this relationship should be the first criterion when one considers the transfer process at the heart of translation. Moreover, Toury defined translation as functions, which map target messages. This is why he put emphasis on the two sets of relationships that should be taken into account:

- (a) relationships between target text and source text (contrastive text-linguistics);
- (b) relationships between target text and target language, and/or between target text and the target-relevant textual system (contrastive linguistics).

⁷⁶ Bühler, K.1934.

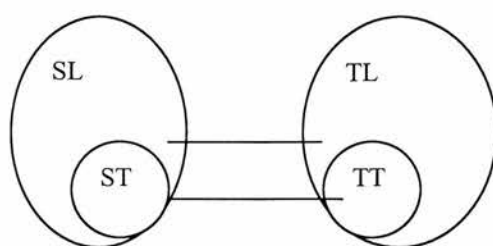


Figure 3-4: system relationships involved in translation

One of the issues in translation is therefore the establishment of correspondences between the systems involved (or codes), as frameworks of the signals that encode the messages (texts).

3.4.2 Semiotic approach to translation

The subsequent paragraph will be developed within a Peircian approach rather than a Saussurean one, and the following terminology will be adopted:

- translating refers to a series of operations whereby one semiotic entity is transformed into, and replaced by, another entity, pertaining to another (sub-code) or semiotic system;
- a translation is an entity pertaining to a single semiotic system, while, at the same time, presupposing the existence of another, logically and chronologically prior entity in another system along with factual equivalence between the two entities (Sebeok et al, 1986).

Saussure's work is mainly concerned with a theoretical typology of entities and aims at isolating the class of 'sign' in order to apply general principles of 'systems of conventions for communication' to the study of language. Peirce's theory covers a larger range and deals with signification from a philosophical viewpoint, presenting the world as a large and complex system of signs. His theory tackles not only symbols (which represent abstract objects and relationships, such as logical, chemical, algebraic formulas, and diagrams; Eco, 1984) but also icons and indices (Hervey, 1982). The term 'semiotics' will be used (referring to 'semiotic' in Peirce's theory, as opposed to Saussure's 'semiology').

For the semiotician, language being only one of many codes, the concept of translation is not limited to interlingual translation (between two natural languages). Translation can be conceived

similarly between two different sign systems, between a natural language and an artificial language, or between two artificial languages. The study of the processes involved in human translation is, within a semiotic perspective, a preliminary step towards a better description of the translation process at large.

Translation being defined as a transfer (of 'meaning') between a source text (in a source language) and a target text (in a target language), or as a communication process, determined by some 'equivalent features' receives a new multidimensional approach from a semiotic viewpoint. Most translation scholars, whatever their main concern (linguistic, text-linguistic, communicative, socio-linguistic) resort somehow to semiotics in their theories:

"No linguistic specimen may be interpreted by the science of language without a translation of its signs into other signs of the same system or into signs of another system."⁷⁷

"Semiotics deals with the processing and exchange of information both within and across cultural boundaries. Translating can be envisaged as the process which transforms one semiotic entity into another, under certain equivalence conditions to do with semiotic codes, pragmatic action and general communicative requirements."⁷⁸

"In translation, the relation between the source language text (A) and its receptor language text (B) stands as the semiotic relation of sign (S) to object (O), or of signifier to signified. In brief, source to translation is as sign to object. Expressed in a formula of ratio equivalence, we have:

$A : B :: S : O$, or source : translation :: sign : object."⁷⁹

'Translating' in its restrictive sense (Sebeok, 1986) selected for the following discussion deals with interlingual translating, referring to processes where the two respective systems (primary codes) are two ontologically equivalent, different natural languages. The initial and resultant entities are both linguistic utterances, i.e. texts, belonging, in addition to the general linguistic systems, also to secondary modelling systems, which are of the nature of text-types, or textual traditions. Translating is therefore an intertextual activity involving cross-lingual procedures.

⁷⁷ Jakobson, R., in Brower, 1959: 234.

⁷⁸ Hatim, B., Mason, I., 1990: 105.

Semiotics allows for an integration of different (sometimes even apparently exclusive) approaches to translation:

- translating is a communicative act, and "semiotics concerns itself with communicative acts/events" (Hervey, 1982);
- translating is a series of operations performed on texts, and texts are sign systems ("The text is an isolated, self-contained semiotic formation, with its integral indivisible meaning and its integral indivisible function", Lotman, 1990: 47);
- translating is a socio-cultural process, a cross-cultural communication. Since culture may be defined as "a system of social codes that permits the expression of information with signs in order to make it the patrimony of human collectivity" (Lotman, in Lucid, 1977: 214), translating builds a bridge between two different cultures, two sign-systems;
- translating is a succession of operations performed on messages belonging to different natural languages, themselves constituting semiotic systems (systems of conventions for communication);
- translating aims at transferring 'meaning' from one message to another, and "in essence, signs are mediators between messages and signals" (Hervey, 1982: 13). The main and most characteristic function of semiological systems (systems of forms aiming at establishing communication by means of arbitrary signs) is to mediate between thought and physical expressions (sounds) (or between the signified and the signifier, in Saussurean terms);
- translating expresses complex cognitive processes, dealing with conceptualisation and symbol manipulation, which is a recent domain of research in semiotics.

The different approaches to translation, reconsidered from a semiotic viewpoint, may become complementary and even allow for a deeper understanding of the processes involved in this linguistic, cognitive and socio-cultural operation. Semiotics consider human experience as organized around the sign (triadic relation between the interpretant and its object).

⁷⁹ Barnstone W.1994: 91.

The preceding paragraphs demonstrated that translation is a complex (cognitive) operation (or rather interrelated operations) performed on texts (ST and TT) between source and target languages. Depending on the approach taken, one may find nearly as many definitions as one wishes. For Catford (1965) (text-linguistic approach), translation is "an operation performed on languages: a process of substituting a text in one language for a text in another." On the other hand, the development of communication studies led to another definition: "translation is the transposition of *messages* between tongues."⁸⁰ Central to this discussion on translation are the operations (that have to be precisely defined, and possibly formalised), the text (or message), the source and target languages. Whether two-phase or three-phase models, most translation scholars agree on the cognitive processes involved in translation. As early as 1969, Nida and Taber underlined the complexity of these processes (defined by Robinson, as *mental* processes) through their presentation of the phases of analysis, transfer and restructuring which allow the 'substitution' of the ST into a TT, as summed up by Catford.

The 'cognitive' processes (or mental operations) involve different operations: the source language (as code A), the target language (as code B), the message encoded, the sender and the receiver (with special attention to the translator who is simultaneously receiver and sender), the 'meaning' of the message (that is supposed to be transferred), the source and target texts, and the relation between the text and its linguistic, social and cultural context (pragmatics). Should translation be included in linguistics (since it deals with source and target languages), in pragmatics (since it deals with the relation between the text and its socio-cultural context), in semantics (considering the central issue of meaning in translation), or in cognitive sciences (as it is the expression of mental operations)? All these disciplines have already been redefined within a semiotic perspective (semiotics being defined by Peirce as "the science of signs/sign systems"), more particularly by Morris⁸¹, who divided general semiotics into three subdisciplines: syntactics (relations between signs and other signs), semantics (relations between signs and their objects), and pragmatics (relations between signs and their users).

⁸⁰ Barnstone, 1994: 89.

⁸¹ Morris, 1938.

Within Peirce's approach, linguistics is seen as a subdiscipline of semiotics. This Peircean approach will be retained for the following discussion, even though the respective positions of semiotics and linguistics (does linguistics include semiotics, or does semiotics include linguistics?) have been at the heart of heated debates, particularly between the European trend (Saussure, Prieto, Barthes) and the American one (Peirce, Morris, Searle). Similarly, translation can be described as a subdiscipline of semiotics, partly overlapping linguistics, but not totally included in linguistics, linguistics being concerned with the study of natural languages, while translation deals with the transfer of 'meaning' (or 'messages' from a communicative viewpoint) between two linguistic entities and involves a sender and a receiver of the message. This communicative dimension of translation cannot be wholly accounted for solely within a linguistic approach. As demonstrated by Hans Strohner⁸² (following Bühler and Morris, but developed from systems theory and information theory), communication can be presented as a complex system, which unites the communicators and the cognitive text processing that takes place in their brains. This communicative dimension of translation deserves as much attention as the structuralist approach that has dominated linguistic studies (hence translation studies) for a while, or the discourse analysis approach that underestimated the cognitive element at stake. The diverging views on language and communication expressed in opposed schools unfortunately led to a misunderstanding of the multidimensional process of translation, and are to be blamed for the apparent failure of machine translation which was developed on a too limited approach to translation (while, as this dissertation is attempting to demonstrate, translation deals *first* with texts - as messages, *secondly* with languages).

Translation must be considered from the viewpoint of all the approaches, and true machine translation can be developed only once the full dimension of this process is analysed and once the operations involved are formalised.

3.5 Conclusion

3.5.1 Filters, shifts, and decision processes⁸³

The issues presented above demonstrate that translation is a complex communication process activating cognitive operations. This process takes place between two messages (texts) belonging to different sign systems (source and target languages, but also source and target cultures). Although most machine translation software packages tend to consider only the linguistic systems involved through grammar formalisation, human translators necessarily take into account not only the source and target linguistic systems but also pragmatic features such as the type of texts to be translated, the potential readership the target text will be addressed to, the literary style of the author and the target culture. The interpretive approach to translation insists on a bilingual and bicultural competence of the translator and on the 'unconscious' performance of the translator, whose total mastering of the source and target linguistic and socio-cultural systems determines the translation strategies.

A translator makes choices and decisions at all stages of translation: analysis of the ST (semantic, stylistic, phonetic elements to be selected), transfer (purpose of the translation, targeted readership) and synthesis of the TT (linguistic possibilities of the target language, stylistic features of the TT within the target culture...). Whether consciously or unconsciously, speakers, writers (and therefore translators) apply different communication strategies according to the languages involved (the difference between 'tu' and 'vous' in French, although absent in the pronoun system in modern English, can be – and usually is – expressed by different linguistic means – use of first names or titles...). The competence of the translator deals with

⁸² Strohner, H., *Textverstehen, kognitive und kommunikative Grundlagen der Sprachverarbeitung*, 1990.

⁸³ For J. Levy, "these decision processes in translation have the structure of a semiotic system, having its semantic aspect (i.e. a repertory of units defined through their relation to their denotata), and its syntax (i.e. rules for combining these units – whether by units we mean paradigms or instructions). As all semiotic processes, translation has its pragmatic dimension as well." (Levy, 2000: 156)

this capacity to know all the linguistic features both of the SL and the TL, to determine the strategies of the author, and to select ‘corresponding’ strategies in the TT he/she is creating. But, as was underlined by Hervey and Higgins (1992: 24), the “transfer of meaning from ST to TT necessarily involves a certain degree of translation loss” (due to the differences of the linguistic systems). Russian, for example, does not have a single verb corresponding to ‘to go’. The translator must choose between ‘to walk’, ‘to run’, ‘to drive’, ‘to ride’, or ‘to fly’. The decision will be determined by the context, and by the short-term memory of the translator, who will infer whether the character is in a boat, a car, on a plane, a bicycle, or is on foot. Such differences of the linguistic systems raise insurmountable difficulties to machine translation, the parser having to select between several ‘correspondences’.

Consciously (and even sometimes painstakingly) applied, or unconsciously used, some ‘translation filters’ are determined by the linguistic systems involved. Despite the debate about the autonomy of semantics and syntax, the only tool at the disposal of the translator is the target linguistic system. The translator may opt for a very literal translation and add some lengthy notes to explain the meaning of the source text and its connotations, as well as the style of the author (an extreme case, where the TT would hardly be a translation). In most cases (fortunately), the translator uses the ‘filters’ and selects the elements (lexical, grammatical, stylistic, social...) he wishes to transfer in the TT. Although some correspondences between languages (such as ‘il peut venir’ ↔ ‘he can come/he may come’) have been determined, these correspondences do not cover all the word combinations that are being used in a ST. Faced with a syntactic structure that can be directly transferred in the TL, the translator must transform the source sentence(s) in order to convey the ‘meaning’ in the target text for a target readership in a target culture. A single grammatical form may have several meanings, or functions. One of these is the primary function. All the others are skewed, because of the mismatch between the grammatical form⁸⁴ and the meaning. Skewing is central to translation, since it explains the

⁸⁴ The grammatical structure refers here to the surface forms of a language.

failure of most "translation software", built on the assumption that a literal word-for-word and grammatical-structure for grammatical structure translation is possible.

"A language is a complex set of skewed relationships between meaning (semantics) and form (lexicon and grammar). Each language has its own distinctive form for representing meaning. Therefore, in translation, the same meaning may have to be expressed in another language by a very different form. To translate the form of one language literally according to the corresponding form in another language would often change the meaning, or at least result in a form which is unnatural in the second language. Meaning must, therefore, have priority over form in translation. It is meaning which is to be carried over from the source language to the receptor language, not the linguistic forms." (Larson, 1984: 9-10)

For Hervey and Higgins (1992), the work of the translator consists in passing the TL text through different 'filters' (lexical, grammatical, idiomatic, stylistic, social...) in order to reduce this translation loss. Manuals like *Thinking Translation* (Hervey, Higgins, 1992), *Stylistique comparée du Français et de l'Anglais* (Vinay & Darbelnet, 1977), *Approche linguistique des problèmes de traduction* (Chuquet & Paillard, 1989) aim at giving tools to the translator to identify and use systematically these filters. Vinay and Darbelnet (2000) list seven different procedures at the disposal of the translator, classified according their characteristics. Direct procedures group borrowing, *calque* (or loan translation), literal translation, while oblique procedures group transposition⁸⁵, modulation⁸⁶, equivalence, and adaptation. For example, literal translation may arise between French and English on the grounds of common metalinguistic concepts, which reveal physical coexistence (periods of bilingualism); or on the grounds of a certain convergence of thought and sometimes of structure, which are present among European languages (as exemplified by the creation of the definite article, the concepts of culture and civilization). Modulation is acceptable, when a literal, or even transposed

⁸⁵ Transposition involves replacing one word class with another without changing the meaning of the message ('Dès son lever' – 'As soon as he gets up/As soon as he got up').

⁸⁶ Modulation is a variation of the form of the message, obtained by a change in the point of view.

translation would result in a grammatically correct, but unidiomatic or unsuitable utterance ('it is not difficult to show' – 'il est facile de démontrer').

What Vinay and Darbelnet call 'filters' are often called 'shifts' by other scholars (Catford, Halliday). Catford⁸⁷ defines shifts as 'departures from formal correspondence in the process of going from the SL to the TL'. He distinguishes level shifts (a SL item at one linguistic level has a TL translation equivalent at a different level – shifts from grammar to lexis and vice-versa), and category shifts (departures from formal correspondence in translation). He also differentiates several kinds of category shifts, structure shifts, class-shifts⁸⁸, unit-shifts, and intra-system shift.

3.5.2 Textual dimension of translation

If we follow the pattern of communication presented by Lotman⁸⁹, and the pattern of translation of Nord⁹⁰, we have the following coexisting sequences:

⁸⁷ Catford, in Venuti, 2000: 141.

⁸⁸ Class-shift occurs when the translation equivalent of a SL item is a member of a different class from the original item. The 'class' is defined (following Halliday) as 'a grouping of members of a given unit which is defined by operation in the structure of the unit next above'.

⁸⁹ Lotman, Yuri M. 1990: 11.

⁹⁰ Nord C., 1991: 34.

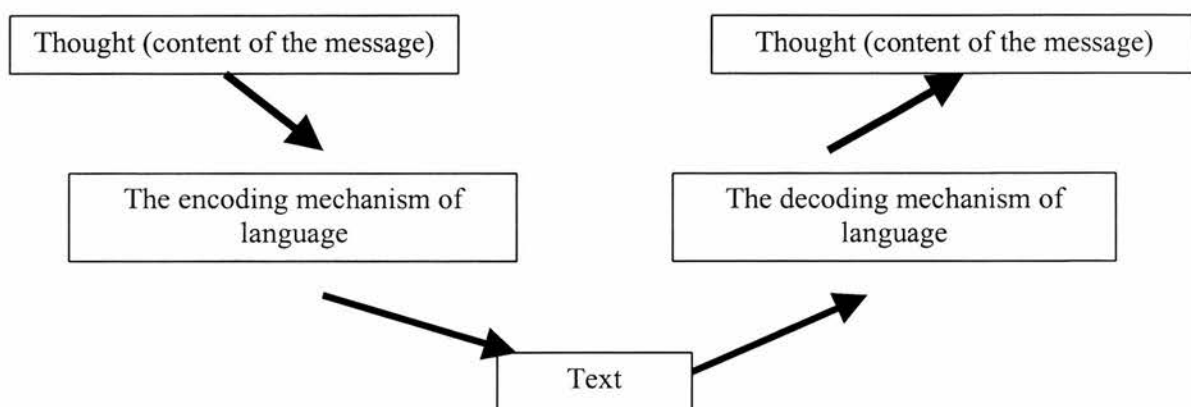


Figure 3-5: The coding process (from Lotman, 1990: 11)

and

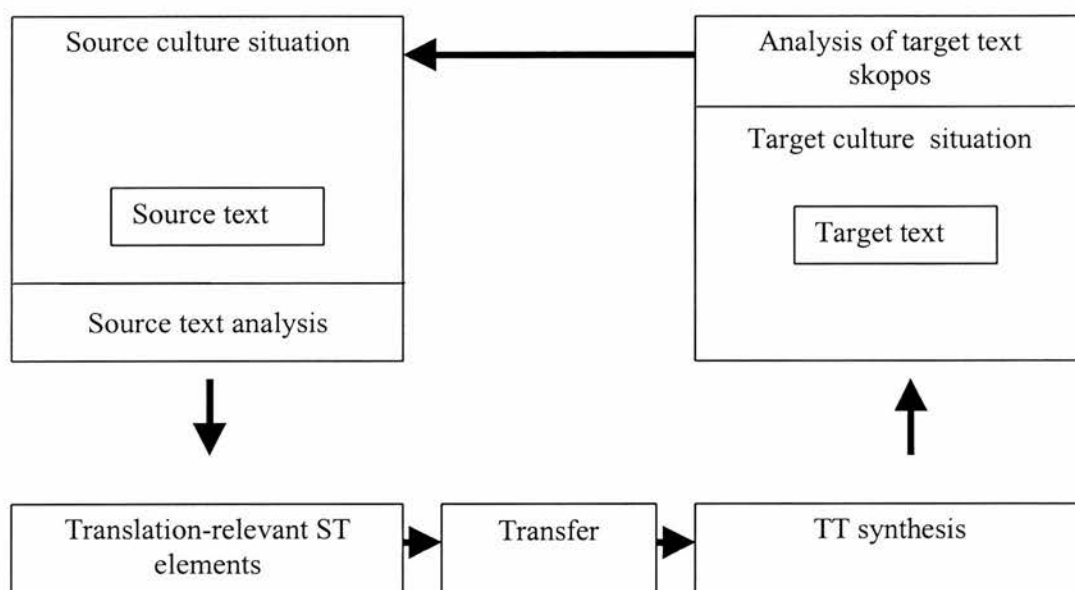


Figure 3-6: The translation process (from Nord, 1991: 34).

As Nord stresses, in many cases, the translation process does not start necessarily with the analysis of the source text and source culture, but rather with the analysis of the target text skopos, which will determine the translator's strategies. The figure above must be read as a circular operation, with constant movements (loops) between the source and the target text skopos.

As shown by Lotman (1990), the restrictive encoding-decoding communication process, although logical and appealing in its presentation, fails to account for the differences in the

codes used (natural languages), and for the interpretation process activated in translation. The analysis of the codes involved is only part of the series of the operations at stake in translation.

Translation is a complex process involving numerous factors. It is a set of operations, but performed by human beings **on texts** belonging to different linguistic systems. The source and target texts are linked by some kind of 'equivalence' relation. Neubert raises the problem of such a wording. For him, 'equivalence' in translation should not be understood in a mathematical logical identity sense, but should be read as the "transfer of something that stands in a *value relationship* to a certain 'subject' or topic expressed in the source language".

« Equivalence in translation (...) is a functional concept that can be attributed to a particular translational situation. » (Neubert, A. 1994: 413).

Linguistic equivalence alone is unable to account for the processes at stake in translation, even though linguistic equivalence between source language and target language must be acknowledged in translation studies, as long as one remembers that linguistically equivalent terms might not be 'translationally' equivalent. This discrepancy may be explained by the fact that « the complex demands on adequacy in translation involve subject factors and transfer conventions that typically run counter to considerations about 'surface' linguistic equivalence. » (Neubert, A. 1994: 414). One of the main elements of this translational 'equivalence' is said to be the 'meaning' of the text, which is supposed to be transferred from the ST to the TT. As developed in the above paragraph about meaning in translation, 'meaning' is far from being precisely determined. Moreover, as Neubert (1994: 414) stresses, denotational identity ('meaning' as defined by some semanticists) between ST and TT is not necessarily the hallmark of a good translation. One hundred per cent equivalence in terms of reference is not only impossible (the target text being destined to a particular readership whose reference will necessarily be different from the source readership's reference) but even not to be systematically sought for.

Although 'meaning' is partly carried by the linguistic signs composing the ST and TT, it is not totally included in these signs. Lexicon and syntax of the ST contribute to the determination of meaning, lexicon and syntax of the TT are supposed to express the meaning 'carried over',

meaning of the ST depends on the interpretation of the reader (and more particularly of the translator), and meaning of the TT similarly depends on the interpretation of the final reader. The 'decoding' of the ST and the transfer of elements (syntactic, lexical and semantic elements, as well as 'skopos') depends on the translational 'competence' (composed of language competence, subject competence and transfer competence) of the translator, or what most translation scholars call the **cognitive system** of the translator:

« Transfer competence refers to the mental equipment that constitutes the translator's unique cognitive set or ability of matching language and subject competences. In detail, these are the processes comprised by ongoing search, retrieval, monitoring and evaluation mechanisms. » (Neubert, 1994: 414).

Different theories attempt to describe and explain the processes involved in the 'translation operation'. The **production system theory**, on the one hand, traditionally presented translation as the application of translation rules according to a planned series of steps, based mainly on linguistic equivalence. The **connectionist theory**⁹¹, on the other hand, presents translation as a transfer of a mental representation of stretches of the source text.

« Like a hologram, a stretch of source text is holistically perceived, 'processed' by the translator as realised or constituted by a number of 'parallel distributed' form and/or context elements. The power of this translational text processing comes from how the units are connected. » (Neubert, 1994: 415).

The second approach is nowadays the commonest view amongst translation scholars, especially those developing their research within an interpretive, or cognitivist framework. Some research has been done on the cognitive operations performed in simultaneous interpreting (prediction, memorization, rephrasing...), but little work has been done on the cognitive operations expressed in written translations. Contrastive translation studies are mainly concerned with the comparison of two linguistic systems, but the similarities and differences that can be determined between a source and a target text are the expressions of the interpretative process

performed by the translator. Languages encompass human cognition and translation is a cognitive operation.

So far, research in cognitive sciences unfortunately has not been able to determine precisely **how** a translator uses his/her cognitive system to operate the complex transfer of lexical, syntactic and semantic elements from the ST to the TT. Meaning is supposed to be contained (at least partly) in these elements, and it is supposed to be carried over from the ST to the TT. If 'meaning' is actually contained in the ST, and transferred to the TT, it is necessarily processed during these cognitive operations, and could therefore (at least in theory) be determined and formalised. The analysis of the translation process, through the comparison of human translations of a source text is a modest, but necessary first step toward a better understanding of the creation and transfer of meaning.

⁹¹ Johnson-Laird, 1988. The connectionist theory is also called parallel distributed theory in cognitive sciences.

4 MACHINE TRANSLATION

4.1 Introduction

4.1.1 History of machine translation

The idea of a machine able to translate languages automatically (via an intermediate 'universal language') dates back to the seventeenth century, but the first concrete proposals for a machine were made in 1933 (patents issued simultaneously by Petr Smyrnov-Troyanski and Georges Artsrouni). These pioneers did not manage to develop their products, and the fatherhood of machine translation is usually attributed to Warren Weaver, who drew a rather simplistic analogy between the process of translation and the process of decoding unknown signs⁹², based on cryptology.

Fully automatic machine translation (FAMT) as foreshadowed by enthusiastic computer researchers⁹³ after the Second World War benefited from highly-subsidised programmes run by military and defence agencies all over the world during the Cold War. For example, the first software packages were devised by the Americans and used by the U.S. Department of Defence to translate Russian documents into English. Linguistic theories of Harris (1954, 1968) and Chomsky (1965, 1969) and the subsequent results of research in generative grammar were supposed to provide the framework for machine translation (MT) on the assumption of universal deep structures. Despite huge investments and the improvement of computers, MT went through a very serious recession after the publication of the Automatic Language Processing Advisory

⁹² "I have a text in front of me which is written in Russian but I am going to pretend that it is really written in English and that it has been coded in some strange symbols. All I need to do is strip off the code in order to retrieve the information contained in the text." (cited in Arnold et al. 1994:13)

⁹³ The first conference on MT was organized by Y. Bar-Hillel at the M.I.T., U.S.A., in 1952.

Council (ALPAC) report⁹⁴ and the poor quality of the first output translations. The ALPAC report evaluated the quality of the output of various MT systems and found it to be so poor that a considerable amount of human post-editing was necessary. 'The report concluded that since there was no shortage of human translators and no cost advantage in machine translation after fifteen years of hard work, there was no justification for further government funding of machine translation for the purpose of developing commercial systems to compete directly with human translators.' (Melby 1995: 30). Temporarily discredited, fully-automatic machine translation (FAMT) gave way to other fields of research more likely to foster the birth of actual useful tools, such as word-processors, data-banks, spell-checkers, on-line dictionaries and so on. Theoretical research also downgraded its objectives and claims, focusing on the study and processing of specific kinds of texts less prone to ambiguities: weather forecast reports (as in TAUM-Meteo) or technical manuals (containing no pronouns, and only a limited range of syntactic structures or grammatical forms). Research intensified in the 1970s and 1980s, and commercial systems gained ground, more particularly in Europe. The MT system Systran (originally developed for the US Air Force) was bought by the European Commission in 1976 and is still extensively used today (up to 2,000 pages are mechanically translated per day). The European Commission also funded the EUROTRA project (multilingual MT system for all EC languages), which ended in 1992, the system being unable to process at a truly multilingual level.

The birth of the worldwide Web in the 1990s marked a renewal of interest in FAMT. Nowadays decrypting Russian confidential documents sounds less important than accessing rapidly all kinds of documents – strategic or commercial – in any language, but there are not enough human translators to cope with the huge and ever increasing demand from users all over the world. Much hope lies in Cross Language Information Retrievers, which should allow any user to get all documents available on the Web (in any language) on any topic, via a query in his/her own language.

⁹⁴ ALPAC, 1966.

The United Nations Organisation (UNO) launched a vast programme in 1996 (involving 17 research teams) which aims at devising a Universal Network Language (UNL). This project marks a shift of approach and attitude in natural language processing (NLP) and machine translation. Contrary to the concept of literal translation (on which most software packages were traditionally based), this new programme is expected to analyse and express the 'meaning' of the input texts. Realistic objective or new utopia? Specialists in the field of natural language processing (NLP) all agree that the issue of meaning is central, but most of them acknowledge the complex problems⁹⁵ still to be faced before the implementation of such a FAMT software.

4.1.2 Definition of machine translation

For many linguists, there is simply no such thing as 'machine' translation, translation being a *human* operation at the crossroads of psychology, linguistics and sociology. On the other hand, computer linguists, although being aware of the limits of computerisation in that field, have attempted to formalise the operations involved in translation ever since the birth of computer science.

Automatic – or machine – translation is one application of computer science. Despite much disappointment in that field of research (no software has so far been able to translate accurately any single text), research done over the last decades led to some breakthroughs, reliable products (as in the case of TAUM-Meteo), and even software good enough to produce a rough translation allowing the user to assess the content of a text and then pass it to a human translator. Nevertheless, machine translation, at its present state of development, does not claim to be equivalent to human translation. Built around syntactic parsers (completed with semantic parsers and knowledge data bases), machine translation falls short from properly analysing the 'meaning' of a text and still seems unable to 'transfer' this meaning. Current computers are still unable to reproduce the cognitive operations performed by a human translator.

⁹⁵ such as anaphora resolution, or syntactic and semantic ambiguities.

The term 'machine translation' (or 'automatic translation'⁹⁶) covers today a wide range of applications, from machine-aided translation tools to human-aided translation software. Machine-aided human translation (MAHT) aims at facilitating the translator's work with powerful electronic tools such as word-processors, data-banks, spelling checkers, on-line dictionaries and so on. Human-aided machine translation (HAMT) requires pre-editing and/or post-editing by a translator, so that the text is disambiguated and converted into a format which can be processed by the program. Machine translation, as defined by Hutchins and Somers (1992: 3), groups all "computerized systems responsible for the production of translations from one natural language to another, with or without human assistance. The central core of MT itself is the automation of the full translation process.'

Human translation is the replacement of a text in a source natural language by a text (supposed semantically and functionally equivalent) in the target natural language, and is usually defined as a 'cognitive' operation performed by the translator. For Holmes (1988: 84), an abstract form (or mental concept) is conceived in the translator's mind, and then transferred in order to be reformulated in the target language. The abstraction process is influenced by contextual, intertextual, situational and individual factors. This abstraction-transfer-reformulation process is said to be uniquely translational. On the other hand, current translation software packages usually analyse the sentences of the source text as separated autonomous units, assign a structure to each of them, and use information about the role of the source language words in this structure to determine their equivalents in the target language, as well as the role of these words in the corresponding target sentence⁹⁷. In case of MT, the target language is artificial, not natural.

"Artificial languages are products of human design, based on and derived from natural languages, which do not have the flexibility and multifunctionality of natural languages. Such languages are usually functionally restricted to the conveyance of information and therefore exclude connotative, emotive, aesthetic and other meanings" (Sager, 1993: 258; 259).

⁹⁶ As it is called in French and Russian, for example.

⁹⁷ Lehreberger, J., Bourbeau, L. (1988).

The very use of the term 'language' is in itself misleading⁹⁸, since one may misjudge the output of a MT system (said to be 'full of mistakes'), when the output is syntactically conform to the grammar of the artificial language so defined, but not to the natural language in which the human translation would be expressed. Sager (1994) stresses this fundamental (but too often overlooked) difference between human and machine translation. Both deal with natural languages at the input level, but the output of machine translation is written in an artificial language that resembles strongly another natural language.

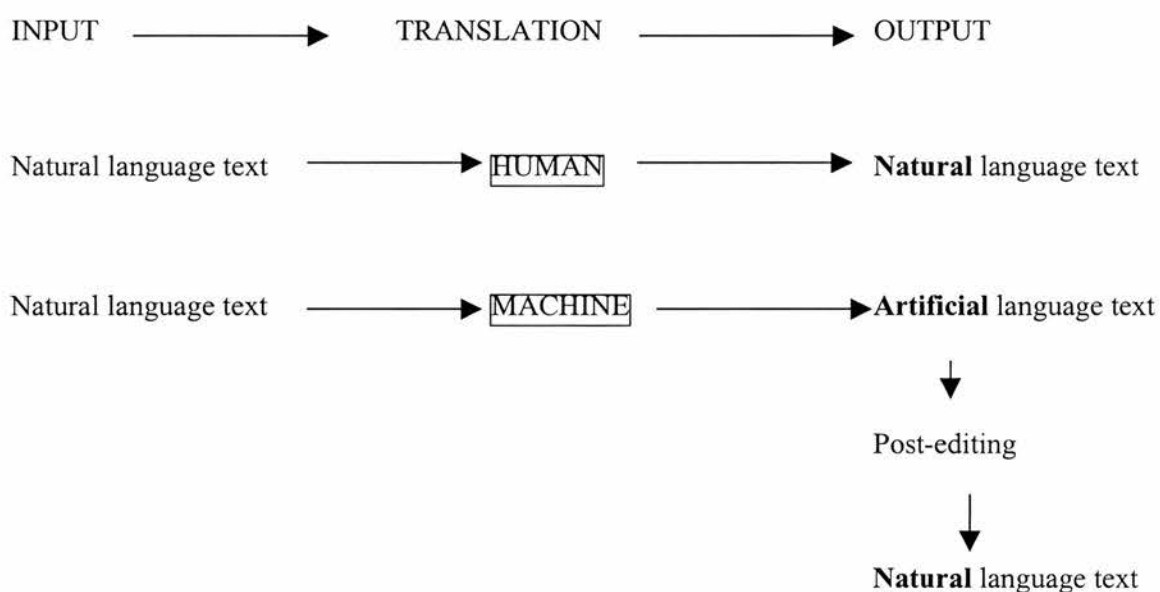


Figure 4-1: Language difference between human and machine translation (Sager, 1993: 258)

FAMT might be conceivable (in a far future, though) by a better analysis of the current shortfalls of most software packages, and by a better understanding of the processes involved in translation (which is far from being a simple decoding-recoding operation). Some problems of MT are inherited from historical mistakes (direct transfer systems based on wrong assumptions

⁹⁸ "Language is a **purely human** and non-instinctive method of communicating ideas, emotions and

about translation⁹⁹) or from structural requirements (most computers are devised along the Von Neumann machines pattern). Others derive from the difficulties encountered in parsing natural languages that cannot be processed as formal languages despite some common features between natural and formal languages. Last but not least, the blind application of some general linguistic rules failed to account for special features of natural languages and of the translation process. A few of these problems will be presented in the following paragraphs in an attempt to identify the main issues at stake.

4.1.3 Fully Automatic Machine Translation

As early as 1955, Erwin Reifler defined the (idealised) MT as “an automatic system which, on the input side, swallows messages in their conventional graphic form and, on the delivery side, spews out these messages in one of the possible conventional forms allowed by the target language for which the system is built” (Reifler, 1955: 144). Twenty years later, R. C. Schank claimed “MT programmes were designed to accept a text in one language as input and produce as output a text in another language that has the same meaning.” For R. C. Schank, mechanical translation is “the *problem* of translating and parsing sentences in one language into some kind of interlingual structure, and then translating back into another language.” Interestingly enough, R. C. Schank did not see MT as something done, but as a field of *research* in order to solve a *problem*. MT as defined by Schank has yet to be born. The systems devised and commercialised correspond more to the definition of R. Reifler. They produce “messages in one of the possible conventional forms allowed by the target language for which the system is built.” In most cases, they transform an SL text into a TL text by applying syntactic and lexical operations on each sentence of the source text (following the pattern of the ‘decoding’ principle). Moreover, most of the systems are bilingual (paired with English) despite extensions to (so-called) new ‘pairs’

desires by means of a system of voluntarily produced symbols" (Sapir, 1921: 8)

⁹⁹ Warren Weaver's approach to translation, which imprinted all the initial research on MT, was based on the cryptology principle (breaking of a code)

of languages, which are always built around the previous SL or TL. Truly multilingual systems based on a language-independent pattern are very rare. And machine translation seems to be (still) an unsolved problem, as soon as one expects high quality fully automatic translation of any input text, despite the numerous software packages available on the market.

4.1.4 Human-aided machine translation

In the case of human-aided machine translation, the machine does the bulk of the translation, and the translator provides any pre-editing or post-editing as needed. Pre-editing consists in adapting the source text to the system, for example by suppressing all pronouns, complex syntactic structures, or ambiguities. The use of pre-editing makes up for the defects of the current translation systems, but limits the development of proper fully-automatic machine translation¹⁰⁰.

4.1.5 Machine-aided human translation

Machine-aided translation (also called computer-aided translation) received various (sometimes conflicting) definitions¹⁰¹ over the last decades, but the term is nowadays understood to cover all kinds of software systems especially designed and developed for use as part of a translator's workstation, but not themselves performing the task of translation as such¹⁰². These systems are for example word-processors, spelling and grammar correctors, terminology databases.

¹⁰⁰"If machine translation is to be viable, the computer must be able to accept any reasonable keypunched text in the source language and to produce a translation without the need for a human being at the interface between the text and the computer." (McDonald, 1979: 94)

¹⁰¹ 'Systems which actually perform the task of translation but rely on the intervention of the human translator at various stages in the translation process' (Blatt et al., 1985: 76).

'A translation strategy whereby translators use computer programmes to perform part of the process of translation' (Sager, 1994: 326).

¹⁰² 'Machine-aided translation occurs in any situation where a machine-readable source text is processed by computerized tools in order to produce a target-language translation, with the translator being in

4.2 Natural Language Processing

4.2.1 Natural and artificial languages

The long-standing lack of consensus between translation scholars and computer linguists may be partially ascribable to the difference between their respective object of study. Human translation is concerned with natural language, while the language produced by translation software is a form of 'artificial language', defined in the BSI Glossary of Documentation Terms as "a language whose rules are explicitly established prior to its use"¹⁰³. According to Sager, "the fundamental difference between human and machine translation is the nature of the language in which they are written, irrespective of quality"(Sager, J., 1997: 36).

John Lyons defines a natural language as "one that has not been specially constructed, whether for general or specific purposes, and is acquired by its users without special instruction as a normal part of the process of maturation and socialization" (Lyons, 1991: 29). He further opposes natural and non-natural languages, defining a non-natural language as any language-system, which is (in whole or in part) the product of human construction. Non-natural languages include but are not limited to "artificial" or "constructed" languages. Artificial languages (Lyons, 1991: 69) are a sub-class of these non-natural languages and they include all kinds of formal languages constructed by mathematicians, logicians and computer scientists (the propositional calculus, predicate calculus, etc.).

Natural languages are characterized by a high freedom of variation at all levels of articulation, since there is no external codification, which can be used as a point of reference. Natural languages evolve under the influence of their speakers, new words are constantly created and even the syntax changes, as in the cases of the use of the subjunctive mode in French. Artificial

control of all stages of this process and performing the intellectual process of translation.' (Somers, in Baker, 2001: 134).

¹⁰³ BSI Glossary of Documentation Terms, 5408-1976, London: British Standard Institute.

languages, on the other hand, have a fixed repertoire of lexical items and associated meanings and rules for their formation and combination. Computers cannot process natural language directly, but only by means of instructions written in the 'artificial' computer language (Sager, 1993: 32-34). The success of a direct machine translation system such as TAUM-Meteo is due to the fact that the source language is itself an artificial language. It was created to improve the communication between weather forecast experts and sailors and pilots. It is based on a code, in order to describe accurately and easily weather phenomena (kinds, directions and force of winds, kinds of rains, force of gales and storms)¹⁰⁴.

4.2.2 Formal languages

Natural languages are usually distinguished from formal languages on the grounds that the grammars of natural languages are described after the usage of these languages have been established, while the grammars of formal languages are developed from a theory, and these languages derive from the theory and the grammar built as a prerequisite. Nevertheless both kinds of languages are made of 'sentences', whose basic constituents are elements, usually called 'words'. But 'words' of formal languages are symbols (logical or mathematical), which have (originally) no 'meaning', and can be interpreted in many different ways (they can theoretically take any 'meaning'). A formal language is defined only by the syntactic order, which states which symbols belong to that language, and which rules will generate the strings of symbols. The way interpretations are assigned to these symbols is not defined by the language itself. The symbols acquire a meaning within the interpretation model chosen for each formal language. On

¹⁰⁴ Example of a weather forecast automatically translated with the use of TAUM-Meteo:

English: LOWER ST JOHN VALLEY UPPER ST JOHN RIVER WIND WARNING ENDED BOTH REGIONS. SNOW AND BLOWING SNOW TONIGHT BECOMING INTERMITTENT NEAR DAWN FRIDAY.CLOUDY WITH PERIODS OF LIGHT SNOW FRIDAY. STRONG GUSTY NORTHEASTERLY WINDS TONIGHT BECOMING NORTHWESTERLY WINDS FRIDAY AFTERNOON.

French translation: VALLEE DU BAS ST JEAN HAUT ST JEAN FIN DE L'AVIS DE VENT POUR LES DEUX REGIONS. CETTE NUIT NEIGE ET POUDERIE DEVENANT PASSAGERES

the contrary, words of natural languages have a well determined 'meaning' ('signification' in French), or several 'meanings', but still determined a priori.

A *formal system* (finite or infinite) consists of:

- (a) a set of primitive entities;
- (b) a set of statements about these entities, called 'axioms', and
- (c) whatever further statements that can logically be derived from the axioms.¹⁰⁵

Similarly, a *formal language* consists of:

- (a) an alphabet;
- (b) a set of formation rules that determine the format of strings of symbols constructed on that alphabet constituting legal expressions in the language, and
- (c) a set of transformation rules for such expressions ('algorithm').¹⁰⁶

Grammars are systems based on finite sets of derivation rules that allow for the generation of all sentences of a language. From a mathematical viewpoint, a grammar is a formal system defined as a set of rules that transform a string of elements (input) into another string of elements (output). It is formally defined as follows:

$G = (V_N, V_T, P, S)$, where

V_N stands for non terminal vocabulary (elements) of the grammar, also called auxiliary vocabulary,

V_T stands for terminal vocabulary (elements) of the grammar, usually called 'words',

P expresses the production rules of the grammar, and should be read as "is rewritten as", and

S is the initial symbol, and should be read as "sentence".

VENDREDI A L'AUBE.VENDREDI NUAGEUX AVEC FAIBLES CHUTES DE NEIGES PASSAGERES. CETTE NUIT VENTS FORTS DU NORD OUEST VENDREDI APRES MIDI.

¹⁰⁵ Hall Partee, B. 1978: 47.

¹⁰⁶ Weizenbaum, J.1976: 60.

V_N and V_T are disjointed and form together the vocabulary V of the grammar.

A **grammar** is a declarative representation of the syntactic facts about a language and specifies the weak generative capacity (set of sentences that are contained within a language) and the strong generative capacity (structure – or possible structures – to be assigned to each grammatical sentence) of the language. The syntax combines the set of non-terminal symbols with the set of rules governing their combination. In most computational natural language systems, the grammar is encoded as declarative data, whereas the parser is a procedural programme; they are therefore separate and independent components.

The language generated by a grammar G is the set of sentences generated by G . The quality of the grammar used to process the syntactic analysis is very important. Whatever the formalism considered, a relatively simple grammar is usually sufficient to analyse a part of the language or even the whole set of a sublanguage. But if some complex features of the language are omitted, hence creating a discrepancy between the generated language and the natural language, ‘overgeneration’ (where the generated language comprises the natural language) may be induced. The problem does not lie so much in the occurrence of ‘overgeneration’ – ‘ungrammatical’ sentences will simply be parsed (considered as ‘understandable’) – but more in the uneven occurrence of ‘overgeneration’, which should then be as homogeneous as possible.

The relation between natural language (NL) and different kinds of formal languages explains the development and importance of pre-editing, which consists in adapting NL texts so that they can be parsed by a specific algorithm. A formal grammar entitles to generate (or derive) sentences with the help of production rules: from the initial symbol of the grammar (S), production rules are applied until the string of symbols does not contain any non-terminal elements (vocabulary). The notion of ‘grammaticality’ depends on the language generated by the grammar, and is equivalent to the concept of syntactic well-formedness in a language (more specifically a formal language). Sentences in a natural language may be *grammatical* but *meaningless*. Parsers are based on the notion of grammaticality (syntactic well-formedness) but

usually fail to account for the meaningfulness of the sentences parsed or produced (in the case of MT).

Moreover, if formal languages and some sublanguages are completely and adequately described in terms of syntactic rules (generating the sentences), natural languages fail to be as adequately described. According to Chomsky's classification¹⁰⁷, languages may be divided into the following sets:

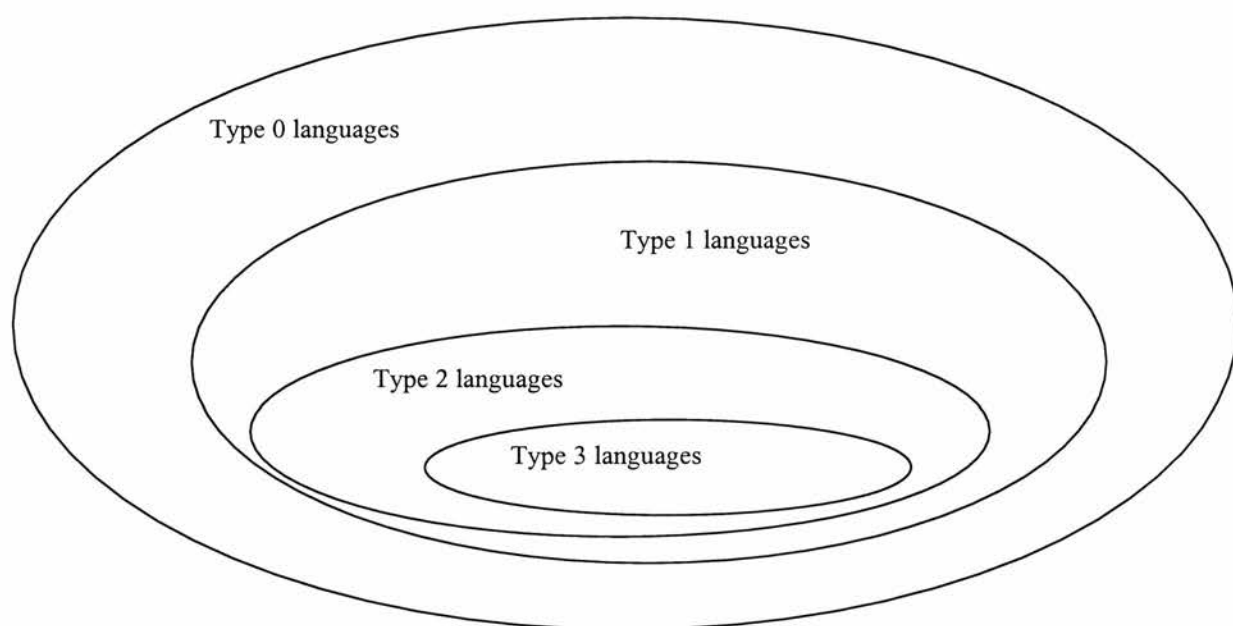


Figure 4-2: Languages' hierarchy

Languages are classified according to the kinds of production rules that generate their elements. Type 3 languages are regular languages (also called finite state languages). In other words, at each derivation step, the string of symbols contains no more than one non-terminal symbol. This

¹⁰⁷ Wehrli, E. 1997: 30.

symbol is either the first (left-linear grammar) or the last (right-linear grammar) of the string. Type 2 languages are context-free languages, and type 1 languages are context-dependent languages. Type 0 languages are languages without restriction. Natural languages are usually considered as a subset of context-dependent languages, also called 'mildly context sensitive', since some of their characteristics place them close to context-free languages. According to Chomsky, context-free grammars, although easily implemented, were inadequate for natural languages parsing¹⁰⁸. Hence the works of Woods (1970) and Winograd (1972) on Augmented Transition Networks (parsing process described as the transition from a start state to a final state in a transition network that corresponds to a grammar of the natural language studied), as well as research on categorial grammar (Montague, in Thomason, 1974; Steedman, 2000 and followers).

Some computational linguists, following Montague (in Thomason, 1974), underlined the similarities between natural and artificial languages, on the ground that syntax and semantics can actually be described and studied within a unique mathematical framework¹⁰⁹.

4.2.3 Syntactic parsing

The main component of natural language processors is the syntactic parser, which function is to determine the syntactic structures of the text (at the sentence level). Once the syntactic structures are determined, the processor computes the different semantic and pragmatic interpretations of each sentence. Syntactic parsing is essential, since the (right or wrong) determination of the structures will lead to a right or wrong interpretation. Any error in the segmentation of the sentence, or in the lexical analysis (cases of homonyms) will affect the upper levels of analysis and processing (more particularly the semantic analysis).

The syntactic analyser attaches to each sentence either a tree structure or a linear one (with parentheses).

¹⁰⁸ For an extended discussion of the in/adequation of context-free grammars for describing natural languages, consult Gazdar, 1982.

¹⁰⁹ Chambréuil, M., et al., 1998: 35-65.

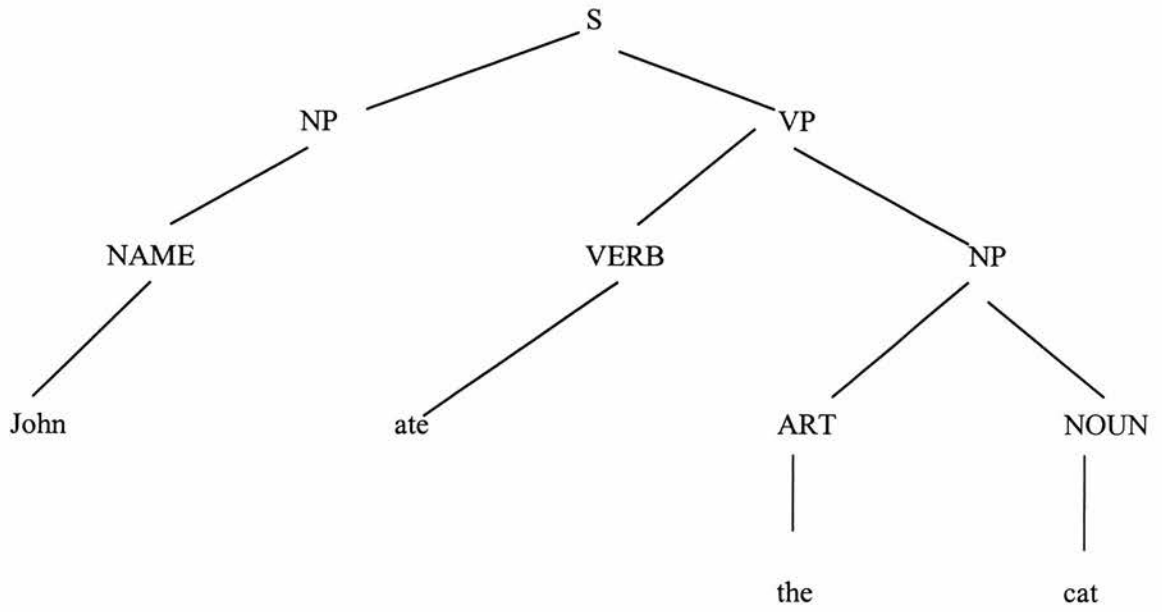


Figure 4-3: A tree representation of *John ate the cat*. (from Allen, 1987: 42)

The linear representation would be [S [NP John] [VP ate [NP the cat]]]

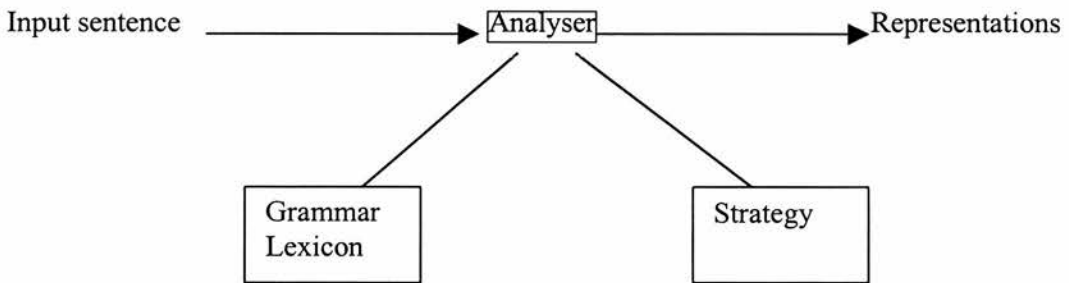


Figure 4-4: Structure of a syntactic analyser (from Wehrli, 1997: 10)

4.2.4 Text typology and sublanguage

The establishment of the source text type (informative, expressive, or operative) is essential in human translation, since it will dictate the strategic choices made by the translator to produce a functionally equivalent target text, as developed in chapter 1.

The development of research on text typology and its application to MT can be explained by two reasons, the obvious failure of parsers to process all possible kinds of syntactic and morphological constructions, and the purposes of MT (only institutions or companies dealing with a huge quantity of documents – usually of a specific type – have ever expressed a demand for FAMT). The recent development of the Internet and the growing need for instantaneous automatic translation of documents has given a new impetus to research on MT and text typology¹¹⁰.

The most accurate translation programmes are devised for highly characterized texts, as exemplified by the TAUM-Meteo and TAUM-Aviation software packages. It has been suggested that MT could be conceivable only for texts with restricted features in terms of semantic domain (domain of discourse/subject field – weather reports, medical reports, engineering manuals), overall discourse type (texts presenting a particular internal format, such as business letters, patent applications...), discourse structure ('descriptive', or 'imperative' sections in a maintenance manual), syntax and morphology (some syntactic constructions, such as the passive, being excluded), or in terms of lexicon (limited number of possible words¹¹¹, limited range of uses and senses of each word). Texts with such limited highly characterized

¹¹⁰ The UMIST developed a sublanguage MT system for the Matsushita Company.

¹¹¹ In the case of weather forecasts, the lexicon is limited to approximately 1,300 words, including morphological variants.

features belong to what has been named a ‘**sublanguage**’¹¹², and defined by Harris as follows: “Certain proper subsets of the sentences of a language may be closed under some or all of the operations defined in the language and thus constitute sublanguages of it.”¹¹³ Kittredge provides a looser definition based on the ‘properties’ of the language: reference to a particular domain of discourse, ‘community’ of speakers, linguistic properties of the subsystem, where no larger system has the same properties.¹¹⁴

One of the most interesting features of sublanguage for MT is that “there is no necessity for the grammar of a sublanguage to bear any interesting resemblance to that of the general language”¹¹⁵. This characteristic allows for the successful development of processors for very specific types of texts, subjected to restrictions (as explained above). On the other hand, successful automatic translation of texts belonging to one sublanguage does not imply a possible generalization of the processing to any text in the same natural language, nor does it allow for the processing of any other sublanguage. But compared with the general ‘failure’ of MT to process ‘any one language’, the success of automatic translation of specific sublanguages raised some hope for the future of MT. Nevertheless, one should keep in mind that one sublanguage may have more common syntactic and morphological features with its ‘equivalent’ sublanguage (for example, aeronautical maintenance manuals in French, English, German...) than with any other sublanguage in the SL, or even with the SL itself. This reason alone could explain the relative success of some translation programmes but should not let us assume that these programmes will be easily extended to other sublanguages, and eventually to any natural language.

Research on particular sublanguages led to the formalization of the grammar generating that specific sublanguage (ex: weather forecast reports, aviation maintenance manuals), but cannot

¹¹² The term ‘sublanguage’ itself is controversial and sometimes wrongly understood by linguists. From the point of view of computational linguistics, ‘sublanguage’ does not simply mean ‘special language of a particular domain’, but ‘the grammar, lexicon, etc., of a particular text-type in a particular domain’.

¹¹³ Harris, Z., 1968.

¹¹⁴ Kittredge, in Nirenburg, S. 1987: 59.

¹¹⁵ Arnold D.1990: 76.

allow for the extension of the parser to other sublanguages or to the natural language 'containing' the sublanguage. For example, aviation maintenance manuals in English present syntactic features that are not representative of English language, and the grammar of this sublanguage is not equivalent to the grammar of English.

'Failure' of MT can be partially ascribed to the difficulty of defining all the rules generating natural languages. Some MT software designers opted for highly regular well defined sublanguages (TAUM-Meteo, for example) and the results are impressively good. Using such programmes for other kinds of texts does not give a proof of the pitfalls of machine translation. It should simply not be done. Just as one would not expect a submarine to fly, one should not expect a programme designed for one sublanguage to process another sublanguage, or more generally the natural language 'containing' that sublanguage. Submarines and aircraft are means of transportation, but their common features (they are made of mechanical parts, powered by engines, and designed for movement) do not entail any 'interchangeability'. Similarly, sublanguages are varied expressions of natural languages, but this quality does not confer on them any equivalence, in terms of grammar and parsing operations.

MT gained a lot – and will gain – from research on text typology, but a few issues still need to be tackled:

- the determination of the specific features of any sublanguage in any language (related to the 'permeability' between a sublanguage and the natural language it belongs to);
- the determination of the common features of corresponding sublanguages in different natural languages.

Computational linguistics failed to devise a syntactic analyser able to process the totality of a natural language (NL), but gave birth to several programmes tailored for specific applications, or devised to study the possibilities of one grammatical formalism or one analysis technique in particular.

4.3 Types of translation systems

The different types of systems developed in MT correspond to the historical evolution of

computational linguistics and machine translation, and are sometimes classified into first-/second-/and third-generation systems. Moreover, the choice of the system retained depends on the purpose of the software (translation of texts between two languages only, or multilingual system), and on the languages concerned. The transfer system is usually preferred in the case of typologically close languages (for example, indo-european languages), since ‘the structures and the lexicon of the source languages usually have equivalents in the target language’, while the interlingua (pivot) system should be preferred in the case of unrelated languages (such as French and Japanese), since ‘only a sufficiently abstract conceptual representation can be common to these two languages.’¹¹⁶

4.3.1 Direct systems

The first translating machines, also called first-generation translation systems, were based on a **direct** (‘transformer’) approach, after Warren Weaver (American mathematician, vice-chairman of the Rockefeller foundation) suggested to apply the decoding principle of the Colossus computer (used by the British to crack the military codes produced by Germany’s Enigma encryption machines): ‘I have a text in front of me which is written in Russian, but I am going to pretend that it is really written in English and that it has been coded in some strange symbols. All I need to do is strip off the code in order to retrieve the information contained in the text.’¹¹⁷ In the direct approach, the morphological analysis is followed by the replacement of the words (through a bilingual dictionary), followed itself by reordering to get the target language output. In these kinds of systems, analysis of the source language is limited to the disambiguation

¹¹⁶ Loffler-Laurian, 1996: 41.

¹¹⁷ The first programme following that principle was devised at the University of Georgetown, U.S.A., on the 7th of January 1954. 1956 marked the birth of Artificial Intelligence, with the Symposium on Information Theory at the MIT, and the Dartmouth meeting in which the term ‘Artificial Intelligence’ was officially acknowledged.

necessary for one given target language. The most well-known direct system is SYSTRAN¹¹⁸, initially designed for Russian-English translation, subsequently adapted for English-French translation. The system has been considerably improved, but the translation process depends largely on the SL-TL dictionaries. More language pairs have been designed, including Spanish, Italian, Portuguese, German... . Limited Semantics Dictionaries added progressively to the system are supposed to solve the problems of homonyms and polysemes (homographic resolution). Not surprisingly, direct systems – although cost-effective and rather easy to implement – render poor-quality translations. Another successful direct system is the TAUM-Meteo¹¹⁹ project, which was implemented at the request of the translators themselves.

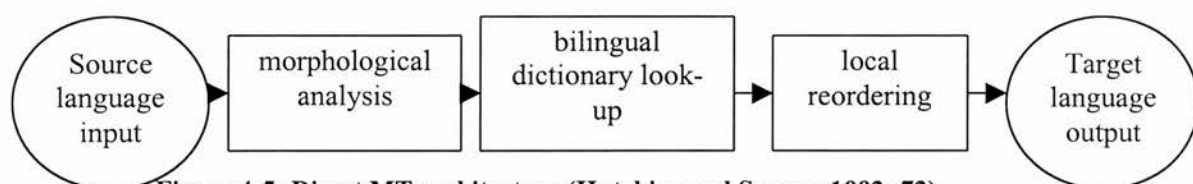


Figure 4-5: Direct MT architecture (Hutchins and Somers 1992: 72)

4.3.2 Transfer systems

The second main type of translating machines, or second-generation translation systems, to be developed were **transfer** systems designed as multilingual systems¹²⁰. One example of such a system is Eurotra (first developed in 1978), in which the source text is analysed (use of source language grammatical and lexical databases) and ‘representations’ of the ST are processed through a transfer mechanism. These representations are themselves transferred into representations of the TL (use of SL-TL dictionaries and transfer rules), which allow for the synthesis of the TT (use of TL dictionaries and grammars). The transfer component in Eurotra is

¹¹⁸ For SYStem TRANslation, developed by P. Tome in 1964.

¹¹⁹ TAUM = Traduction automatique de l’Université de Montréal.

¹²⁰ « In the transfer strategy, a source language (SL) sentence is first parsed into an abstract internal (usually some sort of annotated structure) representation. Thereafter a ‘transfer’ is made at both the

deliberately kept as small as possible and basically reduced to lexical transfer. The assumption on which such a system is based lies in the ‘interface structure’, built on a deep semantic representation of the source and target texts (use of Fillmore’s Case Grammar, for Eurotra). The system is totally modular (separation of analysis and synthesis, separation of morphology, syntax and semantics). The reader should keep in mind that Eurotra was never meant to produce Fully Automatic High Quality Machine Translation, but was supposed to require post-editing.

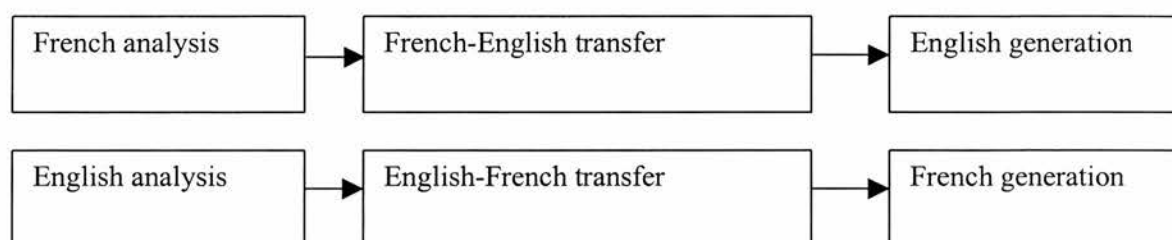


Figure 4-6: Transfer MT architecture (Hutchins and Somers, 1992: 75)

4.3.3 Interlingua

Both direct and transfer systems exemplify a too limited approach to translation. It was assumed that a complete description of SL and TL grammars and exhaustive dictionaries would be sufficient to ‘decode’ a SL text and ‘recode’ it in a TL. These kinds of systems being based on independent operations, they fail to account for the complexity of translation, as underlined by J. Weizenbaum: “Translation must be seen as a process involving two distinct but not quite separable components: the text to be translated has to be understood; and the target-language text has to be produced.”¹²¹ Moreover, in some cases, the first apparently reasonably good output translations were ascribable to convergent lexical and syntactic features between the two languages considered in the translation process. Many difficulties appeared at the stage of

lexical and structural levels into corresponding structures in the target language (TL). In the third stage, the translation is generated. » (Tucker A.B., 1987: 23)

¹²¹ Weizenbaum J., 1976: 186.

extension to other languages belonging to different families of languages (higher morphological and lexical discrepancies).

The third main type of automatic translation (third-generation translation systems) that was considered in the 70s was the **interlingua** approach, based on the use of an intermediate universal artificial language. An interlingual model builds text representations at a level higher than the sentence level, so that some discourse phenomena can be tackled. Such a processing is impossible with a transfer system, which computes sentence after sentence, each independently of the others. An interlingua system would be therefore expected to produce a better translation, more homogeneous and closer to the 'sense' of the source text, even if it is less "faithful", or not as close to the source text form. The main difficulty lies in the fact that current research results do not allow for the description of an operational interlingual language (Loffer-Laurian, A.M. 1996: 42).

Although highly promising in theory¹²², this interlingua strategy has been abandoned for a while on the ground of unfeasibility, but there seems to be a renewal of interest in such a system at the moment, with the application of results in AI research (knowledge representation). The following process is applied: "First, the source text is analyzed and mapped into a language-free conceptual representation. Inference mechanisms then apply contextual world-knowledge to augment the representation in various ways, adding information about items that were only implicit in the text. Finally, a natural language generator maps appropriate sections of the free-language representation into the target language."¹²³ This system is supposed to express the semantic content of the text to be translated, and to allow for the translation of any language into any other one. According to Schank, an interlingua "must represent all the information that is outside a language, that is implicit in a language."¹²⁴ Despite the current renewal of interest in interlingua, the interlingua is still at the stage of fundamental research. So far, all translating

¹²² « The (syntactic) structure obtained after parsing an SL text was declared universal (interlingual) and was supposed to be used directly by the generator », (Tucker A.B., 1987: 24).

¹²³ Carbonell et al., 1981: 376.

¹²⁴ Schank, R.C., 1975:9.

systems produced have been based on distinct operations, while the interlingua approach requires a complex parallel conjunction of different processes. Moreover, the representation of a text for an interlingua language is based on the field knowledge of the given source text. Such a representation is not easily transferable/applicable to a text belonging to a different field. In that respect, a transfer system is more easily applicable to different kinds of texts, since its analysis is more linguistic and less cognitive.

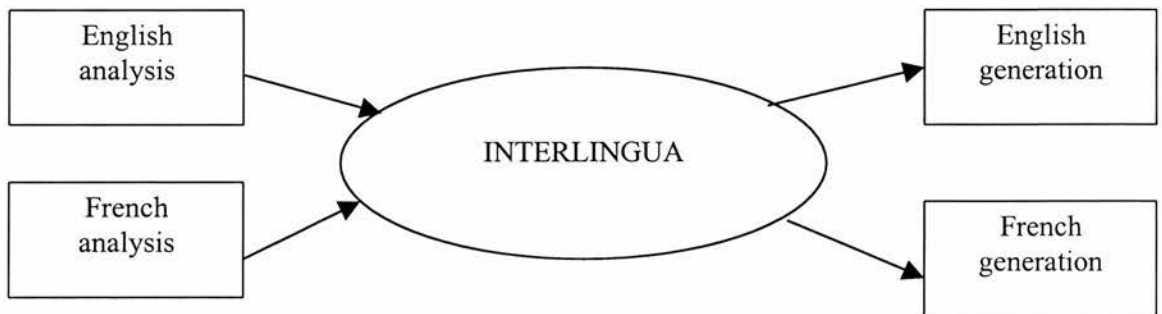


Figure 4-7: Interlingua MT architecture (Hutchins & Somers 1992: 74)

4.4 Issues in machine translation

4.4.1 Segmentation and the unit of translation in MT

Machine translation (MT), an application of Natural Language Processing (NLP), is highly dependent on syntactic parsing of sentences. Words or groups of words must therefore be isolated. More particularly, segmentation is necessary in any multi-pass system where the processor moves freely back and forth through the text. The main problem lies in the fact that MT systems (based on structure-bound translation) tend to translate at the level of word or lexeme¹²⁵, while human translators use higher-level units. Moreover, there is no single concept

¹²⁵ Word considered as a lexical unit, in abstraction from the specific forms it takes in specific constructions (*run, ran, running*)

of 'unit of translation' (UT) that would correspond to a linguistic unit. Vinay and Darbelnet define the UT as "the smallest segment of the utterance where the cohesion of signs is such that they cannot be translated separately"¹²⁶. UTs may correspond to parts of words (sea/shore), to single words, or to more than one word. Compound forms raise particular difficulties, since the meaning cannot always be inferred from the meanings of their components, through a compositional rule. Moreover a letter or a letter sequence can be part of the preceding as well as the following constituent of a compound (as in Russian, where <rybolovu> may mean <to a fisherman>, <rybo/lovu>, or <to the tin of fishes>, ryb/olovu>). The best solution consists in inserting all possible compound forms in the machine memory, but at the expense of loading that memory. Since text cohesion and rhetorical structure play a very important role in human translation, one could be tempted to devise programmes processing larger UTs, but there is no guarantee of improving the quality of MT by adopting a larger unit of translation¹²⁷. This is due to the architecture of the systems, which actually do not process natural languages as a translator would. Much failure of MT could be ascribed to the mechanical and structural limits of the systems allowing only for the processing of independent sentences. Prediction, for example, plays a very important role in human analysis, which proceeds by making assumptions about what the meaning of the sentence is going to be as it is being input. Multi-pass systems allow for forward and backtracking, but the level of ambiguity remains very high. A human translator – through prediction – rates one sense of any single word above the others and subsequently corrects – if necessary – his/her assumption. A parser (that is supposed to extract and identify the conceptualisations that underlie the input) has at its disposal the immediate words or segments of sentences, which are usually insufficient to effect a provisional disambiguation.

4.4.2 Ambiguity

¹²⁶ Vinay, J.P. & J. Darbelnet, 1958: 16.

¹²⁷ Bennett, P., 1994: 12-20.

Ambiguities are of two kinds: lexical and structural. Lexical ambiguities cover polysemy (a single word having two or more related senses) and homography (forms belonging to different grammatical categories, but sharing a common spelling, e.g. to *round* vs. a *round* table). If the latter can often be disambiguated by syntactic parsing (a verb vs. a noun), the former is more often disambiguated by the use of subject fields ('bank' as a financial institution vs. 'bank' as a riverside), or the assignment of semantic features ('human', 'animate subject', 'male', 'female', and so on). Despite the increasing use of world-knowledge data in the disambiguation process, computers are still unable to solve some common ambiguities¹²⁸. Structural ambiguities cover syntactic structures and representations of sentences. Quite often, MT systems are unable to assign the correct grammatical categories, and to analyse the sentence completely¹²⁹; in which case, they resort to a word-by-word translation and copy the structure of the source sentence.

4.4.3 Anaphora resolution

'Anaphora' refers to the relation between a pronoun and another element, in the same or in an earlier sentence, which supplies its referent. The resolution of pronoun references is especially important in the case of languages (such as German) that mark the gender of the pronoun¹³⁰. Moreover, current systems usually operate at the sentence level and are unable to determine cross-sentence connections.

¹²⁸ As shown by Bar-Hillel's example 'the box is in the pen'. The introduction of the world knowledge distinguishing 'pen' (a playpen or a pigpen) from 'pen' (a writing instrument) has proved to be very difficult. (Austermuehl, F., 2001: 171)

¹²⁹ The detached prefixes of many verbs in German (aus/auf/weg/hinaus/heraus/fort) are for example extremely problematic.

¹³⁰ 'The monkey ate the banana because *it* was hungry'. 'It' referring to the monkey will be translated by 'er' (= der Affe) in German.

'The monkey ate the banana because *it* was ripe'. 'It' referring to the banana will be translated by 'sie' (= die Banana). (Hutchins and Somers, 1992: 95).

4.4.4 Idioms

In translation, idioms¹³¹ have to be treated as single units of translation. The methodology usually used in MT systems consists in including them in the system's dictionary, but most systems tend to prevent any translation error by resorting to pre-editing, and the suppression of all idioms.

4.4.5 Syntax/semantic separation

The main focus of (as well as the most controversial issue in) NLP is the semantic representation of the text. This is sometimes called in computational linguistics 'target representation'. This phrasing will not be used in this dissertation, in order to avoid any confusion with 'Source Language-Target Language' used in translation studies. One of the steps towards this interpretation is traditionally supposed to be the syntactic analysis (taking place before the semantic analysis), but some computational linguists (Schank, 1975) argued for a direct semantic interpretation. Actually, all the systems based on that theory include syntactic parsing (of some sort) within the semantic analysis, on an ad-hoc basis, the semantic and syntactic analyses being intertwined. Even though the pioneers in machine translation were aware of the fact that human translators do not seem to 'parse syntactically' before disambiguating and interpreting, the mechanical requirements of the machines conceived to translate automatically are such that successive hierarchical operations are still needed.

E. Reifler wrote in 1955 : « Experience and situational criteria often enable a human translator to grasp intuitively the semantic content of a foreign text whose grammatical problems he does not fully understand. (...) The human translator need not adhere to any fixed sequence of determinative evaluative procedures. This is, however, not yet feasible in MT. Here a *hierarchy*

¹³¹ 'Idioms are expressions whose meaning cannot be completely understood from the meanings of its components.' (Arnold et al., 1994: 122).

of determinative operations is necessary, at least for the present. (...) A translation mechanism will, within the limitations of its design, first determine the grammatical situation of each source form and then, on the basis of the grammatical situation, proceed to the determination of its intended nongrammatical meaning and the supply of the appropriate output equivalent» (Reifler, 1955: 148-149). Machine translation did not seem to have evolved much since 1955, despite Reifler's optimism. The last sentence could still be written today to describe most of the available translation software.

Computational linguists may be divided into two main 'schools', one in favour of independent syntactic and semantic parsing (easier to implement), one in favour of an integrated approach (as exemplified by formalisms such as GPSG – Generalised Phrase Structure Grammar – and Categorical grammar). Fodor and others, for example, argue that syntactic analysis is an independent step that logically takes place before the semantic analysis. Their view is mainly derived from Chomsky's theory (1965) according to which it is possible to have a purely structural theory of language, independent of supporting cognitive mechanisms such as memory and thinking. Proponents of the integrated approach (Schank, 1972, 1975) argue that semantic analysis is used throughout the language understanding process, and not only following syntactic analysis.

The separation between syntactic and semantic parsing is partly ascribable to the historical development of MT (direct systems and transfer systems) as well as to the structural requirements of the first processors (sequential operations). But research in cognitive science done over the past twenty years tends to prove that human beings *do not* perform a syntactic parse first. There is actually very little evidence for the use of standard syntactic notions in understanding. According to some computational linguists, human beings rely heavily on semantic prediction for language understanding, and syntax is seen rather as a tool helping us to find things we are looking for¹³².

¹³² "Our meaning representation tells us what kinds of things to look for and, when we can't decide which one is meant, we use syntactic notions to help us figure out what was really meant." (Schank, R.C., 1975:18).

Another argument in favour of semantic parsing rather than syntactic parsing lies in the evidence that people very rarely remember sentences in the original word encoding. *Meaning structures* rather than syntactic structures are remembered (basis of interpretative theories of translation). It is therefore widely held that, if MT is ever to get close to human translation, it would be through a semantic parsing. Current parsers tend to mark ambiguous representations for each sentence, and then select the most likely representation according to the information stored ('world knowledge'). An even better system consists in selecting preferred meaning representations (based, for example, on expectation) and applying syntactic parsing to check the adequacy of the representation selected.

Some computational linguists argue for an integration of syntactic and semantic processing, taking place at the same time¹³³. Their arguments are twofold:

- inability to resolve syntactic ambiguities using only local syntactic checks (sentences are processed independently)¹³⁴ ;
- build-up of syntactic ambiguities when syntactic and semantic processing are separated. Several syntactic attachments must be considered, hence resulting in a kind of 'combinatorial explosion' in some cases.

Separate syntactic and semantic processing tends to be cheaper, and easier to implement. The main advantage of such a system lies in the fact that it improves the efficiency of the parsing by an efficient syntactic analysis. On the other hand, intertwined parsers (presenting different options of interaction between the syntactic and the semantic processing) would be closer to human understanding but they tend to be less efficient since syntactically impossible semantic interpretations may be built (and discarded later, but consuming time, memory and computing capacity). Recent technologies, such as backtracking, render such intertwined syntactic/semantic parsing more feasible. But semantic parsing is far from being solved. Some attempts have been

¹³³ Lytinen, S.L., 1987: 302-316.

¹³⁴ Lehrberger and Bourbeau (1988) are very sceptical about any claim that adequate parsing can be accomplished by means of local analysis, since even the identification of constituents that perform given functions (subject, object, etc) requires syntactic analysis of the entire sentence.

made to integrate processing on a 'semantic level' on the basis of Schank's conceptual dependency theory, fragments of semantic and distributional information being associated with lexical entries, and a global representation being built up progressively (conceptual dependency network). However, all the so-called 'semantic grammars' and 'semantic processing' programmes use syntactic classifications as well as syntactic rules to build up the structures at some stage. Therefore, it seems impossible to consider syntax and semantics as separate entities computable in isolation, as is the case in most software (mapping of a syntactic representation of a whole sentence onto a representation of its meaning). According to E. Rich (1991), there are arguments in favour of each approach, either immediate application of semantic interpretation every time a syntactic constituent is formed, or interpretation of the whole (syntactically parsed) sentence.

The so-called 'third-generation systems' (interlingua principle), currently at the stage of fundamental research, use techniques from the field of AI, and rely more on semantics, including inferencing both within and beyond the sentence boundary¹³⁵. In this respect, the MT industry would greatly benefit from a better understanding of the cognitive processes involved in translation.

4.4.6 Machine translation evaluation

The major factors in evaluating MT are cost, time, quality of machine output, improvability of the system, possible extension of the system (to related domains, unrelated domains, language pairs), and facility of use. Nevertheless, such an evaluation is problematic, considering the fact that most systems usually require pre- and/or post-editing, in which case not only the actual translation software package, but also the human translators and operators are tested, as well as the additional computer tools coupled with the translation programme (on-line dictionaries, word-processors, spell checkers...).

¹³⁵ Lehrberger & Bourbeau, 1988.

Results may vary drastically according to whether 'raw' texts are processed, or pre-edited ones. Similarly, the assessment of the linguistic quality of the output will depend on more or less post-editing. One central issue in evaluation is the selection of appropriate texts. After the failure of most MT programmes in the 70s, research was re-oriented towards text typology. Texts with a high frequency of incidence of a limited number of linguistic phenomena were – not surprisingly – easier to process. There is no doubt that TAUM-Meteo is more efficient and cost-effective than human translators, but weather forecast bulletins present specific linguistic features (statements, reduced verbal forms, reduced number of articles, extensive use of codes for types of clouds and strength of wind...) which render them ideally 'processable'. Nevertheless, they do not represent examples of common language. The quality of the translations of TAUM-Meteo (and then of TAUM-Aviation) does not imply that the system is able to translate accurately any random text. Therefore the difficult question of overall assessment remains. Results may vary drastically depending on whether the programme is submitted with highly specific types of texts, or with texts comprising as complete a range of linguistic phenomena as possible.

Once the type of texts to be translated has been determined, system evaluation is usually made through linguistic assessment (mainly error analysis, but also 'intelligibility', 'comprehensibility', 'style') by asking the human post-editor to count the corrections he has to implement (word addition/deletion/substitution/transposition). Some assessments are not based on the result expected by a human translator, but aimed at comparing the actual output with the output expected by the programmer. Some automatic translations might therefore be judged 'wrong' from a human translator viewpoint, but 'right' if they match the system's design goals. At the current state of advancement of machine translation, assessment based on 100% accuracy expectancy (in terms of errors, intelligibility, comprehensibility, style) would discard all the software if they were tested with texts containing the widest range of linguistic phenomena. The assessment of the data as presented hereafter (chapter 6) does not aim at rating the accuracy of the translation software, but rather at pinpointing main areas of research to be developed in order to improve meaning representation and translation. A similar assessment of human

translation (with an expectation of 100% accuracy) would also discard most translations. There is no such thing as a 'perfect' translation, but one has to define the criteria for a 'good enough machine translation', or 'acceptable' translation.

4.5 Meaning in MT

4.5.1 Introduction

The great majority of research done in NLP is based on the assumption that it is possible to compute a literal meaning (which belongs to semantics) and then to interpret it thanks to the knowledge of the reference world, to the context, and to the speakers involved (this interpretation process belongs to pragmatics). In NLP, meaning is computed from the following three aspects of meaning:

- lexical aspects (the problem lies in establishing the links between words and their significations);
- syntactic aspects (words take place in syntactic structures, and the problem lies in establishing the meanings carried by these structures);
- pragmatic aspects (determination of the world knowledge and situation contributing to the determination of meaning).

This separation of meaning into different entities correspond to the 'non-contextual meaning'/'contextual meaning' dichotomy, or to the difference in French between '*signification*' and '*sens*'.

The data (texts) are constituted of a form on the one hand (material support, letters forming words, forming sentences, forming texts), and of a meaning on the other hand. The problem lies in the fact that AI aims at formalising the meaning through a syntactic processing (form), but this very focus on the form excludes the meaning. Moreover, it has not been proved that

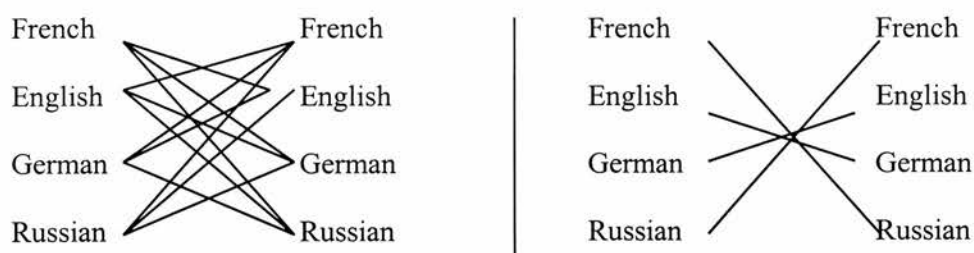
languages are formalisable, and it seems that human beings do not process languages in a formal way.

At the dawn of MT, Erwin Reifler (1955) considered that “the first concern of MT must always be the highest possible degree of source-target *semantic agreement* and intelligibility. The MT linguist, therefore, must study the languages that are to be mechanically correlated in the light of source-target semantics.”¹³⁶ Later, Sergei Nirenburg also pointed at the central issue of meaning in machine translation, as much as in human translation.

”The computer must be able to obtain as input a text in one language (SL, for source language) and produce as output a text in another language (TL, for target language), so that the meaning of the TL text is the same as that of the SL. It is clear that finding a way of maintaining invariance of meaning is the crucial *problem* in MT research.”
(Nirenburg, 1987: 2)

Despite this acknowledgement shared by all computer linguists, most machine translation systems still rely on a ‘transfer’, which simply shortcuts the meaning related parts of the process, triggering several problems, such as the resolution of ambiguities, or the necessity to build additional transfer modules, ($n(n-1)$ transfer modules are necessary for n languages), hence increasing the cost of the system, every time a new language is added.

In theory, the use of an intermediate (‘pivot’) language would allow to avoid the problem. Moreover, if the intermediate language is a workable semantic representation language, the meaning can be represented and processed by the system, which more closely mimics the human translation process model.



¹³⁶ Reifler E., 1955: 138.

Transfer system

Pivot system

Figure 4-8: Transfer modules respectively in a transfer and in a pivot system (from Everaert et al.)

Several projects on an interlingua module have been launched (GENESE, EUROTRA, among others), but so far none has been successfully implemented.

Natural language processing implies three problems: simulation of the thought process (function of the quantity and quality of the information we have about a subject, as well as the reasoning capacity), understanding of the linguistic input (representation of the *meaning* of the linguistic input) and simulation of the world knowledge used in the understanding phase.

Two main approaches have been put forward for the modelling of cognitive functions:

- the functional modular approach, inspired from Fodor, assumes that the brain system can be adequately modelled by separate specialized subsystems, which process individually parts of the information ;
- the connectionist approach, assumes that the brain can be best modelled using connectionist models as neural networks. These networks mimic (in a very simplified way) the neuronal organisation in the brain.

Nevertheless, many translation programmes transform a text from one language into another language without any semantic representation (direct approach).

4.5.2 Meaning and knowledge representation

Since the early 70s, the major issue in MT has been the creation of a meaning representation that relates the concepts underlying a sentence.

On the other hand the representation of meaning has led to various controversies over the last 30 years, mainly between word-based representation advocates, and interlingua advocates. The latter developed several theories about meaning representation through the use of primitives¹³⁷ (as in R. Schank's conceptual dependency theory¹³⁸), frames¹³⁹ (as developed in Minsky, 1975), scripts¹⁴⁰, plans and goals. Frame and script systems allow for a representation of expectations, hence their use in NLP (reduction of the number of possible sentence structures). Whatever the model developed, the word-independent theory is the most widely held view within computer linguists, who base their approach on the observation that human beings usually do not have a word-based representation. When told a story, they can easily re-tell it, using a different phrasing and different words, but they keep the gist of it, the meaning. Most cognitive theoreticians argue for a language of thought that is language-independent (interlingua), and several programmes (called parsers) have been developed, taking natural language as input, and producing some kind of 'meaning representation' as output. In linguistics, parsing refers to the derivation of the grammatical structure of a linguistic input. Schank adopted the term to refer to the derivation of the meaning of a linguistic input, and reflects his integrated approach to semantic/syntactic parsing. For her part, Rich (1991) defines parsing as syntactic analysis ('conversion of a flat input sentence into a hierarchical structure that corresponds to the units of meaning in the sentence'). The common goal of all these programmes is to develop a single conceptualisation that represents the meaning of the input sentence and they are all (more or less) based on the following axioms (particularly useful for MT):

1. "For any two languages that are identical in meaning, regardless of language, there should be only one representation."

¹³⁷ « Primitives are the basic element of knowledge representations in AI programmes and theories. They represent the lowest, most elemental, nondecomposable layer of the representation. Primitives represent knowledge and/or meaning and organize the inferences that can be made on the basis of knowledge » (Shwartz, 1987: 29).

¹³⁸ Conceptual Dependency Semantic representation includes primitives based on notions such as transfer (abstract, physical, or mental), bodily activity (apply force, move a body part, grasp, ingest, expel) or mental actions (conceptualise or think). These primitives are completed by cases (actor, objective, to, from, recipient, directive) and are linked by temporal connectives and the instrumental relation.

¹³⁹ A frame organizes a set of slots ('actor', 'object', 'from', 'to'...) that are filled by some nominal value (a person, animal, place, or object).

¹⁴⁰ Scripts are high-level knowledge structures containing sets of scenes that have a strict ordering.

2. "Any information in a sentence that is implicit must be made explicit in the representation of the meaning of that sentence."¹⁴¹

Much research has been done on knowledge representation¹⁴², particularly on different ways to represent pragmatic knowledge. The main issue deals with the content of representation, rather than the 'envelope' used.

The most important techniques for knowledge representation are predicate logic, semantic networks, production systems, and frames and scripts. Predicate logic uses standard forms of logical symbolism and is associated with techniques for the analysis of many conceptual structures in our common thought; its drawback being its extremely formal kind of knowledge representation. Semantic networks express general facts and assumptions by means of a link, or arc, joining two points, or nodes in a network. Production systems contain a set of production rules, usually of the form IF <condition > THEN <action >; they are particularly useful where the knowledge consists of many loosely related facts and independent actions). Frames and scripts are both conceived with the idea that our knowledge of concepts, events and situations is organized around expectations of key features of those situations. Marvin Minsky defined frames in 1975 as data-structures for representing a stereotype situation, composed of networks of nodes and relations. The top levels of a frame are fixed and represent things that are always true about the supposed situations. The lower levels have many terminals, 'slots' that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet. Scripts are defined as predetermined causal chains of conceptualizations that describe sequence of things in a familiar situation, each script containing a set of scenes that have a strict ordering (Schank, 1975). Both frames and scripts have been judged very useful since, according to Wilks (1979), they express the dynamic order of events that is part of the meaning of certain concepts, in some intuitive sense (Wilks, 1979: 27-43). These techniques for knowledge representation are based on natural language theories, and particularly on different aspects of language, defined by linguists as syntax, semantics, and pragmatics. "First, language is infinitely

¹⁴¹ Schank and Abelson, 1977: 11.

¹⁴² Barr and Feigenbaum, 1981; Wilks, 1978, Winograd, 1972.

productive: human beings are, in principle, able to produce and understand an infinite number of novel sentences. Secondly, we can use language to convey meanings and thoughts to others, to get at the *content* behind the purely physical signals of sound waves or marks on paper. Finally, humans do not merely form meaningful sentences, but can give an 'appropriately meaningful answer to what is said'; that is, they can use language in ways appropriate to the context."¹⁴³

4.5.3 Meaning in MT, semantics and Transformational Grammar

Part of the problem of meaning representation lies in the imprint of transformational grammar on NLP, although the objectives and focus of transformational grammarians and computational linguists are different. The main preoccupation of researchers in Transformational Grammar is to write grammars that should not generate sentences that are either meaningless or ungrammatical, whereas the computational linguist's preoccupation is to explain how a computer could perform as a human would perform with regard to such sentences.

According to R. Schank, "semantics, as defined by transformational grammarians, is inapplicable to the problems of computational linguistics" (Schank, 1975: 7). Moreover, most of research in Artificial Intelligence and Machine Translation is based on the assumption that literal meaning (semantics) can be computed, and then interpreted according to world knowledge (pragmatics)¹⁴⁴. However, this strict separation is a highly debated topic, on the grounds that – as some linguists claim – the interpretation depends on the interpreter, hence the cognitive approach to MT. The distinction between literal meaning and interpreted meaning corresponds to the mechanical distinction between syntactic and semantic parsing. But as Sabah (1997) stresses, if the objective of AI (more particularly MT) is the treatment of meaning, its

¹⁴³ Sharples M. et al., 1989: 130.

¹⁴⁴ Most French computational linguists use the word '*signification*' for literal meaning, and '*sens*' for interpreted (contextual) meaning ('sense'). (Sabah, G,1997).

object of study (language) is processed through a syntactic parsing (material expression of language: written sentences). Hence an apparent vicious circle traps computational linguists: any formalization of meaning implies a syntactic processing based on the form, not on the meaning. Works of Schank, Minsky, Johnson-Laird, for example, converge towards an attempt to build formal representations ('mental' representations, for Johnson-Laird). The problem of computational semantics is twofold: "the finding of a step by step procedure that assigns a meaning to input sentences; and the creation of meaningful ideas and their encoding into the sentences of a language"¹⁴⁵. Schank's requirements are first a language-free meaning representation, and secondly the definition of step-by-step procedures conforming as much as possible to what is known about human behaviour. As far as this last condition is concerned, computational linguists are faced with a difficult problem. It seems that human translators intuitively grasp the semantic content of a ST without resorting to any fixed sequence of determinative evaluative procedures, whereas determinative operations must be hierarchically organised in a programme (mechanical requirement).

Despite much fundamental research in semantics, few results have been successfully implemented in MT. Two main types of approaches to meaning in Natural Language Processing can be defined: one based on formal semantics (classical logic, higher-order logic, typed logic, modal logic, non-monotonous logic, illocutory logic) and one based on different kinds of representation, like semantic networks (Quillian and followers), frames and scripts to model beliefs. Semantic (context-free) grammars combine syntactic, semantic and pragmatic knowledge into a single set of rules in the form of a grammar. Although useful for producing restricted natural language interfaces, semantic grammars fail to capture important linguistic generalizations. Case grammars (Fillmore, 1968) include some semantic information in the structure that is built by the parser, but they tend to provide an incomplete semantic description of a sentence. Conceptual parsing combines syntactic and semantic knowledge into a single interpretation approach (use of strong expectations for particular sentence structures).

¹⁴⁵ Schank, R.C., 1975: 7.

Conceptual parsing is driven by a dictionary that describes the meaning of words as conceptual dependency structures, and requires additional knowledge about the context in which the sentence appears. Compositional semantic interpretation applies semantic processing to the result of the syntactic parse. Compositional semantics (Montague, in Thomason, 1974) is based on the idea that for every step in the syntactic parsing process, there is a corresponding step in semantic interpretation; and the meaning of each sentence constituent is defined as a composition of the meanings of its constituents with the meaning of the rule that is used to create it.

4.5.4 Conclusion

An overview of the current available automatic translation software leads to the conclusion that meaning is absent from MT. Semantic components are added to syntactic parsing, but there is no implemented programme to process at a purely semantic level. Most MT applications are based on improved direct systems (TAUM-Meteo) or transfer systems. The only possible treatment of meaning for MT would imply an interlingua architecture, which does not seem implementable in the near future. Nevertheless, the United Nation programme (UNL) launched in 1996 marks a shift of concern and the birth of a new field of research, characterized by a cognitivist approach to MT.

4.6 Conclusion: how does machine translation relate to translation?

Despite numerous attempts, there seems to be no single reliable theory of translation, for the very reason that translation is a cognitive and social process. Grammars of (artificial and natural) languages may be written, but operations ruling translation cannot be easily modeled (Vinay and Darbelnet's translation model fails to account for the actual processes at stake).

Computer scientists, becoming more aware of the complexity of 'translation', now recognise the limits of programmes such as direct ones, which assumed that translation equalled a decoding-recoding operation. On the other hand, they have now at their disposal highly powerful tools (combinatory logic, fuzzy logic...) opening new approaches to NLP. MT (as one main application of NLP) would greatly benefit from the study of the actual cognitive/subcognitive operations performed by human translators. The following chapters will attempt to show the limits of the current automatic translation tools available, and compare the output of translation software (transfer approach) with the published translation of the same texts. Translation (done by human beings) is a set of operations performed between two texts, one in the source language, the other in the target language. Part of the apparent failure of MT may be ascribable to the wish of computer scientists to consider only the two natural languages involved, and to the assumption of some of them that the input of an extensive dictionary coupled with grammar rules and 'world knowledge' data would be sufficient to render a good translation. Linguists know how complex a process translation is, although no one has ever been able to describe all the operations involved, in such a way that a 'scientific' theory of translation could be developed. If, following Steiner (1998), we take as a starting point that translation is, first and foremost, understanding, we must study the process of understanding, hence the transfer of meaning involved in translation. To do so, it is necessary to take an extract of natural language, a text, translated in different languages. Under the assumption that meaning is the invariant element between a ST and its translations, the comparison of the different target texts (in different languages, preferably not too close syntactically and morphologically) should provide some useful data on the way meaning is generated by languages, as well as on the syntactic and semantic operations performed by the human brain during the translation process.

Much research has been done on specific well-defined languages, and automatic translation of highly technical texts is already fairly successful, even with a direct system (without meaning representation). But until quite recently, it was assumed that automatic translation of 'literary' texts was totally impossible, on the ground of the linguistic complexity of such texts. It is therefore interesting to choose precisely such a piece of literature, displaying fairly good

example of linguistic variation. The selection of a *text* (*Le Petit Prince*, Antoine de Saint-Exupéry) belonging to a well-defined type (Children's literature) is due to several reasons. Not only does *Le Petit Prince* present the features requested for such a study (literary text, displaying a large variety of linguistic features, although containing a rather simple lexicon and syntax), but it also represents an extremely interesting tool for the study of meaning (role of interpretation, the text being understood differently by the potential readers – parents reading the book to their children, or children reading the book themselves). It has sometimes been argued that children's stories provide suitable texts for research on automatic translation, since they are short and seem linguistically simple. However, the simplicity is only apparent. There might be fewer sentences, as well as a reduced lexicon, but some difficulties appear, such as semantic and situational combinations (talking animals...), which would be unsuitable for other types of translation¹⁴⁶. They are therefore a particularly interesting object of study for meaning representation.

The most powerful and successful translation software currently available are transfer systems. The study of the output of such a programme (in English and German), as well as their comparison with the published translation will give insights in the pitfalls of the formalisms used for machine translation. On the other hand, a study of meaning representation derived from Montague Grammar and combinatory logic represents an interesting – although purely theoretical – alternative. An attempt at modelising translation will be valuable in the long term for research in AI, just as translation studies will greatly benefit from research in cognitive sciences. Most translation studies tend to compare linguistic features of SL and TL but leave the actual process of translation aside. 'Operations' are assumed to be performed by the brain, seen as a 'black box'...but there is little data on the so-called 'operations' whether at the syntactic or semantic level. On the other hand, formalization of NL does not necessarily imply rightness of the modelised translation process. The aim of modelising is not to reproduce reality in all its

¹⁴⁶ « The morphology and syntax are probably not significantly altered, but the semantic apparatus, depending on how it is conceived, may have to be considerably different from the semantic apparatus necessary for a suitable translation of a newspaper editorial or a text in biology », (McDonald, 1979: 93).

complexity, but rather to capture what is essential to understand some aspect of its structure or behaviour.

5 APPLICATIVE AND COGNITIVE GRAMMARS AND TRANSLATION

5.1 Introduction

The study of language as expression of ideas/concepts, and the determination of meaning has always been a central issue in philosophy. The Greek tradition (through the works of Plato, Aristotle and the Stoics¹⁴⁷) deeply imprinted the European approach to language, logic, grammar and meaning. The Middle Ages saw the development of the view of language as a system of significance (following the platonic and neo-platonic traditions), and marked the birth of semantics, as the science studying the link between logic and meaning. Most of the research at that time was devoted to the study of the properties of terms, in order to distinguish the different roles played by words or expressions when used as terms within an expression. Grammar (as the expression of thought and a link between form and content) is distinguished from logic (centred on true and false propositions). The acknowledgement of the fact that there was no absolute correspondence between language and thought led to a search for logical structures through the complex grammatical forms, the only tool readily available. In its turn, this triggered two different approaches, one leading to the creation of an artificial language supposed to express precisely the "logical operations" of the mind (this will be Leibnitz's approach¹⁴⁸), the other one developed by the Port-Royal school¹⁴⁹, that considered the form of discourse and the logical operations of the mind as two different entities.

¹⁴⁷ The Stoics used the term 'logic' for the general study of language, including rhetoric and grammar and aimed at preserving the logical structures in agreement with the grammatical structures. They called 'dialectic' what we call now logic.

¹⁴⁸ 'Lingua characteristica universalis', Leibnitz's project, was an attempt to create a 'philosophical language', or 'universal calculus', according to which all logical queries would be answered to through a calculus, i.e. an automatic procedure, from which judgment would be absent.

¹⁴⁹ The 'Port-Royal logic', as it is called today, is based on the work of Arnauld and Pierre Nicole published in 1662.

For the Port-Royal logicians, these mental operations were "externalised" through language, and the purpose of logic was to "extract" meaning through the intricacies of verbal forms. At the heart of the theory developed by Port-Royal scholars, lay the distribution and organisation of grammatical categories, and the proposition (composed of the verb and the name). Discarding variables and focusing on concrete examples taken from the grammatical analysis, the Port-Royal School distinguished itself from formalism. Logic then tended to divide itself into two main approaches: one called classical logic and included in philosophy, mainly centred on syllogistic, the other one developed by mathematicians introducing new methodologies and concerned with formalization. George Boole (1815-1864), Gottlob Frege (1848-1925), and Bertrand Russell (1872-1970) were the main proponents of this new approach. Boole aimed at treating thought (as expressed in our language) as an algebra, and elaborated a parallelism between algebraic syntax and logical syntax ($xy = yx$ /white sheep = sheep white; $x + y = y + x$ /sheep and oxen = oxen and sheep). Frege, considered as the founder of modern logic (as it is currently developed), saw logic as a means rather than an end, and focused his study on propositions rather than on concepts. Russell developed a whole system of symbolic notation (to express functions, variables, constants), which has become the common language of logicians. In recent times, logic has become a field of research on its own, field that can be divided (roughly) into four main sub-fields: semantics (concerned with the relations between discourse and models), the theory of demonstration (concerned with the syntactic study of the proofs), the theory of calculability (concerned with the analysis of calculation methodologies), and "philosophical logic" concerned with the application of logical methods to the study of notions such as necessity, condition, or belief. Works of logicians such as Hilbert¹⁵⁰, Church, Kleene, Rosser, Gentzen and Tarski marked the development of a more 'mathematical' logic, whose applications have been extremely important in computer sciences.

¹⁵⁰ Hilbert's formal project consisted in separating the mathematical concepts and their subjective interpretation. Mathematical symbols were defined by their inter-mutual relations, independently from their 'sense' inferred from natural language. This aim of this project was to build mathematics through a language without any signified, defined only by its syntax (hence without any semantic), and whose all results would be inferred only from its syntactic rules.

Another approach to the analysis of natural languages was the one developed by a logician, Richard Montague, in the late 1960s and early 70s. In Montague's approach (in Thomason, 1974), rooted in formal semantics, there is no such thing as deep structure, the semantic component being truth-conditional. But Montague semantics operates with a concept of relative truth (as opposed to absolute truth-conditional theories), or truth-in-a-model. Following Frege's distinction between reference (*Bedeutung*) and sense (*Sinn*), in model theory, the sense (or intension) of a sentence refers to its propositional content, while the reference (or extension) refers to its truth-value (under an interpretation, within the model theory considered). In Montague's theory, the core issue is compositionality, the central function of syntax being not simply to generate the well-formed expressions of a language, but to do so in such a way as to provide the necessary structural basis for their semantic interpretation. Syntactic rules put expressions together to form more complex expressions, and corresponding semantic rules interpret the whole as a function of the interpretations of the corresponding parts. This interpretation is made possible thanks to the homomorphism mapping elements of the syntactic algebra (the 'syntax') onto elements of the semantic algebra (the 'semantics'). Montague's aim was to establish a correspondence between syntax and semantics, by adopting a particular kind of categorial grammar and by matching syntactic categories to intensional categories. Most linguists and philosophers originally rejected Montague's distinction between the description of a language and the description of the language-user's of the language, this approach contradicting the Chomskyan programme that was equating grammar with what was in the head. This opposition was exacerbated by the controversy about the possible-worlds basis of the analysis of intentionality in Montague's semantics¹⁵¹. Montague grammar, 'rediscovered' after Montague's death, seems nowadays to be a particularly powerful and successful tool for the semantic analysis of natural languages. Even though meaning was treated as essentially static

¹⁵¹ Within Montague's framework, all logically equivalent sentences are treated as having the same semantic interpretation when propositions are analysed as sets of possible words.

and sentence-bound in classical Montague grammar, recent years have seen the development of models including the incremental flow of information through discourse¹⁵².

5.2 Translation and formal semantics

One main mathematical approach to translation is based on the use of logic (mainly standard propositional logic until quite recently), and the relation between logic and semantics is also highly debated. McCawley (1979), Lakoff (1988), Postal and others argue for a correspondence between the basic syntactic categories and the primitive terms of symbolic logic, while Brame (1976) and others strongly oppose the adoption of the object language of symbolic logic as a syntactic or semantic basis for natural language. Many attempts to translate natural languages into formal systems (such as first-order or higher-order logic) failed because the formal systems had to be extended, e.g. with the introduction of a new symbolic notation, that was generally not universally accepted. The numerous attempts to use mathematics in translation are ascribable to the attractiveness of finite set theory for the formalisation of the transfer of meaning.

‘The meanings that can be expressed in a human language can be represented by the set of sentences of the language. The set is infinite, but can be (very roughly speaking) represented in a finite way as a two-part structure: one, a finite set of basic sentence types, and two, a finite set of operations which recursively generate complex structures from simpler ones. This two-part structure also represents the set of meanings expressible in the language. The meanings of basic sentences are in general lexically specific, but the meanings of complex sentences are normally understandable as a function of the meanings of the structures the complex ones are formed from. That is, if we know the meaning of the underlying structure, and we know what meaning change is introduced by the operation which forms the complex one from the simpler one, then in general we know the meaning of the complex sentence.’ (Keenan, 1975: 164).

¹⁵² Such as the Discourse Representation Theory (DRT).

Keenan (1975) raises some logical issues related to translation and logic and challenges the Exact Translation Hypothesis (anything that can be said in one language can be translated exactly into another language) in view of logical analysis. His arguments are twofold:

- the set of syntactically unanalysable predicates in any given language will not have a complete set of semantic equivalents in most other languages;
- the operations which form complex structures from simpler ones in different languages are also not exactly the same (as exemplified by passive-like rules across languages).

Despite the controversy over the use of logic in translation, some scholars developed models of semantic interpretation. Richard Montague (in Thomason, 1974) advocated a model-theoretic semantics through the transformation of excerpts of English into a formal language. The interpretation of this formal language is supposed to induce the interpretation of English. Each expression of a language is related (through the translation procedure) to a **single** expression of the formal language, and to a **single** meaning (each expression of the formal language being related to one meaning). The core issue in Montague's theory is compositionality (which is ascribable to Frege, 1953) and is also called the homomorphism requirement: the meaning of a complex expression is a function of the meanings of its parts and the way they are syntactically combined. In other words, translation of any constituent is seen as the result of some particular operation applied to the translations of its immediate subconstituents. Moreover, translation between formal languages is defined with the help of functions mapping expressions between the two formal syntaxes considered. Translation from a language L into a language L' is consequently defined as "a certain function from the syntactic analyses of expressions of L' of the same semantic type".

Most formal semanticists claim – although their theory deals with formal syntaxes and semantics – that such a translation process can be extended to natural languages. For the time being, however, no natural language may be characterized in such a way. Montague's work paved the way to a formalization of the translation process since it allowed for some kind of semantic interpretation of syntax and lexicon. His work was based on excerpts of a disambiguated version of English (English – like every natural language – being an ambiguous

language, some of its expressions are related to a set of meanings). Cooper (1975) developed Montague's theory and mapped (directly) expressions of English into sets of expressions of the logic.

Logic is not an aim, but rather a tool used by logicians to build a system whose purpose is understanding (understanding of natural languages, more particularly), therefore logic is for semanticists a 'door' to meaning. Logic is often said to be unable to account for the processes at stake in translation, since it is based on the notion of truth, and one may wonder how questions, or fairy tales can be related to 'truth'. This distrust is due to a misunderstanding about 'truth', which has taken in logic (at least since Tarski, 1933) the meaning of adequacy between a well-formed formula and an interpretation structure. But logic does not provide any statement about this structure and the external reality. Logic is a tool used to infer links between sentences (when the little prince says 'In one of the stars I shall be living', the logician does not ask himself if this is possible, or true, but he uses this sentence to deduct inferences). The sentence is simply taken as 'valid' in the interpretation structures of the knowledge base. Meaning is then deduced, extracted from the processing of these data, in the framework of a logical theory of inferences. Logicians originally called 'syntax' and 'semantics' what most computer linguists nowadays call 'operational' and 'denotational' approaches to the logical theory of inferences. Therefore there is no difficulty for a formal semanticist in building an interpretation of a text like *Le Petit Prince*. Whether it is 'true' that snakes and roses speak is irrelevant. The interpretation is perfectly feasible within the possible world so defined. Meaning is the outcome of a dynamic process of understanding derived from the semantic representation. The linguistic structure determines only part of the initial process.

5.3 Lambda-calculus and combinatory logic

Combinators and lambda-calculus are abstract programming languages aiming at describing some of the most primitive and general properties of operators and combinations of operators.

- **Lambda-calculus**

The logician Alonzo Church (1941) invented the λ -calculus, as part of a system of higher-order logic and function theory. λ -calculus is actually a collection of several formal systems, based on a function-notation. They are designed to capture the most basic aspects of the ways that operators or functions can be combined to form other operators. The ' λ ' is an auxiliary symbol, used to construct functions.

Ex: ' $x - y$ ' can be thought as defining either a function f of x ($f: x \rightarrow x - y$) or a function g of y ($g: y \rightarrow x - y$). Church converted these functions into $f = \lambda x.x - y$, and $g = \lambda y.x - y$.

The lambda-calculus is nowadays used in many programming languages, such as Lisp. In linguistics, the lambda-calculus is used to determine the semantic representation of sentences. A sentence such as John loves Mary will be written: $(\lambda x. x \text{ love } y)$. Different operators and functions can express notions (such as possession, tenses, progressive forms, or reflexivization) in a mathematical way. Operations (substitution, β -reduction, ...) can be performed on the equations to reduce a sentence to a basic semantic expression.

- **Free and bound variables**

An occurrence of a variable x in a term P is bound if and only if it is in a part of P with the form $\lambda x.M$ (it refers to a formal parameter introduced in the expression), otherwise it is free (it is not a formal parameter in the expression). If x has at least one free occurrence in P it is called a free variable of P .

- **Combinatory logic**

Combinatory logic is a functional object-system. It has nothing but objects and the application operation for combining primitive objects to produce complex objects. For example, the mathematical operations $(a + b)$ and $(a \times b)$ are themselves considered as objects. An operation is an action, while an operation seen as an object is called an operator. In combinatory calculus, the fundamental primitive concept is the concept of application, indicated by juxtaposition and

parentheses, such that if X and Y are expressions, then (XY) is an expression. Combinatory logic is used to show that bound variables can be eliminated without loss of expressiveness. The main drawback is that the expressions are usually less clear than the λ -expressions.

5.4 Montague Grammar (MG)

Richard Montague's semantic theory was derived from his approach to language analysis, which, according to him, belonged to mathematics, considering the similarities between natural and artificial languages¹⁵³. Montague's texts did not receive much attention when he wrote them (most were actually published after his death), since they contradicted the main views developed at that time. For philosophers like Quine (1960), mathematical analysis of natural language was not possible, the latter being too unruly and therefore requiring a transformation. For others (like Austin, 1962), no result of research on artificial languages should be applied to natural languages, on the grounds that natural languages were considered as too different from artificial languages. For Montague, syntax and semantics of natural and artificial languages could be encompassed in a unique mathematical theory. His work aimed at describing, analysing and formalising the elements of this mathematical theory, whose languages (whether natural or artificial ones) represent as many realizations.

5.4.1 Montague's Universal Grammar

Montague's 'Universal Grammar' is not to be taken as the linguistic Universal Grammar theory (where it refers to common features of all natural languages). Montague's Universal Grammar refers to a mathematical framework allowing for a description of any language system, whether a natural language or an artificial one. Within Montague's framework, syntax is formalised as an

algebra¹⁵⁴, semantics is an algebra, and there is a structure-preserving mapping (homomorphism) between them¹⁵⁵.

The central tenet of Montague's theory is that the function of syntax is not only to generate the well-formed expressions of a language but also to do so in such a way as to provide the structural basis for their semantic interpretation, meanings being associated with a set of unambiguous syntactic expressions. The interesting feature of his theory lies in the fact that meanings can even be assigned to ambiguous expressions, by first assigning meanings to expressions of an unambiguous (artificial) language¹⁵⁶, and by pairing these unambiguous expressions with the expressions of the ambiguous language.

A disambiguated¹⁵⁷ language is a system $\langle A, F_\gamma, X_\delta, S, \delta_0 \rangle$ such that :

- (1) $\langle A, F_\gamma \rangle_{\gamma \in \Gamma}$ is an algebra ;
- (2) for all $\delta \in \Delta$, X_δ is a subset of A ;
- (3) A is the smallest set including as subsets all the sets X_δ (for $\delta \in \Delta$) and closed under all the operations F_γ (for $\gamma \in \Gamma$):
- (4) X_δ and the range of F_γ are disjoint whenever $\delta \in \Delta$ and $\gamma \in \Gamma$;
- (5) For all $\gamma, \gamma' \in \Gamma$, all sequences a in the domain of F_γ , and all sequences a' in the domain of $F_{\gamma'}$, if $F_\gamma(a) = F_{\gamma'}(a')$, then $\gamma = \gamma'$ and $a = a'$;
- (6) S is a set of sequences of the form $\langle F_\gamma, \langle \delta_\xi \rangle_{\xi < \beta}, \varepsilon \rangle$ where $\gamma \in \Gamma$, β is the number of places of the operation F_γ , $\delta_\xi \in \Delta$ for all $\xi < \beta$, and $\varepsilon \in \Delta$; and
- (7) $\delta_0 \in \Delta$.

¹⁵³ "There is no important theoretical difference between natural languages and the artificial languages of logicians". Montague, R., 1974: 1222.

¹⁵⁴ For example, the expression 'love him' is generated by using a certain operation to combine 'love' and 'he', and there is a structural relationship among these 3 expressions.

¹⁵⁵ The advantage of such a general definition is that it leaves a greater freedom as to what sorts of things the elements and the operations of these algebras are.

¹⁵⁶ A structurally unambiguous language is usually understood as a language in which each expression generated by the syntactic rules as a member of a certain category has only one syntactic analysis (is generated in only one way) as a member of that category.

The sets X_δ are regarded as the categories of basic expressions of the disambiguated language, the operations F_γ as its structural operations, the set A as the set of all its proper expressions (that is expressions obtainable from basic expressions by repeated application of structural operations), δ_0 as the index of its category of declarative sentences, and S as the set of its syntactic rules.

5.4.2 The Proper Treatment of Quantification in Ordinary English (PTQ) and Syntax

One of the problems faced by computational linguists is the complexity of natural languages. We have not been able so far to determine a well-defined set of rules to define such a system, as in the case of formal languages. Montague's option was to write rules designed to generate a precisely defined fragment of this language (a sublanguage), regulating artificially the fragment under study in a way similar to the procedure used by logicians towards formal languages. The idea behind Montague's process is that the formalization of the weaker languages can be carried over into the stronger ones. The Proper Treatment of Quantification in Ordinary English (PTQ) is a theory of semantic representation in which the model theoretic interpretations of natural language sentences (and expressions generally) are built up by rules operating in strict correspondence with the syntactic operations that generate their structural representations.

Montague's syntax defines¹⁵⁸ many sets of sentences, one set for each category of the language. The definition determines the expressions of the language, and indicates the categories the expressions belong to. The base (as defined in note 10) is the lexicon of the language, i.e. the lexical items (or basic elements) of the language. Montague uses several recursive clauses,

¹⁵⁷ Thomason, 1974: 225.

¹⁵⁸ Montague uses a recursive definition, consisting of a base, a recursive clause, and an exclusive clause. For example, to define the set of all strings consisting of a's and b's and containing at least either a or b, i.e. {a, b, aa, ab, ba, bb, aaa, ...}, the following recursive definition could be used:

BASE: $a \in X$ and $b \in X$.

RECURSION CLAUSE: If $\alpha \in X$ and $\beta \in X$, then $\alpha\beta \in X$.

EXCLUSION CLAUSE: Nothing is a member of X except as required by the base and recursion clause. (Cooper, 1980: 21)

which represent the rules of the language. For example, another typical recursive rule in Montague's syntax can be the following one:

If $\alpha \in P_C$ and $\beta \in P_D$, then $F_i(\alpha, \beta) \in P_E$, where $F_i(\alpha, \beta)$ is ...

Which reads as:

If α is a phrase of category C , and β is a phrase of category D , then the result of applying the syntactic operation F_i to α and β is a phrase of category E , where the result of applying F_i to these arguments is...

All the members of the set are generated by applying the recursive clauses. One important feature of Montague's syntax is that each recursion clause must yield well-formed expressions of the language (well-formedness constraint, Partee, 1979).

5.4.3 Montague Grammar and Translation

In PTQ, English sentences are not interpreted directly but are translated into expressions of Montague's intensional logic, by means of a set of translation rules. The translation procedure is purely compositional and can be viewed as assigning 'meanings' to all English expressions, though it assigns only indirectly senses and denotations to them.

1. There is an assignment function of meanings to basic expressions of English;
2. There are operations corresponding to each syntactic rule that operate on translations and give new translations;
3. There will be a set of all the translations of basic expressions and translations that can be produced from them by applications of the operations given in the translation rules.

The translation of an expression into intensional logic in PTQ can be seen as a 'semantic representation', this translation representing in a way the semantical object which is to be the extension (and also indirectly the one which is the intension) of the English expression.

The use of the compositional principle reduces the risk of errors, since phrases and their meanings are constructed step-by-step. Moreover, the focus on semantics allows to assess the correctness of a translation.

5.4.4 Montague Grammar and Machine Translation

Although most research on Montague Grammar was concerned with the theoretical aspect of semantic analysis, a few computer applications were also developed. For example, Friedman and Warren (1978) devised a context-free parser based on the PTQ grammar, while Landsbergen developed a parser for a class of grammars, called M-grammars, based on Partee's transformational extension (Partee, 1976). His research was subsequently applied to machine translation in the Rosetta project, but with a few changes.

The theoretical assumptions of Landsbergen (2003: 244-245) are as follows:

- the linguistic aspects can be separated from all other aspects, hence allowing for the building of a 'stripped' system, using linguistic information only;
- the systems considered translate isolated sentences;
- the translation systems are not able to translate sentences unambiguously, but they define a set of possible translations;
- the function called F-PTR is defined as the function that operates on a sentence of the source language and yields the set of possible translations into the target language. F-PTR is reversible, i.e. if s' is a possible translation of s , then s is a possible translation of s' :

$$s' \text{ in F-PTR}(s) \quad \longleftrightarrow \quad s \text{ in F-PTR}'(s')$$

- the 'correct' or 'best' translation of s (selected on the basis of extra-linguistic information) should be an element of the set $\text{F-PTR}(s)$;

- the translation function of F-PTR must be defined in such a way that correct sentences of the source language (SL) are translated into correct sentences of the target language (TL);
- the information (or 'meaning') that must be conveyed during translation must be clearly defined.

Landsbergen then defines a Montague Grammar (or his M-grammar, or any other grammar developed from Montague Grammar) for the source and target languages, and derives analysis and generation components. At the analysis level, the system translates a syntactic derivation tree into the logic according to the semantic component of the grammar. At the generation level, the system performs the reverse function.

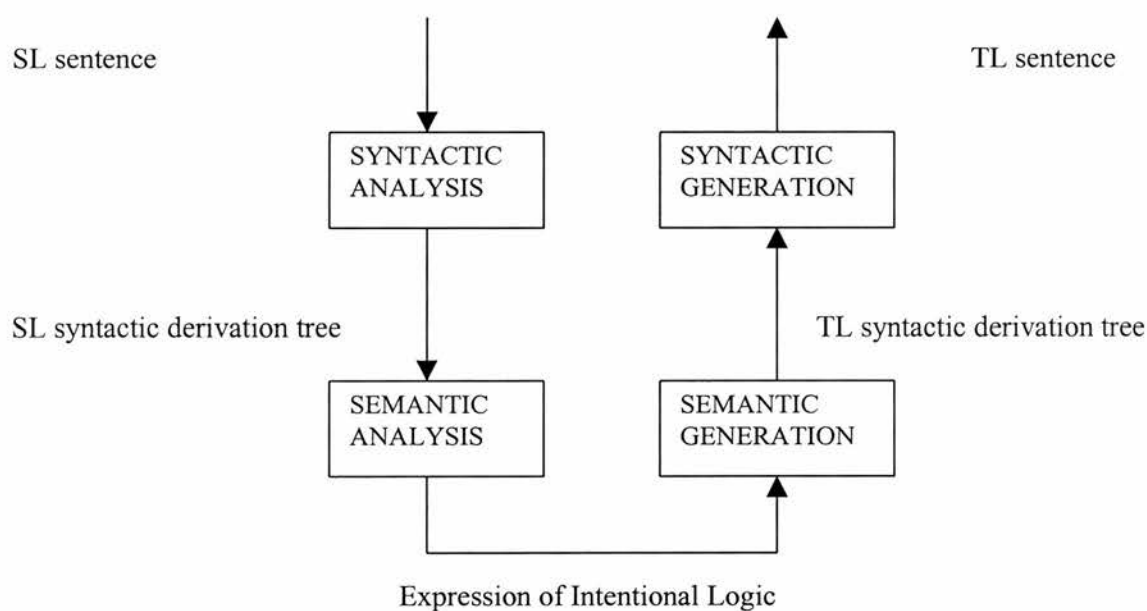


Figure 5-1: Interlingua system using MG and intentional logic (Landsberger, 2003: 245)

The actual system developed in the Rosetta project is not using intentional logic, but a transfer system, considering the following problems:

- most of the research done on Montague grammar deals with detailed semantic analysis of small fragments of one source language. In machine translation, it is necessary to know whether a term or a construct of a source language means the same as a term or construct of the target language, and not enough research has been done in that respect;

- intentional logic seems to be unable to carry all the information (pragmatic or stylistic aspects) that must be conveyed between the source text and the target text;
- Montague Grammar translating natural languages into a subset of intentional logic, two Montague Grammars for two languages might not map them onto the same subset.

The analysis of the human translation of *Le Petit Prince*, as presented in chapter 6, addresses some of the issues presented above. The determination of equivalent terms or constructs between source and target sentences, as well as the determination of the pragmatic and stylistic aspects to be conveyed, will be helpful in developing the next generation of translation systems, based on a true interlingua architecture.

5.5 Categorical grammars

The term 'Categorical grammars' refers to several generalisations of a core context-free grammar formalism (based on a correspondence between syntactic categories and semantic types), which was originally developed by a Polish logician, Ajdukiewicz, in 1935. This formalism was rooted in the works of Frege (1892), Carnap (1956) and Tarski (1933). All these theories are based on the following principles:

- the main burden of syntax is borne by the lexicon;
- the constituents are characterized, both syntactically and semantically, as 'functions' and/or 'arguments';
- the relation between syntax and semantics is characterized as 'compositional'; and
- a tendency to 'freer surface constituency' than traditional grammar.

In a categorical grammar, all grammatical constituents, and in particular all lexical items, are associated with a 'type' or 'category'. The category defines the potential of the grammatical

constituent for combination with other constituents to yield compound constituents. The category is either one of a small number of ‘basic’ categories (such as NP, or VP) or a ‘functor’ category. A functor category has a type, which identifies it as a function mapping an argument of some type onto the result of some (possibly different) type.

Example: English intransitive verbs like *walks* are defined as functions from noun phrases NP on their left to sentences S. (S\NP).

English transitive verbs like *married* will similarly be written as follows: (S\NP)/NP

Different notations were developed, most of them using the ‘slash’ principle, as used originally by Bar-Hillel and Lambek, with variations, as in combinatory categorical grammars (CG).

Lambek notation: sees:=(np\s)/np

Combinatory CG notation: sees:=(S\NP)NP

The main problem faced by the original categorial grammar formalism was the discrepancy between the non-context freeness of English, and the context-freeness of this categorial grammar¹⁵⁹. Several logicians addressed this problem in different ways. Lyons (1968) added a (meaning-preserving) transformational component to a categorial base. Richard Montague used a categorial grammar nomenclature to establish the homomorphic category-to-type correspondence among generated expressions, but also used syntactic operations more powerful than simple concatenation for putting expressions together. Cresswell added free-permutations to his categorial grammar. Categorial grammars were subsequently adapted (for example by van Benthem) to the demands of natural languages, and opened new possibilities to logicians and computational linguists to treat English as a context-free language (Gazdar, 1982; Generalized Phrase Structure Grammar).

- **Combinatory Categorical Grammars**

¹⁵⁹ Chomsky argued that many phenomena in natural languages could not be naturally expressed using context free grammars of any kind.

Combinatory Categorical Grammars were developed in order to account for coordinate constructions (such as ‘Harry cooked and ate apples’), by using rules (or families of rules), such as:

Coordination Rule (<&>): $X' \text{ conj } X'' \Rightarrow X'''$

X' , X'' , and X''' are categories of the same type X but represent different interpretations, and the rule is a schema over a finite set of rules.

Similarly, more complicated coordinate constructions (such as ‘Harry cooked, and Mary ate, some apples’) can be accounted for with the help of some operations that combine functions in advance of their arguments. The term ‘combinator’ was coined by Curry and Feys (1958) who built a mathematics similar to the λ -calculi (refer to 5.3).

5.6 Applicative and cognitive grammars

5.6.1 Applicative languages

The so-called 'applicative languages' (or 'functional languages') take their origin in the theory originally conceived by G. Frege (1953) and developed (amongst others) by Schönfinkel (1924). Different languages were consequently developed, such as the lambda-calculus (Church, 1941), combinatory logic (H.B. Curry, F.Fitch), Quine's attempts (1960) or the functional programming languages of J. Backus (1978). These languages were developed as a reaction to the more classical programming languages, which were considered by some computer scientists as too closely linked to the architecture of computers. They rapidly became popular with cognitive scientists, since they allow for a formalization of notions such as "operator", "concept", "process", categorization", "abstraction", "substitution", "intension", which could not be adequately computed by classical logic (proposition calculus/predicate calculus) more concerned with the study of mathematical coherence.

The name of these artificial languages is derived from the 'applicative operation' seen as the fundamental notion from which sets are derived.

Applicative Grammar

The term 'Applicative Grammar' refers to a formalism developed by a Russian logician, S. Shaumjan, who was influenced by generative semantics and the lexical decomposition work of Fillmore. He tried to combine categorial grammar and what is called now situation semantics. The basic units of the grammar are the term (represented by α), and the sentence (represented by β). Other units are derived by combining the two basic ones, and are interpreted as follows:

$\alpha\beta$ - a unit or category applying to a term (α) to yield a sentence (β), i.e. a verb;

$\alpha\alpha$ - a unit or category applying to a term (α) to yield a term (α), e.g. an adjective;

$\alpha\beta.\alpha\beta$ - a category applying to a verb to yield a verb, e.g. an adverb;

$\alpha\alpha.\alpha\alpha$ - a category applying to an adjective to yield an adjective, e.g. an adverb, such as *very*.

Combinations of items are generated by a single operation called 'application' (an operation applying an operator to an operand). In the case of $\alpha\beta$, β labels the category of the operator, α labels the category of the operand, and $\alpha\beta$ labels the category into which the operand is mapped.

Shaumyan's goal was to conceive a higher-level language (called Genotype) likely to be used in **any** computer system. Described by Shaumyan as a 'software machine', Genotype is in theory appropriate both for functional and logic programming¹⁶⁰. The Genotype concept is based on the Applicative Universal Grammar, which was originally a mathematical theory of language derived from three sources: semiotics, categorial grammar and combinatory logic. Genotype is

¹⁶⁰ A functional programme (such as Lisp) is an algorithm, that is a set of prescriptions for computing functions. A logic programme (such as Prolog) is a description of inferences made from some initial statements (axioms) by rules of logic.

defined by S. Shaumyan as a 'universal model of language' that can be used to design compilers, but also object-oriented programming systems, knowledge representation systems, and machine translation systems. The Genotype model is based on a semiotic law called the Principle of Representational Relevance, and formulated by S. Shaumyan as follows:

"If we abstract from everything in the language used what is irrelevant to the representational function of the communication we have to recognize as essential only three classes of linguistic expressions:

- a) the names of objects (called terms, which are either nouns or noun-phrases);
- b) the names of messages (called sentences);
- c) the means for constructing the names of objects and messages (expressions called operators).

An operator is a function that acts on one or more expressions called its operands to form an expression called its resultant. When an operator acts on one operand, we effect an operation called application: the application of an operator to an operand" (Shaumyan, 1990: 142).

According to Shaumyan's hypothesis, there is a hypothetical system of linguistic operators which constitutes the semiotic basis of the languages of the world. The above concepts as formulated by Shaumyan are supposed to determine the simplest possible sign-system for a natural language. Such a language has neither nouns, nor verbs, neither adjectives, nor adverbs, but only terms (names for objects). This minimal sign system is also said to be an object-system, that represents anything in the real world as an object. Things, properties of things, processes, states, actions, messages are all objects. Natural languages being more complex than this minimal sign-system, they are derived from the minimal sign-system, through the use of combinatory logic.

Even though Shaumyan's theoretical framework was extensively published in Russia, the lack of practical implementation limited its development in Applied Artificial Intelligence. Despite the attraction of the rigor and formal consistency, Shaumjan's applicational grammar is often criticised too limited in scope, being too remote from natural languages.

5.6.2 Cognitive Grammars

Cognitive grammars regard grammar as being inherently meaningful, and maintain that grammatical structure is 'symbolic' in nature, being fully describable in terms of symbolic links between semantic and phonological structures. Cognitive grammars are based on the following principles¹⁶¹:

- meaning is not identified with truth conditions, but with 'mental experience' or 'conceptualization';
- a linguistic category is typically 'complex': its adequate description requires not just a single structure, but a set of structures linked by relationships of instantiation and extension to form a network (a lexeme's meaning comprises a network of related senses);
- linguistic semantics is 'encyclopaedic' in scope (everything we know about an entity can in principle be regarded as contributing to the meaning of an expression that designates it, there is no sharp distinction between semantics and pragmatics, or between 'linguistic' and 'extralinguistic' knowledge).

Cognitive grammars aim at determining cognitive invariants independent of languages. Such invariants are supposed to be deductible from the different categorisations (lexical as well as syntactic) operated by languages. These categorisations are seen as semiotic organisations of more abstract cognitive organisations, and they are the expression of human cognitive capacities. Their main characteristics¹⁶² are:

- research aiming at identifying cognitive invariants between languages;
- cognitive grammars oppose any clear-cut separation between morphology, syntax, and semantics (even in some cases, between lexicon and syntax, language being seen as a continuum);

¹⁶¹ Brown, K., J. Miller (ed.) 1996: 51.

¹⁶² Desclés, 1994a: 74.

- semantic categorisations are essential, hence the central attention given to conceptualisations, prototypes, and type degrees;
- the representations developed within cognitive grammars are non-formal, and non-propositional, but rather iconic, imaginal or figurative;
- language is not seen as an autonomous cognitive operation, but rather as interacting with perception, action and reasoning;
- cognitive grammars are closely linked to anthropology and Gestalt.

Langacker's cognitive grammar

Langacker developed his own theory of cognitive grammar, based on the following principles:

- A language can be described completely in terms of semantic structures, phonological structures and symbolic links between these structures;
- Cognitive Grammar establishes an equivalence between semantic structures and conceptualisations (Meaning=Conceptualization), conceptualisation been seen as a cognitive process, it must be included into a context and an environment. In that respect, cognitive grammar opposes the representations based on truth-conditions, and rejects Montague grammar;
- Semantic structures are not universal, on the contrary they depend on each language;
- Grammar does not represent any formal level of representation, but is a conventional representation of semantic structures;
- There is no important difference between lexicon and grammar, lexicon, morphology and syntax forming a continuum of symbolic structures.

5.7 Translation, levels of interpretation and applicative and cognitive grammar

5.7.1 Translation and Genotype

Shaumyan¹⁶³ developed the theory of a translation software, called Genotype translator, based on the following principles:

- a. The syntactic structure of the text in a source language is transformed into the syntactic structure of Genotype. The result is a set of hybrid sentences having the syntactic structure of Genotype, but the lexicon of the source language;
- b. The native lexicon of hybrid sentences are replaced by the universal lexicon of the target language;
- c. Correspondences are established between the lexicon of Genotype and the native lexicon of the target language;
- d. The lexicon of Genotype is replaced by the native lexicon of the target language. The result is a set of hybrid sentences in the target language, having the syntax of Genotype and the native lexicon of the target language;
- e. The Genotype syntax of the hybrid sentences of the target language is transformed into the native syntactic structure of the target language.

For example, the sentence 'Ann purchased a new book', is translated in Russian following the steps:

1. translation into a hybrid sentence with the Genotype syntax and the English lexicon:

((purchased (a (new book)) Ann)

2. translation of the hybrid sentence into the pure Genotype sentence:

(PAST ((BUY (INDETERMINATE (NEW BOOK)) ANN)

'purchase' is replaced by 'buy', the purpose of the Genotype language being to use only a minimal lexicon for every language.

3. establishment of correspondences between the English and the Russian lexicon:

(PAST ((KUP (INDETERMINATE (NOV KNIG)) ANNA)

¹⁶³ Shaumyan, 1990: 152-154.

'buy' = 'kyp'; 'new' = 'nov'; 'book' = 'knig'. The proper endings found in Russian are suppressed since they vary according to the gender, number and function in the sentence. The Genotype words (whether in English, French, German or Russian) are only roots, which are variants of the same concept.

4. The previous sentence is transformed into the hybrid sentence:

((kupila (novuju knigu)) Anna)

5. Which is itself transformed into the native Russian sentence:

Anna kupila novuju knigu.

Within the Genotype Translator, the syntactic structures of the sentences of the source language and of the target language are unified on the basis of Genotype, and the universal lexicon (considered as an interlingual conceptual interface) is reduced to approximately 1000 words-concepts. Genotype is an interlingua in the sense that it is an intermediate language between the source and the target languages.

5.7.2 Cognitive archetypes

The importance of knowledge representation is presented in chapter 4. Different schools in Artificial Intelligence developed several approaches to knowledge representation, frames, scripts, plans and so on. Another approach, developed from the Applicative and Cognitive Grammar, is based on the notion of cognitive archetypes. A cognitive archetype is defined by J.P. Desclés as 'a universal semantic space structure of a sentence independent of particular languages'. The universal space structure of a sentence is a mathematical notion related to the concepts of general typology (a mathematical discipline studying the positions of points in space)¹⁶⁴.

¹⁶⁴ Shaumyan, 1990: 157.

5.7.3 Levels of representation, language and meaning

Today, one of the main areas of research in cognitive sciences is the determination of the levels at which cognitive representations are situated. On the other hand, meaning is clearly a cognitive process, and it must be analysed as such (Langacker, 1987). Computer scientists are faced with two kinds of problems:

- how to articulate the linguistic level and the conceptual level, or in other words how to articulate these two levels of representations;
- how to determine all concepts underlying natural languages.

At the linguistic level, one must distinguish between the discourse and contextual organisations of the signifiers, the organisations of the signifieds, and the 'conceptual organisations', located at a higher level than the signifieds.

Applicative and Cognitive Grammars aim at defining these different levels of interpretation, which are dependent one from the other. Whether a sentence is active or passive, it expresses at a deeper cognitive level an action exerted by an object on another object, but it also expresses the 'direction' of the action. The determination of the functions (applications) aim at linking the two different constructions in order to extract a possible common 'invariant' between the two constructions. The determination of linguistic primitives (such as possession, localisation, mouvement, change, identity, conservation, contrôle, and so on) aim at determining all the operators involved in natural languages, as well as the successive levels of interpretation that contribute to the building of concepts, or cognitive invariants.

5.8 Natural Language Processing, Cognitive Science and Artificial Intelligence

The war between translation theorists (often opposed to machine translation) and computational linguists (sometimes too keen on claiming that human translators can easily be replaced by computers) is rooted in the concept of mind and intelligence. Most linguists would oppose MT

on the grounds that translation equates to mental processes, hence is irreducible to mechanical replication (imprints of the first computers, devised to make operations, but not to ‘reason’). If, following Turing (and Descartes), we assume that intelligence is the ability to reason and communicate; MT actually raises the question of machine and intelligence, as well as the one of ‘mechanical thought’.

The problem of machine translation cannot be seen as a straightforward decoding-encoding process, as Weaver assumed half a century ago. Translation (as performed by human translators) being a cognitive and social operation transferring messages from a SL into a TL, machine translation cannot be reduced to complex but purely linguistic operations.

A computer is a symbol manipulating device with the capacity to perform a certain list of fundamental operations, and is capable of thought, according to the symbol system hypothesis, since it has the same properties as the human mind (universal symbol system – all human thinking consists of symbol manipulation)¹⁶⁵. Operations are performed according to the logical structure (programme). As predicted by von Neumann, part of the difficulty faced by artificial intelligence (AI) lies in the commonest architecture of computers (serial) as opposed to the assumed architecture of the brain (parallel). It appeared quickly that classical computers (sequential systems based on the von Neumann architecture) were unable to deal with symbolic calculus, hence the development of new tools: functional (or applicative) programming, declarative programming, and combinatory logic. Recent years also saw the development of parallel computers, several central processing units being linked and working together to process data simultaneously.

If we assume an analogy between computers (artificial automata) and brains (natural automata), research in cognitive sciences tends to underline a radical difference between these two kinds of automata in terms of logical approach and structure. Nevertheless, the basic assumption of Turing (developed and extended by von Neumann), according to which any machine can

¹⁶⁵ Newell and Simon, Information Processing System Paradigm.

produce another machine equivalent to itself, led to the theory of the possibility of the replication of natural language by machines.

« There exists a Turing machine [set of quintuples: L current state, symbol read, next step, symbol written, direction of state motion] U (actually a whole class of machines) whose alphabet consists of the two symbols '0' and '1' such that, given any procedure written in any precise and unambiguous language and a Turing machine L embodying the transformation rules, the Turing machine U can imitate the Turing machine L in L's execution of that procedure. » (Weizenbaum, 1976: 62.)

The comparison between the brain and the computer is based on the brain/hardware and mind/software analogies¹⁶⁶. Although language performance is not transmitted genetically, but acquired, language competence is genetically transmitted. Children may acquire any language, whatever their ethnic origins. This observation led computer scientists to consider the brain as a kind of universal machine similar to the Turing machines: the 'hardware' is general enough to allow for the acquisition of *any* natural language. By admitting that one product of brain activity (language) can be described with the help of algorithms (hence with one Turing machine), each language can be seen as one particular representation of this 'universal code', equivalent to any other natural language. Translation practice strongly supports this theory, since any message in any language seems to be 'translatable' into any other language. Recent research, based on that assumption, marks the renewal of interest in the interlingua pattern. Cognitive scientists aim at modelling the universal operations underlying languages (but do not necessarily suppose universal structures of languages), or conceptualisations, as developed by R. C. Schank and others: "there exists a conceptual base that is interlingual, onto which linguistic structures in a given language map during the understanding process and out of which such structures are created during generation (of linguistic utterances)"¹⁶⁷. Schank's basic idea is that every natural language utterance is a manifestation, an encoding, of an underlying conceptual structure. Other authors prefer the term 'symbolic representation', hence allowing for an easier analogy between brains and computers, while refuting Fodor's argument, and freeing themselves from the

¹⁶⁶ the relation between the mind's information processing and the brain's biological states.

'hardware' component. For J.P. Desclés, "the brain, in order to process highly structured data, like that existing in natural languages, proceeds by compilation: that is, by generating intermediate symbolic representations between the most external expressions and the compiled expressions directly compatible with the physical structures of the processing organs"¹⁶⁸. Different levels (for example in language analysis, linguistic, cognitive and neural levels) are interrelated by a compilation process, and the main goal of cognitive sciences should be to search for a cognitive level where language and perception (representation of spatial positions and movements in space) become analysable with the same basic concepts and the same relations. On the other hand compilation is already used in computer sciences to design high-level programming languages abstractly and independently of the execution processes (therefore reducing the influence of the structure of digital computer hardware). The generation of intermediate representations, through a compilation process¹⁶⁹, allows the handling of complex symbolic representations by relatively elementary components. If such a theory may be applied to NLP, MT would greatly benefit (through the interlingua architecture) from the breakthrough in cognitive sciences.

¹⁶⁷ Schank, R.C., 1972: 553-554.

¹⁶⁸ Desclés, J.P. 1989: 132.

¹⁶⁹ « A compilation process (...) consists of a set of programmes, arranged in a hierarchy, that automatically insures the translation of external expressions accessible to human users into internal representations directly compatible with the electronic structures of information processing machines (computers). The intermediate symbolic representations correspond to each of the processing phase. » (Desclés, J.P 1989: 125).

6 HUMAN AND MACHINE TRANSLATION OF LE PETIT

PRINCE

6.1 Introduction

Le Petit Prince, written by Antoine de Saint-Exupéry in 1942¹⁷⁰ (and illustrated by the author himself), is a text of twenty-seven chapters of differing lengths, completed by a short text (a kind of afterthought) directed to the reader. It relates the meeting between the author stranded in the desert and the little prince (a child) who left his own tiny solitary planet to find a friend. The book gives an account of the little prince's travels, his meetings with people on other planets, and eventually his arrival on earth, his taming of a fox, and his friendship with the pilot Antoine de Saint-Exupéry¹⁷¹ in the Sahara. *Le Petit Prince*, written shortly before Saint-Exupéry's accidental death, resumes many themes developed in his previous books, and can be seen as an introduction to his philosophical novel, *Citadelle*, which was published in 1948 after his death. *Citadelle* develops more fully Saint-Exupéry's philosophical and theological quest (the "meaning of life") and illustrates the author's main concerns: truth, friendship, fear and courage, death.

Eugen Drewerman (1992: 11) claims that, in a few centuries, if our libraries were to contain only a few books representative of each era, the works of Dante would exemplify the Middle Ages, the Elizabethan period would be seen through the works of Shakespeare, and only two books would suffice to describe and sum up our epoch: *The Castle*, by Franz Kafka, and *The Little Prince* by Antoine de Saint-Exupéry. One cannot know if *Le Petit Prince* will be read by the coming generations, but its message appears timeless and independent of cultures. Its

¹⁷⁰ First published in New York in 1943 .

¹⁷¹ Antoine de Saint-Exupéry (1900-1944) related in most of his books his experience as a pilot (*Courrier Sud*, 1929, *Vol de nuit*, 1931, *Terre des hommes*, 1939, *Pilote de Guerre*, 1942).

success amongst millions of readers could be taken as a proof of an assumed ‘universal meaning’. Still, how can such a meaning override barriers of cultures and languages? Is there really such a ‘universal meaning’ located in the apparently simple sentences making up a nice fairy tale about a prince-child, his tiny planet, his rose, his friend the fox and his encounters during his travels in the universe?

Like many books for children, *Le Petit Prince* opens a maze of interpretation levels, and raises questions about the deep meaning hidden behind the words telling the story of the little prince, his rose, his planets with its volcanoes and the threatening baobabs, his friendship with a fox, his travels and meetings with strange grown-ups on different planets. *Le Petit Prince* even opens with a discussion about the problem of meaning (through the topic of the author’s drawings when he was a child – boas swallowing elephants drawn from the inside, or from the outside). The whole book, behind simple words and statements, contains several ‘philosophical lessons’ that are supposed to be inferred by the adult readership through the interpretation of the symbols contained in the book.

This book represents an interesting object of translation study for several reasons. First of all, despite being only moderately successful when it was published, it has become over the years one of the best-known French books in the world. Translated into over one hundred languages¹⁷², it offers a rare wealth of material for comparative linguistics. Secondly, belonging to children’s literature, it presents some specific features highly valuable for the study of translation of meaning, as well as for the study of the controversial issue of syntax and semantics. Written for children, its syntax and lexicon are deliberately kept simple: short sentences, reduced variety of words, limited range of grammatical devices (especially, a small number of relative clauses). But children’s literature is also targeted at adults (who read the book to their children, before they can read it by themselves), and can therefore be interpreted on two different levels (the child's one – in the world of fairy tales, and the adult one – philosophical enquiry).

¹⁷² *Comité de l'Association des Amis de Saint-Exupéry*, 1981: 11.

The analysis of the operations performed by a human being during the process of translation requires a comparison of translations into several languages (preferably not too close syntactically and morphologically) of the same source text. Discrepancies and similarities between these translations will provide interesting data about the actual process of translation. An additional comparison between these translations and the output of a machine translation software will enable us to identify the limits of machine translation, will allow for an improvement of the software used for this experiment, and will open new areas of research on translation. Although translation software packages currently available are supposed to be used for non-literary texts, the text selected is a literary one displaying a large variety of linguistic features, despite a rather simple lexicon and syntax. Its interest lies in the fact that it belongs to a sublanguage (children's literature), defined by specific norms governing this text type. This choice is amenable to two reasons. First of all, even direct systems render reasonably good translations of highly limited sublanguages (such as in the case of weather forecasts) and the most common criticism about machine translation refers to its impossibility to render even a 'good enough' translation of a literary text. The purpose of this dissertation being to study the limits of machine translation, it was necessary to analyse the automatic translation of a literary text. Secondly, considering the architecture of most translation programmes currently in use, computers at their present stage of development are unable to tackle natural language as a whole. Sublanguages, on the other hand, are a better object of study, since their syntax and lexicon are limited and can be completely described. Too many translation programs failed because they were devised at the level of natural languages, and as explained in the preceding chapter the dimension of natural languages render the formalisation at this level impossible. If one wishes to devise translation programmes, one should customize the software to different text types, whose syntax and lexicon can be completely analysed and stored in the data base.

It is also important to keep in mind that translation is a complex set of operations performed on texts rather than languages, and the comparison of the source and target texts aims at determining to what extent the source text and the respective target texts are in an 'equivalent' relationship. The additional comparison between the linguistic structures of selected source and

target sentences aim at defining the 'transfer' linguistic operations, supposed to express cognitive processes. The study of the output of the translation software does not aim at demonstrating that machine translation is unable to translate such a text, but aims at pinpointing the various difficulties encountered by the program (syntactic analysis, anaphora resolution, lexical selection in the target language). Although belonging to literature, *Le Petit Prince* can be considered as an example of a sub-language (children's literature) whose features should be of interest in human as well as in machine translation.

6.2 Text type, text variety, function and translation of *Le Petit Prince*

The determination of the text type is now acknowledged as essential in the process of translation (translation theory has evolved from a linguistic approach to a more communicative and text-oriented approach, aiming not only at a functional equivalence between the source and target texts, but also at an adequacy of the target text in the target culture), even though texts being essentially hybrid, text typologies can only account for predominant tendencies. The analysis of the text type will define the structure and the general features of the target text(s), and determine the general method adopted by the translator. According to the equivalence principle, the function of the target text in the target culture must be equivalent to the function of the source text in the source culture. Even though the target texts should fit into the 'fairy-tale' pattern in each of the target cultures, discrepancies may nevertheless appear, since fairy tales are also anchored in beliefs and rituals specific to each culture. Following Reiss' typology¹⁷³, *Le Petit Prince* would belong to the expressive type (as opposed to the informative and operative types)¹⁷⁴, articulated around artistically organized contents. The contents in the target text should therefore be conveyed in an analogous artistic organization. The translator would be expected to translate by identification with the artistic and creative intention of the SL author.

¹⁷³ Reiss, in Venuti, 2000: 163.

¹⁷⁴ Other authors (Newmark, 1982) adopt the following distinction: expressive, informative, and vocative texts. Such a distinction can be linked to Bühler's organon theory, in which language is taken to be a tool for dealing with emotive meanings, referential meanings, and connotative meanings.

At the same time, the determination of the text variety (the text is classified according to specifically structured socio-cultural patterns of communication belonging to specific language communities) will entail specific choices in terms of structure conventions. The translator must adopt the target culture children's literature pattern ("il était une fois" would have to be translated into "once upon a time", "es war einmal", "zhil-by!"), to fulfil a similar function in the target literature and culture.

Even though *Le Petit Prince* departs in some ways from the classical fairy-tale (it does not start with "Once upon a time...", and it does not end with the 'ever-after happiness of the married hero/heroine'), it can be said to belong to the fairy-tale type: hero with whom the child-reader can identify, hero's quest through perilous travels, and magical powers of the hero. It is an untypical fairy-tale, like *Alice in Wonderland*. Both authors deliberately wrote for children, but they also targeted these books at an adult readership. In the case of *Le Petit Prince*, the first chapter opens with the test of the drawing, setting the problem of meaning. In *Alice in Wonderland*, words are taken 'at face value', and the queen of the card game has therefore all the attributes of a human queen. Both authors were fascinated by language, although in different ways, but both were mathematicians, and approached language not only as a tool to express ideas (Saint-Exupéry was concerned with the notion of 'concept', while Dodgson was more interested in the link between language and logic).

On the other hand, some features of the tale (initiation through the travels and the meetings) place *Le Petit Prince* within the myth type. Even though *Le Petit Prince* presents the destiny of a single character, and therefore conforms to the pattern of the fairy-tale, it also tackles the general issues faced by the humanity as a whole, which is a characteristic of the myth. *Le Petit Prince* is concerned with timeless messages addressed to all human beings: the importance of the time and care given to someone else, the importance of friendship, the vanity of useless activities, the search for truth and, at a deeper level, the search for the meaning of life.

6.2.1 Fairy-tales morphological analysis

In some respects, this text fulfils the pattern developed by V. Propp¹⁷⁵ about tales. V. Propp identified a total of 31 functions fulfilled by the characters. A **function**¹⁷⁶ is the action of a character, defined from the viewpoint of its meaning in the story process. The functions are constant, permanent elements of the tale, whatever the characters, and whatever way these functions are fulfilled in any single tale. The functions are the fundamental elements of any tale. All these functions may not appear in any single tale, but they are always limited in number, and always appear in the same chronological order (following the hero's quest).

For example, one may find in *Le Petit Prince* some of the functions listed by V. Propp such as:

- departure;
- reception of the magical object (the drawing of the sheep);
- move in space;
- recognition and discovery of deceit (the rose is not unique);
- home return (symbolized by the disappearance/the death of the little prince);
- main character (hero) endowed with magical powers (ability to talk to animals, to make a well appear in the desert, to travel from planet to planet).

Similarly, V. Propp identified in fairy tales 7 types or "constant roles" attributed to the characters: the hero, the false hero, the sought-for-person, the donor, the dispatcher, the villain, and the helper. Each of the seven types possesses a specific action area and one or several functions.

A type is not necessarily one character, since a character can cover several types. In the case of *Le Petit Prince*, one can identify the prince with the hero, the pilot with the false hero, the rose with the sought-for-person, the pilot and the snake with the donors, the rose, the geographer and the king with dispatchers, the baobab with the villain, and the pilot with the helper. In many

¹⁷⁵ Propp, V. 1965.

¹⁷⁶ Propp, 1965: 31.

tales, the hero, in his quest for truth, arrives at a border zone (between the inside and the outside, the superficial level and the deep one), where he meets animals that come to his help, talk to him, and show him the right path. In *Le Petit Prince*, this role (*function*, in Propp's analysis) is held by the fox, who will explain to the prince what differentiates his solitary rose on his tiny planet from the 5000 roses he saw on earth. The characters in *The Little Prince* also have the features of fairy-tale characters defined by Propp, they have no first name, no surname. The characters are 'the pilot', the 'little prince' (and the reader eventually knows very little about him), 'the geographer', the 'king', 'the tippler', 'the businessman', 'the rose', 'the fox', 'the baobab', 'the sheep'. These characters do not have any 'motivation', they simply fulfil predetermined roles in a well-structured narrative.

The behavioural rules and the structure of fairy-tales represent a complete semantic system, in which the functions reveal complementary logical links, independent of their syntagmatic links. The fact that the structure of the text and the functions of the characters are constants¹⁷⁷ (as defined by Propp), may also account for the success of the translations of *Le Petit Prince* in so many languages. No reader is really taken aback or feels uneasy, when discovering the text, since it refers to some universal issues (such as friendship), and it follows an expected structure and development. In some other respects, *Le Petit Prince* does not fully follow the fairy tale pattern. A fairy-tale necessarily reaches a point of equilibrium at the end (the hero/ine gets married ...'and lives happily ever after'), a kind of closing, while *Le Petit Prince* ends with an opening ("...Send me word that he has come back"). Far from leaving the reader with a feeling of balance, *Le Petit Prince* ends in a tone of fear and anguish.

The textual cohesion¹⁷⁸ structuring *Le Petit Prince* must be taken into account in the translation process. The story is articulated around the following stages:

¹⁷⁷ One might call these 'cultural and cognitive universals'.

¹⁷⁸ "Cohesion will be considered as an overt relationship holding between parts of the text, expressed by language specific markers" (Blum-Kulka, 2000: 298)

- departure from the little prince's planet; travels - visit of the first asteroid, the second planet, and so on until the hero arrives on the seventh planet, the Earth;
- questioning and information collection on each planet;
- meetings with all the other characters.

These successive steps can be analysed as the different stages of a script (as defined in 4.5.2), and contribute to the representation of the meaning of the text. If one can establish that a text - in this case *Le Petit Prince* - belongs to a specific type (fairy tale), then its characteristics can be listed, analysed and loaded in the memory of translation software. Then the 'sentence-by-sentence' processing can be put in parallel with the structure requirements imposed by the text type, hence allowing for a better translation.

6.2.2 Double articulation, ambivalence of children's literature

Le Petit Prince is a literary text, and as such conforms to the literary norms governing French literature. The translation of a literary text is expected to conform to the literary norms of the target culture. The translation of a poem is usually a poem. A detective story remains a detective story (but might be slightly adapted to conform the 'detective story' norms in the target culture). A text normally has an unequivocal status in the system it has entered: either the text is for children, or it is for adults. But children's literature raises the issue of texts that are actually targeted at adults and children. This feature of children's literature is called ambivalence¹⁷⁹, such texts belonging simultaneously to different and exclusive systems and being therefore read differently (but concurrently) by two groups of readers, the adults and the children. According to Shavit (1986), children's literature allows for a greater freedom in translation, on account of its equivocal status, at the periphery of the literary polysystem. The polysystem theory (Even-Zohar, 1978a, 1978b) assumes that translations never function as totally independent texts, and

¹⁷⁹ "Ambivalent texts give the system its flexibility and the heightened degree of non-predictability in its behaviour. It is for this reason that the internal capacity of the object for creating information (the

that translators always belong in one way or another to a literary and/or cultural environment. In such a particular case, the text may be changed, enlarged or trimmed, as long as the translator keeps in mind that the target text must comply with what the target society considers as 'educationally correct', as well as with what this society perceives as 'understandable by a child'. Within a semiotic perspective, translation of children's literature is not only translation from one language (sign system) into another, but also from the adult system into the child's one.

6.3 Interpretation of *Le Petit Prince*

Children's literature is also particularly interesting for the study of translation of meaning, in the way that semantic analysis of such texts reaches new dimensions. In the case of *Le Petit Prince*, the reader, through straightforward descriptive sentences, enters an imaginative world where animals speak, where drawings give life to sheep, where the little prince travels from planet to planet without any kind of mechanical device, and where European wells (with ropes and pulleys) suddenly appear in the Sahara.

Le Petit Prince opens with the problem of meaning and the whole book makes constant reference to this opening, indicating to the child-reader that he (and only he of course!) will understand the true message hidden in the text. The drawing of a boa constrictor digesting an elephant is seen as the drawing of a hat by the grown-ups, who lack insight and therefore are unable to understand. This double level of interpretation of drawings illustrates the double level of interpretation of the whole text. Within a semiotic perspective, it opposes the icon (hat) to the symbol (shape of a hat, which represents the elephant in the boa). By explaining the lack of understanding of grown-ups, the author creates a kind of complicity between himself and the child-reader. But an interesting feature of the book lies in the fact that the adult-reader should

inexhaustibility of hidden possibilities) is far greater than its description would indicate". (Lotman, 1977a: 201)

not feel excluded or offended by the constant hints about the lack of insight of the grown-ups. Saint-Exupéry subtly makes it clear to the adult-reader who kept his/her child-heart that this wonderful world where true friendship and love are saved up is within his/her reach. Most people have read (or listened to) *Le Petit Prince* when they were children and many read it again when they become adults, attracted by the poetry of the text, as well as by its humanist dimension. Numerous analyses of this book focus on the message(s) carried by the fairy-tale, especially on the secret revealed by the fox to the little prince (“it is only with the heart that one can see rightly”), but as McCrory stresses (1990), the overall meaning of the book is a lot richer. Key themes include: friendship, utility and devotion in life, opposition between material wealth and wealth in human ties, despair, political systems, socio-cultural prejudices, meaning of life and death. The world-wide fame and popularity of *Le Petit Prince* is probably due to the attraction of the purity and truthfulness of the little prince to all men, whatever their culture, education or age. The little prince is actually interpreted by many critics as representing the ‘dead child’ that every adult carries within her/himself. Saint-Exupéry simply says to the adult-readers that this child is not dead, but lost deep inside us, and that we must look for him and wake him up by changing our behaviour, and by ‘seeing with our hearts’.

Different interpretations of *Le Petit Prince* may be inferred, since, according to - McCrory “the text makes use of symbols which mean different things to different people, (...) the gap between statement and comprehension being filled out by levels of meaning” (McCrory, 1990). Drewerman (1992) even provides a psychoanalytical interpretation of *Le Petit Prince*, in which the child-prince is a symbol of Jesus, the characters on the planets are depicted as solitary people desperately in quest of a goal in their life, who represent our misgivings, where the desert is the symbol of wisdom and love, and where the meaning of life is symbolised by the bond between the little prince and his flower. Capestany (1982) makes an analogy between *Le Petit Prince* and Plato’s simile of the Cave, claiming that it is materialism, which ‘covers up things’. Several biographers of Antoine de Saint-Exupéry have seen in *Le Petit Prince* the story of the author, represented by the pilot stranded in the desert but also by the little prince himself. The rose is Consuelo Suncin, his wife, and the book can be seen as the story of the difficulties

encountered by the couple, who seemed to be unable to understand each other. The fox is undoubtedly a reminder of the fennec tamed by Saint-Exupéry himself when he was posted in the Sahara.

There are as many interpretations as there are readers, but it seems that every reader actually finds in this text his own beliefs, wonders, sorrows and joys. The multiple interpretations may be considered as a consequence of the ambivalent status of the text, which belongs to children's literature (Jakobson, 1960, 1971; Lotman, 1977b; Shavit, 1986), but it is basically the expression of the semantic autonomy of the text (Ricoeur, 1976). Ricoeur opposes the unidimensional approach of language (Saussure's concept of sign as 'the only basic entity' composed of two aspects, the signifier and the signified) and argues for the discourse that dissociates speaker's intended meaning and sentence meaning¹⁸⁰. In the case of written texts (as opposed to oral discourse), the author is disconnected from the verbal meaning of the text, the author's intention and the meaning of the text cease to coincide. Being addressed to unknown (and potentially indefinitely numerous) readers, a written text is – by definition – open to as many interpretations as there are readers. On the other hand, such a theory contradicts the view of a single 'meaning' inherent to the text and 'transferred' during the process of translation. The fact that millions of readers around the world seem to have interpreted *Le Petit Prince* in a similar way would support the view of a 'two-sided meaning', simultaneously referential and not referential, the 'transferred meaning' being the referential one.

« There are two aspects of meaning: one which is referential, objective and cognitive and hence, refers to the shared property of the speech community which uses the language of which the word or sentence forms a part; one which is not referential, but associational, subjective and affective. This kind of meaning, being personal, may or may not be shared by the community at large. » (Bell, R.T. 1991: 102).

¹⁸⁰ « Discourse grounds the very existence of language since only the discrete and unique acts of discourse actualize the code (...). Discourse is realized temporally and in present moment, whereas the language system is virtual and outside of time (...). If all discourse is actualized as an event, it is understood as meaning. But the concept of meaning allows two interpretations: to mean is both what the

As Eco stated (1992b), there is an infinity of interpretations, or an infinity of 'virtual' texts. The author writes for an intended readership (one may think that Saint-Exupéry had in mind for example the son of his friend Leon Werth, to whom the book is dedicated, as the targeted reader), he makes presumptions about the reader. On the other hand, the reader makes hypotheses about the author, and constructs a 'virtual' text (or image) from the text at his disposal. This 'virtual' text varies from reader to reader, and it seems unlikely that two different readers would construct the same 'virtual representation', since this representation depends on several factors (such as the socio-cultural environment, the age or the gender of the reader). Nevertheless, the success of *Le Petit Prince* in almost all countries of the world, would indicate a possible unique 'virtual representation', despite the translation process, which is adding another stage of interpretation.

6.4 Stylistic features of *Le Petit Prince*

The main stylistic feature of *Le Petit Prince* is its overtly simple style, written as if by a child (short sentences, numerous repetitions, naive questions...). de Galembert (2001) parsed *The Little Prince* with Flesh for Word 2000 (readability software), and the results of this parsing show the extreme linguistic simplicity displayed by the text, as if the author made a great effort to address his young readers in the most appropriate way. The style is very simple (index < 20/100 with only 2% of passive constructions), and the lexicon is very limited (index 11/100). Many sentences are overtly 'instructive' and pedagogical, the drawings playing a very important role in the explanations given (as in the case of the hat/boa and the elephant).

The stylistic features as presented thereafter correspond to the features of children's literature: short sentences, numerous repetitions, references to other fairy-tales, and drawings.

- **Fairy-tales references**

speaker means (what he intends to say), and what the sentence means (what the conjunction between the identification function and the predicative function yields), (Ricoeur, 1976: 11-12).

Although the book does not start as a stereotypical fairy-tale ('Il était une fois...'), Saint-Exupéry reminds the reader that it is a fairy tale by placing some hints in the text, as in chapter IV (p.16, English translation):

'I should have liked to begin this story in the fashion of the fairy-tales. I should have liked to say: "*Once upon a time* there was a little prince who lived on a planet that was scarcely any bigger than himself, and who had need of a friend..." To those who understand life, that would have given a much greater air of truth to my story.'

The author also borrows sentences from other fairy tales, as Le Hir (1954: 22-23) noticed:

'Approach, so that I may see you better' (chapter X, page 33, English translation), and 'That is a (funny) hat you are wearing' (chapter XI, p.39) are inspired by *Little Red Riding Hood*.

- **Repetitions**

Repetitions play a very important role in the structuring of the text (building its coherence), and adds to the connotative meaning of the tale.

The chapters, which seem unrelated at first sight (some chapters begin with a dialogue with an unknown and undescribed character, as if there was no link to the previous parts), are actually co-referenced through recurrent sentences opening and closing the chapters (with a gradation), as in the chapters X to XII:

Chapter X: "Le premier astéroïde était habité par un roi. (...) "Les grandes personnes sont bien étranges", se dit le petit prince, en lui-même, durant son voyage".["The first of them was inhabited by a king. (...) "The grown-ups are very strange", the little prince said to himself, as he continued on his journey."]

Chapter XI: "La seconde planète était habitée par un vaniteux. (...) "Les grandes personnes sont décidément bien bizarres", se dit-il simplement en lui-même, durant son voyage". [The second planet was inhabited by a conceited man. (...) "The grown-ups are certainly very odd", he said to himself, as he continued on his journey".]

Chapter XII: "La planète suivante était habitée par un buveur. (...) "Les grandes personnes sont décidément très très bizarres", se disait-il en lui-même durant le voyage". [The next planet was inhabited by a tippler. (...) "The grown-ups are certainly very, very odd", he said to himself as he continued on his journey".]

Chapter XIII: "La quatrième planète était celle du businessman. (...) "Les grandes personnes sont décidément tout à fait extraordinaires", se disait-il simplement en lui-même durant le voyage." ["The fourth planet belonged to a businessman.(...) "The grown-ups are certainly altogether extraordinary", he said simply, talking to himself as he continued on his journey".]

The constant reference in the text to the 'grown-ups' (unable to understand, as opposed to the children who understand things), and to the 'matters of consequence' that worry them so much, contributes to the overall coherence of the text:

- 'J'ai ainsi eu, au cours de ma vie, des tas de contacts avec des tas de *gens sérieux*. J'ai beaucoup vécu chez les grandes personnes.' (p.10) ['In the course of this life, I have had a great many encounters with a great many people who have *been concerned with matters of consequence*. I have lived a great deal among grown-ups.', (p.6)]
- 'Mais non ! Mais non ! Je ne crois rien ! J'ai répondu n'importe quoi. Je m'occupe, moi, de *choses sérieuses*. (...) – *De choses sérieuses !* (...) – Tu parles comme les grandes personnes !'(p.28) ['Oh, no ! I cried. No, no, no ! I don't believe anything. I answered you with the first thing that came into my head. Don't you see – I am very busy with *matters of consequence !* (...) – *Matters of consequence !* (...) – You talk just like the grown-ups !' (p.24-25)]
- 'Hein ? Tu es toujours là ? Cinq cent un millions de...je ne sais plus...J'ai tellement de travail ! *Je suis sérieux*, moi, je ne m'amuse pas à des balivernes ! Deux et cinq sept...'(p.45)['Eh ? Are you still there ? Five-hundred –and-one million-I can't stop...I have so much to do ! *I am concerned with matters of consequence*. I don't amuse myself with balderdash. Two and five make seven...'(p.42)]
- 'Des abeilles ? – Mais non. Des petites choses dorées qui font rêvasser les fainéants. Mais *je suis sérieux*, moi ! Je n'ai pas le temps de rêvasser.'(p.47) ['Bees ? – Oh, no. Little golden objects that set lazy men to idle dreaming. As for me, *I am concerned with matters of consequence*. There is no time for idle dreaming in my life.' (p.43)]
- '*Je suis sérieux*, moi, je suis précis.'(p.47) [*I am concerned with matters of consequence*: I am accurate.'(p.43)].

The 26th chapter, which is the climax of the tale (disappearance/death of the little prince) is built like a poem with several repetitions giving rhythm to this particularly long chapter: '*petit bonhomme*', '*et il rit encore*', '*je ne te quitterai pas*', '*moi je me taisais*'. These repetitions are kept in the translations retained for this study (with a slight amendment in German - '*und er lachte noch immer*' instead of '*er lachte wieder*' only once, but this is enough to change the stylistic effect of the text, as well as the dynamic equivalence. Whether this was a purposeful choice by the translators or an unfortunate omission, the target German text lacks in terms of dynamic equivalence.

- **Numbers**

Another feature of *Le Petit Prince* is the constant use of precise numerical references (numbers of the asteroids, numbers of stars, numbers of possible sunsets on planets visited by the Little Prince, and the numerical description of the Earth – 111 kings, 7,000 geographers, an army of 426,511 lamplighters, 900,000 businessmen, 7,500,000 tipplers, 311,000,000 conceited men, about 2,000,000,000 grown-ups), although, at some point, the author claims, that figures are useless and meaningless, important only for adults. Saint-Exupéry was himself a mathematician by formation, and he devised many aeronautical instruments. In his work, figures were “matters of consequence”, since he could not possibly fly without headings and navigation references. His use of numbers in *Le Petit Prince* could reflect his natural mathematical facility, as well as his deep feeling that figures had only the meaning attributed to them by the writer and the reader. The typography deserves some attention: it is not common in French to write in letters such big numbers, and A. de Saint-Exupéry certainly aimed at playing with the words: was he planning to teach the numbers to his young readers, or trying to annoy the grown-ups, much more accustomed to ‘numerical writing’ (typography in figures)? No writer would bother about such an accurate numerical description, unless he wanted to laugh at this characteristic of grown-ups too keen on counting (as exemplified in other chapters – price of a house, account of

stars...). Interestingly enough, the English translator converted the typography (using Arabic figures), while the German and Russian ones kept the numerals as in French. The fact that the Earth is the seventh planet visited by the Little Prince is certainly not a random choice by Saint-Exupéry, who was a mathematician. A week counts *seven* days, there are *seven* deadly sins, there were traditionally *seven* wonders of the world. The first pages of the Bible explain how God created the world in six days and rested on the *seventh*. This seventh day became the Jewish day of rest, the Sabbath, Saturday. It is assumed that there are *seven* principles of the mind; there are *seven* principal tones. Newton discovered *seven* primary colours. In the earliest systems of the world, there were *seven* planets (Saturn, Jupiter, Mars, the Sun, Venus, Mercury, and the Moon).

6.5 Human translation of *Le Petit Prince*

The human translations considered are in English, German and Russian. The versions selected are the pocket editions in French, English and German. The English translation, by Katherine Woods, is the British one (Mammoth, London, 1991, first published by William Heinemann, 1945). The German version is the one published by Karl Rauch Verlag (Düsseldorf, 1952 and 1956) translated by Grete and Josef Leitgeb. The Russian text is the version published in 1963 by '*Molodaïa Gvardia*', translated by Nora Gall. The purpose of this thesis being to study the transfer operations taking place at a cognitive level, it seemed necessary to select languages presenting linguistic discrepancies, especially in terms of syntax. Too close languages would inevitably lead to lexical or syntactic similarities. Common linguistic features (for example between French and English, or between German and Russian) might have been detrimental to a purely bilingual comparison. On the other hand, the limited scope of a dissertation like this one does not allow for a complete and detailed assessment of all the translation features involved. Nevertheless, the most striking results and data likely to be useful in this analysis will be

developed. Apart from a few general comments on the respective translations, no quality assessment of the published translations will be made, the aim of this research being to determine any cross-linguistic systematic cognitive feature of the operations performed by a translator, as well as its possible formalisation.

Translation studies do not suffer from lack of theories, but rather from a dearth of approaches, as shown in chapter 3, the main problem lying in the conflicts between some of these theories. Even though it seems that translators generally processed *Le Petit Prince* sentence-by-sentence, the whole text is necessarily the framework of the translation operations. Translators (at least in the case of the translations selected for this study) definitely took into account the typological features of the text (children's literature/fairy tale). The deliberate intertextual references ('matters of consequence') or the slight adaptations of the text (especially noticeable in the case of the Russian translation) to match the format expected by the target readership support the textual approach to translation.

Overall the human translations present all the features that translation scholars would expect: search for dynamic equivalence, use of syntactic and lexical devices in the target text, which correspond to the format expected by the target readership. Nevertheless, the study of these translations allows us to determine a few important characteristics.

- The English version sometimes sounds clumsy, very close to the syntactic structure of the French source text, but reveals also in some parts a great freedom of interpretation and rephrasing. There are many instances of expansions, through the use of repetitions, or the expression of implied meanings. For example, in the following instances:

a) (F): Elles m'ont répondu : « Pourquoi un chapeau ferait-il peur ? »

(E): But they answered: « *Frighten?* Why should any one *be frightened* by a hat? »

« Frighten? », taken from the previous question is repeated, adding strength to the sentence and underlining the surprise of the grown-ups.

b) (F): J'ai alors dessiné l'intérieur du serpent boa, afin que les grandes personnes puissent comprendre.

(E): *But since the grown-ups were not able to understand it, I made another drawing : I drew the inside of the boa constrictor, so that the grown-ups could see it clearly.*

The statement of the author is reinforced. What was hinted at in French is clearly expressed in English : grown-ups are unable to understand...

c) (F): L'astronome refit sa démonstration en 1920, dans un habit très élégant. Et cette fois-ci tout le monde fut de son avis.

(E): *So in 1920 the astronomer gave his demonstration all over again, dressed with impressive style and elegance. And this time everybody accepted his report.*

The emphasis is put on the elegance of the astronomer's dress, by the adding of another adjective. 'Etre de son avis' is expanded with the introduction of the 'report' which has not been mentioned in French.

- The German version is very close to the French syntactic structure, and presents a wide range of grammatical and lexical features, but also a rather high occurrence of repetitions.

a) (F): J'ai dû vieillir.

(G) : Ich musste ja *im Laufe der Zeit* älter werden.

The adding of 'with time' emphasizes the length of time, hence the ageing.

b) (F): Mais bien sûr, nous qui comprenons la vie, nous nous moquons bien des numéros.

(G) : Wir freilich, die wir wissen, was das Leben eigentlich ist, wir machen uns nur lustig ueber die albernen Zahlen.

The adding of 'albernen' in German ('not serious') emphasises the uselessness of figures, this message of the author being rendered by constant repetitions and references to figures.

- The Russian version is the one presenting the greatest stylistic and syntactic freedom. It displays a great deal of poetic features and could be described as an attempt – and true success – by the translator to transform the French text into a proper Russian fairy tale. The following few examples show how the translator freed herself from the source text, creating a new text in Russian.

a) (F): N'importe où. Droit devant lui.

(R): -Malo li kuda ? Bse priamo, priamo, kuda gliaza gliadiat.

The Russian translation presents a semantic adding ('straight ahead, straight ahead, where the eyes see'). The sentence is also rendered more poetic, thanks to the alliteration (pria/glia).

b) (F): L'astronome refit sa démonstration en 1920, dans un habit très élégant. Et cette fois-ci tout le monde fut de son avis.

(R): V 1920 godu tot astronom snova dolozhil o svoiem otkrytii. Na etot raz on byl odet *na poslednei mode*, i vce c nim soglasilis'.

The elegance of the astronomer is rendered in Russian by 'dressed according to the latest trend'. This interpretation gives some interesting information about the definition of 'elegance' for the Russian translator.

c) (F): J'éprouve tant de chagrin à évoquer ces souvenirs. Il y a six ans déjà que mon ami s'en est allé avec son mouton.

(R) : Serdtse moe bol'no czhimaetcia, kogda ia vspominaiu moego malen'kovo druga, i nelegko o nim govorit'. Proshlo uzhe shest' let s tekh por, kaka on vmeste so svoim barashkom pokinul menia.

Expansion in Russian ("*my heart is compressed with pain*" for "I have suffered too much grief"), and reintroduction of the little prince ("these memories" become "my little friend").

Despite the instances of adaptation and expansion presented above, the overall impression contradicts the view according to which the act of translation is nearly invisible in the case of fairy-tales and children's literature. In most cases of translated children's literature, the adoption of the target literary norms is so complete that the majority of readers remain unaware of the foreign origins of the text. In the case of the English and German translations, though, the source texts are discernible in the target texts. This literal translation might be due to the translation norms (linguistic approach to translation) in Europe at the time of publication (1945 for English, 1952 for German). Similarly, the freedom displayed in the Russian translation could be ascribable to the ideological and cultural norms in the Soviet Union at the time of publication (1963), as exemplified by the 'sultan' image. There is a famous Russian history painting by Zlya Repine, of the Zhaparoghe kossacks refusing the ultimatum from the Sultan of Constantinople. Either the translator made a conscious choice (hence playing the role of an 'educator', and probably following the guidelines for a 'politically correct' translation), or this choice was unconscious, and demonstrates the role of the cognitive processes at stake in the choice of words by the translator. The drawing of the Turkish astronomer (as a sign) might have interfered with the source text, and created a different concept from the one put forward by Saint-Exupéry.

6.5.1 Drawings in *Le Petit Prince* and inter-semiotic references (Annex 8.1)

Whereas the English and Russian versions present all the original drawings, and basically at the same places as in the French versions, the German one contains only a limited number of drawings, hence limiting the scope of the text, which makes constant reference to these drawings, and which cannot be separated from them. The German edition used for this study

was the pocket edition of Karl Rauch Verlag, since the French and the English versions were the pocket editions. There are several other complete editions in a larger format by Karl Rauch Verlag (same translation). The limited number of drawings in this pocket edition having some consequences on the intersemiotic dimension of the text, as well as on the translation itself, the features of this 'truncated' edition are important. The first chapter sets the central issue of meaning through the problem of the drawing of the boa constrictor, and the whole book conveys as much through the drawings as through the text. Moreover, within a semiotic perspective, the drawings represent a sign system in parallel with the sign system constituted by the text, and these systems are inter-twined and co-referenced. Last but not least, drawing was essential to Antoine de Saint-Exupéry, who studied science, then art, before becoming a pilot. According to his friends, he was constantly drawing (mainly sketches of the numerous aeronautical devices he devised), and it seemed that a drawing of the little prince existed before he was asked to write this book. Several versions of this child were drawn on napkins on several occasions. The picture probably existed before the story was made up, and the child became alive through the magic of the drawings, just as the sheep became alive in the drawing of the box containing it. The drawings – although coloured in the original version – are in black and white in the English and German, hence limiting again the scope of the whole book. Pictures are essential in children's literature, since they represent the first means of conceptualisation for children. They learn how to read, first through the association of the pictures and the sounds, then through the association of this picture-sound with the written words. The lack of colours, as well as the suppression of some drawings may not matter much if the book is targeted at an adult readership, but this commercial choice (limited printing budget) is likely to reduce drastically the number of child readers. One may wonder about the publication of a children's book that is unattractive to children. On the other hand, these pocket editions might have been aimed specifically at an adult readership (larger books, containing the exact colour copies of the original drawings, being published both in English and German). Whatever the reasons for such a choice (probably due to a wish to reduce costs), the consequences are crucial for the translation. As shown in the annex 8.1, some translated sentences lose their meaning, since they

refer to a drawing. The overall meaning and pedagogical dimension of the text is also reduced, the child-reader losing the co-reference to the drawings.

6.5.2 Compensation (Annex 8. 2)

An important feature of translation deals with the so-called 'compensation' technique, which usually aims at preserving the dynamic equivalence between the source text and the target text, hence transposing the source text (expressed in a source language aimed at a source readership) and 'adapting' it to the target socio-cultural and linguistic environment. The table presented in the annex 8.2 indicates the kinds of transfer (or compensation) techniques used when the text was not rigorously translated sentence by sentence, but when sentences were merged, or some elements were deliberately included in the previous or the following sentence. The commonest features of linguistic variations ('il me dit' put before the dialogue, but put after the dialogue in the translation – 'he said'), which correspond to a stylistic choice by the translator will not be addressed, since they are of little interest in the framework of this dissertation. The aim of this study is to determine systematic –if any- un/conscious syntactic operations (processed in translation) partially expressing some kind of 'meaning transfer' that might be formalised.

Compensation seems to take place mainly within each sentence (adding of elements every time the translator wished to put emphasis on any word - for example, '*sad and monotonous*' for '*mélancolique*'), or between two consecutive sentences (through merging or splitting). In some cases, especially in the English translation, the translator tended to include the reader, or tried to establish a link between the reader and himself ('vous vous demanderez' becomes 'you will ask me'). In Russian, there is a frequent transformation of 'vous' (formal 'you') into 'tu' (familiar 'you'), emphasizing the wish from the translator to address the children's readership.

Some compensation features could be described as 'syntactic transfers' operated probably unconsciously by the translators, but they belong to the kind of systematic transfers "advised" to translation scholars: passive for active (and vice-versa), noun phrase for verbal phrase (and vice-

versa), extensive use in Russian of the pronominal form 'sja' which has several functions (pronominal form, but also impersonal form and passive form). As shown in the annex, these syntactic transfers seem to occur in all the target languages considered. Whether these transformations are due to an unconscious cognitive process (pointing toward a possible 'universal language') or whether they express a deliberate choice by the translators, they definitely show that some 'translation operations' are performed at the syntactic level.

Compensation (refer to annex 8.2) takes the following forms:

- splitting of sentences, stressing some elements in the target language (as in the case of the Turkish astronomer), more frequent in English and Russian;
- merging of sentences, especially in German, modifying the style of the text (children's literature is usually characterized by small sentences, easier to read, understand and memorize);
- compensation between sentences, with parts of speech being moved from one to the other sentence, most cases concerning the 'he said', or 'he added' (stylistic preference of the translator);
- repetitions and expansions aiming at putting emphasis on the characters or the situation;
- adaptation to children's readership, mainly through the transformation of the French 'vous' into 'tu';
- lexical and syntactic transfers entailing a change in meaning;
- expansions of deictic and anaphora (reducing the risk of misunderstanding, but also indicating a tendency to explain);
- modifications due to the differences between the two linguistic systems (nominalization, transformation of noun phrases into verbal phrases, indirect/direct discourse transformation).

A more detailed analysis of the transformations between the source and the target sentences is presented in the table 8.2.

Two important features must be noted:

- the adaptation to children's readership in Russian, which takes the form of changes of pronouns (informal 'you' for the formal more polite 'you', expressed in French by 'tu' and 'vous', in German by 'du' and 'sie') and in Russian by 'ty' and 'vy'). The Russian translation also contains a high number of expansions, and explanations, following the didactic norm of children's literature. It also contains many repetitions, hence contributing to the overall cohesion of the text. Nevertheless, in some cases, the didactic repetitions in French are lost in the Russian translation (non-systematic adaptation), as shown in the annex 8.2;
- the similarities between the English and the Russian translations in some instances, while they both diverge from the French source text. One could assume that the Russian translator used the English version as a source text, but this assumption is refuted by other instances where the Russian translation is close to the source French text, while the English translation is different. Several examples are also presented in the annex 8.2.

The high occurrence of splitting and merging (common practice in human translation) cannot be formalized to be applied to machine translation. Software are devised to parse sentences, and merging (or splitting) of sentences is not possible. A comparison of the occurrences of splitting/merging with the size of the respective sentences still provides useful information about the processing speed and the short-term memory capacity of translators. With respect to pre-editing, a maximal size of sentence can be inferred, all sentences longer than the size determined can be automatically selected, and passed to a human translator for pre-editing.

The repetitions and expansions (under the form of explanations, as listed in the annex 8.2) cannot be formalized either. The occurrences are too variable and impossible to systematise. The high occurrence of these expansions is interesting in respect to text typology. Children's literature has a didactic function, and such expansions could be considered as necessary norms to be applied when translation this type of texts.

The most interesting features for this study are the stylistic transformations (active/passive, nominalisation, use of pronouns, and use of the reflexive form –sja in Russian). These can be selected as linguistic correspondences and used as a basis for a transfer system, or considered as 'equivalent structures' between two phenotype languages, in the framework of the Applicative and Cognitive Grammar.

a) (F): J'éprouve tant de chagrin à évoquer ces souvenirs. Il y a six ans déjà que mon ami s'en est allé avec son mouton.

(G): Ich empfinde so viel Kummer beim Erzaehlen dieser Erinnerungen. (...)

Nominalization of "à raconter" in German.

b) (F): Et un jour *il me conseilla de m'appliquer* à réussir un beau dessin, pour bien faire entrer ça dans la tête des enfants *de chez moi*.

(E): And one day *he said to me* : "*You ought* to make a beautiful drawing, so that the children *where you live* can see exactly how all this is."

Indirect discourse translated by a direct discourse in English, with the subsequent necessary changes (I/you).

c) (F): - J'aime bien les couchers de soleil. Allons voir *un coucher de soleil*.

(R): Ja ochen' liubliu zakat. Poidem posmotrim, *kak zoxodit solntse*.

Noun phrase translated in Russian by a verbal phrase ('how the sun sets').

d) (F): Chaque jour j'apprenais quelque chose sur la planète, sur le départ, sur le voyage.

(R): Kazhdyj den' ia uznaval chto-nibud' novoe o *ego* planete, o tom, kak *on pokinul i kak stranstvoval*.

Verbalization in Russian of 'departure' and 'journey'.

e) (F): Il me répondit après un silence méditatif :

(R): On pomolchal v pazdum'e, potom skazal :

The noun phrase ("après un silence méditatif") is expressed in Russian by a verbal phrase ("il se tut dans une méditation") and the chronological order of the two events is established by the adverb 'potom' - thereafter.

f) (F): Mais je ne suis pas tout à fait sûr de réussir.

(E): But I am not at all sure of success.

Nominalization in English of 'to succeed'.

g) (F): Il *me* croyait peut-être semblable à lui.

(E): He thought, perhaps, *that I was* like him.

(G): Er glaubte wahrscheinlich, *ich sei* wie er.

(R): Mozhet byt', on dumal, *chto ija* takoi zhe, kak on.

In English, German and Russian, the dative 'me' is expressed through a complete verbal phrase (= 'il croyait *que je...*')

h) (F): Les baobabs, avant de grandir, *ça* commence par être petit.

(E): *Before they grow so big, the baobabs* start out by being little.

(R): Baobaby sperva, *poka ne vyrastut*, byvaiut sovsem malen'kie.

Both in English and Russian, the infinitive form 'avant de grandir' is translated by a verbal phrase equivalent to 'avant qu'ils ne grandissent'.

6.5.3 Shifts of cohesion and coherence

Coherence is defined by Blum-Kulka (in Venuti, 2000: 298) as a *covert* potential meaning relationship among parts of a text, made overt by the reader or listener through processes of interpretation. Cohesion is defined by the same author as an *overt* relationship holding between parts of the text, expressed by language specific markers. Both coherence and cohesion

contribute to the meaning of any text, and are necessarily present in the source and the target texts. Cohesion creates the semantic unity of the text. Nevertheless, some shifts appear in the translation process, in terms of coherence as much as in terms of cohesion. The target text might be more or less explicit than the source text, or the explicit and implicit meaning potential of the source text might change. Two factors contribute to this modification. First, the grammatical and lexical differences between the two linguistic systems concerned (marking of gender, grammatical system presenting two different aspects for each verb, more extended lexicon in some languages, as for the notion of 'snow') will entail a change in the markers used in the target text. Secondly, the text and discourse processing performed by the translator will also entail a cohesion change. These shifts are due to the interpretation process involved in any translation process (refer to chapter 3). For example, many translations present a higher degree of explicitness, and the 'explicitation hypothesis' can be put forward (human tendency to explain, or 'interpret' the source text).

Shifts in cohesion are easily detectable through a comparison of the source and the target texts, while shifts in coherence are not easily quantifiable. The translator (as reader) uses his world knowledge and subject matter knowledge to understand the text. In Fillmore's terms (1981), this process leads to an 'envisionment' of the text in the reader's mind, which corresponds to the 'virtual translation' notion developed by Neubert and Shreve (1992).

Different kinds of shifts are listed in the annex 8.2, but the most striking examples are presented hereafter. Some words are deliberately changed ('8 jours' becomes 'a week' - even though a week counts only 7 days, '8 days' is actually used in French with the meaning of 'a week'; more surprisingly, the 43 sunsets become 44 in the English edition - considering the fascination of Saint-Exupéry for figures, this transformation definitely modifies the connotative meaning of the original text¹⁸¹). The most interesting lexical change is probably the '*dictateur turc*' who becomes a '*sultan*' in the Russian translation. One should keep in mind that this edition dates back to 1963 (soviet regime). The deliberate mention of 'sultan' can be due to different reasons,

¹⁸¹ *Le Petit Prince* was published in 1943, when Antoine de Saint-Exupéry was 43 years old. Is it the reason for the choice of the 43 sunsets?

a wish to sound 'exotic' - the source text describes the traditional costume as the source of all the troubles of the astronomer whose discovery is ignored, or a choice dictated by political pressures (revolt of the Zaporoghe Cossaks refusing the ultimatum from the Sultan of Constantinople).

- **Change or break in the rhythm:**

French (p.72) : "Vous êtes belles mais vous êtes vides, leur dit-il encore. On ne peut pas mourir pour vous. Bien sûr, ma rose à moi, un passant ordinaire croirait qu'*elle* vous ressemble. Mais à *elle* seule elle est plus importante que vous toutes, puisque c'est *elle* que j'ai arrosée. Puisque c'est *elle* que j'ai mise sous globe. Puisque c'est *elle* que j'ai abritée par le paravent. Puisque c'est *elle* dont j'ai tué les chenilles (sauf les deux ou trois pour les papillons). Puisque c'est *elle* que j'ai écoutée se plaindre, ou se vanter, ou même quelquefois se taire. Puisque c'est ma rose".

The French paragraph is articulated around the anaphora repetitions and the 'e' sound. The German and the English translations keep the repetitions, but in Russian the repetition is missing, 'eje, a ne vac' is repeated twice, but the intertextual reference is lost, hence reducing the cohesion of the text.

English (p.68) : "You are beautiful but you are empty, " he went on. "One could not die for you. To be sure, an ordinary passer-by would think that my rose looked just like you-the rose that belongs to me. But in herself alone she is more important than all the hundreds of you other roses: because it is she that I have watered; because it is she that I have put under the glass globe; because it is she that I have sheltered behind the screen ; because it is for her that I have killed the caterpillars (except the two or three that we saved to become butterflies) ; because it is she that I have listened to, when she grumbled, or boasted, or even sometimes when she said nothing. Because she is *my* rose."

German (p. 57): "Ihr seid schoen, aber ihr seid leer", sagte er noch. "Man kann fuer euch nicht sterben. Gewiss, ein Irgenwer, der voruebergeht, koennte glauben, meine Rose aehnle euch. Aber in sich selbst ist sie wichtiger als ihr alle, da sie es ist, die ich begossen habe. Da sie es ist, die ich unter den Glassturz gestellt habe. Da sie es ist, die ich mit dem Wandschirm geschuetzt habe. Da sie es ist, deren Raupen ich getoetet habe (ausser den zwei oder drei um der Schmetterlinge willen). Da sie es ist, die ich klagen oder sich ruehmen gehoert habe oder auch manchmal schweigen. Da es meine Rose ist."

On the other hand, cohesion is reinforced in the English translation, with the introduction of the 'matter of indifference'. In chapter 3, the English translator alludes to the 'matters of consequence' (chapter 13): 'Mais, bien sûr, nous qui comprenons la vie, *nous nous moquons bien des numéros*' = 'But certainly, for us, who understand life, *figures are a matter of indifference*'.

- **Lexical or syntactic modifications entailing a change in meaning**

1) (F): C'était pour moi une question de vie ou de mort. J'avais à peine de l'eau à boire pour huit jours.

(E) : It was a question of life or death for me: I had scarcely enough drinking water to last a week.

(R) : Ia dolzhen by ispravit' motor ili pogibnut'. Vody u menia edva khvatilo by na nedeliu.

Both in English and Russian, the '8 days' are replaced by a week, '8 jours' having precisely this meaning in French. The German version kept the precise '8 Tage'.

"It was a question of life or death for me " is translated in Russian by "I had to repair the motor or to die". The Russian expression is more literal, and adds a repetition in the text, making it more accessible to a younger readership.

2) (F): Je fis remarquer au petit prince que les baobabs ne sont pas des arbustes, mais *des arbres grands comme des églises* et que, même s'il emportait avec lui tout un troupeau d'éléphants, ce troupeau ne viendrait pas à bout d'un seul baobab.

(E): I pointed out to the little prince that baobabs were not little bushes, but, on the contrary, *trees as big as castles*; and that even if he took a whole herd of elephants away with him, the herd would not eat up one single baobab.

(G): Ich erklarte dem kleinen Prinzen ausführlich, dass Affebrotbaeume doch keine Stauden sind, sondern *kirchturmhohe Baeume*, und selbst wenn er eine ganze Herde Elefanten mitnaehme, wuerde diese Herde nicht mit einem einzigen Affenbrotbaum fertig werden.

(R): Ia vozprazil, chto baobaby ne kusty, a *ogromnye derev'ia, vyshinoi s kolokol'niu*, i ecli dazhe on privedet tseloe ctado slonov, im ne c"ect' i odnogo baobaba.

Particularizing in German and Russian (church tower for the church itself), while in English the image is altogether different, the baobabs being compared to castles (different connotative meaning).

3) (F): Car je n'aime pas qu'on lise mon livre à la légère.

(G): Denn ich moechte nicht, dass man mein Buch leicht nimmt.

"Lire" is translated by "to take". The translator might have missed the difference in French between the two expressions ("prendre à la légère" being a fixed locution), or he might have consciously used that reference to give more weight to the remark. On the other hand, 'leicht lesen' does not carry the 'lightly' connotation, and would have changed the meaning of the sentence.

4) (F): Les grandes personnes sont comme ça.

(R): Uzh takoi narod eti vzroslye !

Emphasis in Russian by the elliptic nominal phrase ('*What a people these grown-ups !*')

5) (F): S'il s'agit d'une *brindille* de radis ou de rosier, on peut la laisser pousser comme elle veut.

(E): If it is only a *sprout* of radish or the *sprig* of a rose-bush, one would let it grow wherever it might wish.

(G): Wenn es sich um einen *Radieschen* - oder *Rosentrieb* handelt, kann man ihn waschen lassen, wie er will.

Both English and German require two different words for 'brindilles', and necessitate a distribution of the noun phrases. Note: the Russian sentence follows the French structure (two subordinate clauses of one noun).

• **Coherence shifts**

1) (F): Mais personne ne l'avait cru à cause de son costume.

(E): But he was in Turkish costume, and so nobody would believe what he said.

(R): No nikto iemu ne poveril, a vce potomu, chto on byl odet' po-tupetskii.

In English, the emphasis is put on the reason why he was not believed, by the expansion into a verbal phrase, and by the place (beginning of the sentence) of this phrase.

Explanation and expansion both in English and Russian, by the mention of the Turkish costume, while in French, it is understood by the context (Turkish astronomer).

2) (F): L'astronome refit sa démonstration en 1920, dans un habit très élégant. Et cette fois-ci tout le monde fut de son avis.

(E): So in 1920 the astronomer gave his demonstration all over again, dressed *with impressive style and elegance*. And this time everybody *accepted his report*.

(R): В 1920 году тот астроном снова доложил о своём открытии. На этот раз он был одет *на последней mode*, и все с ним согласились'.

The emphasis is put in English on the elegance of the astronomer's dress, by the adding of another adjective. 'Etre de son avis' is expanded with the introduction of the 'report' which is not mentioned in French.

In Russian, the elegance of the astronomer is rendered by 'dressed according to the latest trend'. This interpretation gives some interesting information about the definition of 'elegance' for the Russian translator.

3) (F): *Alors* seulement elles croient le connaître.

(E): *Only from these figures* do they think they have learned anything about him.

Semantic expansion of 'alors' rendered by 'only from these figures' (implicit in French, emphasized in English).

4) (F): Mon ami ne donnait jamais d'explications.

(E): My friend never explained anything *to me*.

(G): Mein Freund hat *mir* nie Erklärungen gegeben.

(R): Moi друг никогда *mne* ничего не обяснял.

Systematic adding ('me') in English, Russian, and German, although it is not explicit in French, the sentence can be understood as a habit of the Little Prince, whoever he is addressing to.

- 5) (F): Mais moi, malheureusement, je ne sais pas voir les moutons *à travers les caisses*.
(E) : But I, alas, do not know how to see sheep *through the walls of boxes*.
(G) : Aber ich bin leider nicht imstande, *durch die Kistenbretter* hindurch Schafe zu sehen.
(R) : No ia, k sozhaleniiu, ne umeiu uvidet' barashka *skvoz' smenki iashchika*.

Adding in all target languages of 'the walls/plates' of the boxes, which is only implicit in French.

- **Modifications due to differences in the linguistic systems**

- 1) (F): D'où viens-tu mon petit bonhomme ?
(R): Otkuda zhe ty priletel, malysh ?

Semantic adding ("d'où viens-tu en volant") necessary in Russian, on the ground of the complex system of verbs of movement. In French, the means used to come (or go) is implied by the context, the verb itself being semantically vague, while in Russian there is a different verb for each kind of movement (on foot, by car, on bicycle, by boat...)

- 2) (F): Il me répondit après un silence méditatif :
(R): On pomolchal v pazdum'e, potom skazal :

The noun phrase ("après silence méditatif") is expressed in Russian by a verbal phrase ("il se tut dans une méditation") and the chronological order of the two events is established by the adverb 'potom' - thereafter.

- 3) (F): Il *me* croyait peut-être semblable à lui.
(E): He thought, perhaps, *that I was* like him.
(G): Er glaubte wahrscheinlich, *ich sei* wie er.
(R): Mozhet byt', on dumal, *chtoija* takoi zhe, kak on.

In English, German and Russian, the dative 'me' is expressed through a complete verbal phrase (= 'il croyait *que je...*')

4) (F): Chaque jour j'apprenais quelque chose sur la planète, sur le départ, sur le voyage.

(R): Kazhdyj den' ia uznaval chto-nibud' novoe o *ego* planete, o tom, kak *on pokinul i kak stranstvoval*.

Verbalization in Russian of 'departure' and 'journey'.

6.5.4 Convergence between the target languages syntax and/or lexicon through translation

1) (F): Les baobabs, avant de grandir, *ça* commence par être petit.

(E): *Before they grow so big, the baobabs start out by being little.*

(R): Baobaby sperva, *poka ne vyrastut*, byvaiut sovsem malen'kie.

Both in English and Russian, the infinitive form 'avant de grandir' is translated by a verbal phrase equivalent to 'avant qu'ils ne grandissent'.

2) (F): Quand un astronome découvre l'une d'elles, il lui donne pour nom un numéro. Il l'appelle par exemple : "l'astéroïde 3251".

(E): When an astronomer discovers one of these he does not give it a name, but only a number.(...)

(R): Kogda astronom otkryvaet takuiu planetku, on daet ei ne imia, a prosto nomer. Naprimer : asteroid 3251.

Both English and Russian versions present a totally similar semantic interpretation for "il lui donne pour nom un numéro", and the syntax of both sentences is identical.

3) (F): Vous imaginez combien j'avais pu être intrigué par cette demi-confiance sur 'les autres planètes'.

(E): You can imagine how my curiosity was aroused by this half-confidence about the "other planets"

(R) : Mozhete sebe predstavit', kak razgorelos' moe liubopytstvo ot etovo polupriznaniia o "drugix planetax".

Both the English and Russian versions render "j'avais pu être intrigué" by "ma curiosité a été éveillée". Such a similarity might be explained by the fact that the Russian translator used the English translation. Otherwise, one might wonder about such a convergence of semantic interpretation.

4) (F): Et, sur les indications du petit prince, j'ai dessiné cette planète-là.

(E): So, as the little prince described it to me, I have made a drawing of that planet.

(R): Malen'kii prints podrobno mne vce opisal, i ia narisoval etu planetu.

Use of a verbal phrase both in English and Russian for 'on the indications'. Noun phrase in German.

6.5.5 Units of translation

Apart from purely stylistic features of minor importance (such as 'and' for a comma or a full stop, and vice versa), it seems that three translators processed *Le Petit Prince* sentence-by-sentence, with very little processing at the level of the text itself (let alone chapters, or even paragraphs), even though the text as a whole is taken into account. There are relatively few splitting, mergings and reports from one sentence to another. The context is clearly used as an input in the determination of the 'equivalent' target sentences, for example when a noun is expanded into a noun-phrase, or when a pronoun is replaced by its antecedent, but the translation processing is done at the sentence level. This observation contradicts some theories about translation, particularly the text-oriented approach and the communicative one, according to which translation being a communicative process, the unit of communication (hence of translation) is the discourse or the text, as opposed to the syntactic structure of language based on the clause and sentence (N. E. Enkvist, 1978). For Enkvist, the discrepancies between translations (close to the ST syntactic patterns or free from these patterns) depend on the choice made by the translator, who may have to change the syntax drastically in order to be faithful to information dynamics. Similarly, by opting for a closer syntax, information dynamics is likely to be lost. Still, "information dynamics should be counted as one of the criteria of translational adequacy" (Enkvist, 1978: 181). Although the published translations used for this study may not be the best ones (but there is no 'best' translation of any work, with the exception of an ideal construction), they obviously fulfilled one of their aims, since each of them represents the published version that was read by most of the (respectively) British, German and Russian

readers. I used *Le Petit Prince* in classes (adult students) in Great Britain, Germany and Russia, and I realised that almost all students had read this book (translated in their own native language) and had a similar interpretation of the text. Much information dynamics (which is central in text linguistics) should be lost by a sentence-by-sentence translation, but that does not seem to be the case with *Le Petit Prince* and the translations considered. These contradictory data may be interpreted in different ways:

- Information dynamics – although a central issue – is not necessarily generated at the text level. The sentence may carry it as well;
- Much of the information dynamics lost by a sentence-by-sentence translation is actually restored by the reader, hence the necessity to study more attentively the process of interpretation and cognition;
- This type of text (children’s literature) presents very specific features, in which information dynamics does not have the same characteristics as in other kinds of texts. Rhetorical cohesion is implemented through the repetition of simple sentences at regular intervals and does not suffer from a sentence-by-sentence translation.

Chapter X: The first of them was inhabited by a king. (...) "The grown-ups are very strange", the little prince said to himself, as he continued on his journey."

Chapter XI: The second planet was inhabited by a conceited man. (...) "The grown-ups are certainly very odd", he said to himself, as he continued on his journey".

Chapter XII: The next planet was inhabited by a tippler. (...) "The grown-ups are certainly very, very odd", he said to himself as he continued on his journey".

Chapter XIII: "The fourth planet belonged to a businessman.(...) "The grown-ups are certainly altogether extraordinary", he said simply, talking to himself as he continued on his journey".

(“Les grandes personnes sont bien étranges”, se dit *Le Petit Prince* en lui-même durant son voyage [chapter X] ; “Les grandes personnes sont décidément bien bizarres”, se dit-il simplement en lui-même durant son voyage [chapter XI] ; “Les grandes personnes sont décidément très très bizarres”, se disait-il en lui-même durant le voyage [chapter XII] ; “Les

grandes personnes sont décidément tout à fait extraordinaires”, se disait-il simplement en lui-même durant le voyage [chapter XIII])

Three different translators seem to have used a similar technique to translate the same text into three different languages. One would usually expect a rephrasing of some paragraphs, or the use of elaborated compensation techniques at paragraph or chapter level¹⁸².

6.6 Machine translation of *Le Petit Prince*

The machine translation produced by the software package used for this analysis also presents some very interesting features, although, as stated above, only the English and German versions could be assessed, since the Russian package was not yet available. The transfer system displays, in a very obvious way, its limits in the case of English. Almost every sentence would require post-editing. It seems that the assumption, according to which French and English are syntactically close enough to be easily translated by a ‘word-for-word’ process complemented by a grammatical matrix, does not hold, even for a text with a rather simple lexicon and syntax. The German version is comparatively better, or at least sounds ‘more German’. This characteristic could be explained by the following reasons:

- convergence of French and German syntax;
- better parsing system, which would account for a better grammatical formalisation;
- transfer system better adapted to this language pair.

The software (Power Translator Pro) was designed by the American company Globalink (bought over by Lernout and Hauspie) and is based on a technology called Barcelona. The translation programme belongs to the category of transfer system model. The document on the Barcelona technology presented in the annexe 8.7 gives some details on the architecture of the system, although no data on the actual programming language are available, for obvious confidentiality and commercial reasons.

¹⁸² Hervey and Higgins, 1992.

The machine translations display both in English and German a range of errors as presented in the annexes 8.3 and 8.4. Only the main types of errors will be presented, considering the necessarily limited size of this dissertation, but the analysis of these errors is important in many respects. On the one hand it can help to improve this product specifically and more generally all transfer systems; on the other hand, the comparison of the systematic errors (or untranslated lexemes) with the human translation may help to identify the translational operations and processes at hand that would require formalisation.

- Some errors are rather minor, such as the untranslated words ('décoiffée', 'démodent', 'margelle', or 'assoiras' - 'asseoiras', the irregularity of this verb in the future tense has obviously been missed by the programmers). It would be sufficient to update the lexical database of the software to get a translation of these words. The main point to be noticed in this respect is the fact that some words are translated into German, but not into English (*sotte*), and vice versa (*Histoires Vécues*, businessman). This tends to prove that the software was designed on a language pair basis, and the lexical databases are not coherent.
- In German, the conjunction 'parce que' is not recognized by the parser, 'parce' being left untranslated and 'que' being translated by a relative pronoun adjusted to the case assumed by the programme (das/den).
 - (F): Parce que chez moi c'est tout petit.
 - (G): *Parce das* bei mir dies ganz klein is.
 - (F): Parce que les ivrognes voient double.
 - (G): *Parce, den* die Trinker dopplet sehen (3).
 - (F): Parce que les fleurs sont éphémères.
 - (G): *Parce, den* die Blumen Eintagsfliege sind.
- In many cases, verbs presenting similar forms at different tenses (present/past historic for example) are not disambiguated, even though a simple parsing of the previous and following sentences might have been sufficient to determine the tense (if all other verbs are

in the past tense, the chances for that verb to be also in the past tense are quite high). Such an analysis would be possible, and the error rate would be rather low. It is feasible to load the database with a list of all verbs presenting similar forms at different tenses, and a statistical computation of the verbs syntactically close to the ambiguous one would allow for a better quality machine translation.

- a) (F): Quand je réussis enfin à parler, je lui *dis*.
 (E): when I finally *succeed* in speaking, I *tell* to him.
- b) (F): Il *rougit*, puis *reprit*.
 (E): He/it *blushes*, took then (the second verb is rightly put at the past tense).
- c) (F): Sur tout ça? dit le petit prince. Sur tout ça, répondit le roi.
 (E): On all that ? *tells* the small prince. - On all that...*replies* the king.
- The unrecognised locutions ('à mon tour', 'les grandes personnes', 'du premier coup d'œil', and so on) presented in the tables 8.3 and 8.4 could easily be loaded. But it would make sense to list locutions according to the text type. Some locutions are highly unlikely in some contexts, some others are to be expected. In that respect, machine translation would greatly benefit from research on text typology (based often on statistical occurrences). In the case of a text like *Le Petit Prince*, it would be necessary first to parse the whole text and to load the database with the locutions likely to occur. Another methodology would consist in loading all possible locutions (based on a standard dictionary, such as *Le Robert*) but such a methodology would heavily burden the database (hence the memory) and overload the parsing process (each word being likely to occur in at least one locution in any language, the parser would have to stop at each word, and check the string of words in the sentence to diagnose any possible locution).

to my tower/à mon tour - the big people/les grandes personnes - of the first squint/ vom ersten kurze Blick/du premier coup d'oeil - if he/it pleases you/s'il vous plaît - to the rising of the day/au lever du

jour - von Seite zu lassen/laisser de côté - Tote unter Androhung/sous peine de mort- nach der Art von den Erzählung von Feen/à la façon des contes de fées

- In English, the highly complex problem of anaphora resolution (as well as the similar problem of untranslated or badly translated possessives) is simply overlooked by the software used for this study. As shown in the annexes, pronouns and possessives are left untranslated, the reader being prompted to choose between 'she/it', or 'he/it', or 'himself/itself', or 'his/her/its'.

but *he/it* didn't answer me - retiring my sheep of *his/her/its* pocket, *he/it* immersed *himself/itself* in the contemplation of *his/her/its* treasure - I endeavored therefore of in to know longer - but where want you that *he/it* goes? - *he/it* calls him for example "the asteroid" 3251" - *he/it* had made a big demonstration of *his/her/its* discovery then - I got to *his/her/its* range - I immediately glimpsed a gleam, in the mystery of *his/her/its* presence retiring my sheep of *his/her/its* pocket, *he/it* immersed *himself/itself* in the contemplation of *his/her/its* treasure - it is that *his/her/its* planet of origin was hardly bigger than a house (4) - *he/it* had made a big demonstration of *his/her/its* discovery then - but no one had believed him because of *his/her/its* costume (4) - what is the sound of *his/her/its* voice ?

Anaphora resolution is certainly one of the most challenging issues in natural language processing. A transfer system is obviously totally inadequate for such an analysis, even though, the software used for this study seems more adequate to the French/German language pair than to the French/English one. In some cases the resolution itself is erroneous. Unsolved anaphora resolution would probably be preferable to erroneous resolution (which can be semantically highly misleading if the targeted readership does not have any access to the source language), but commercial standards usually require a full translation of the source text. One cannot expect the reader to select the right word in a list of two or three options. It is clear from the examples selected that anaphora resolution cannot be considered at the surface syntactic level. A proper and complete semantic parsing is necessary to determine the word the anaphora refers to.

- The tables 8.3 and 8.4 present all kinds of errors, as well as their analysis, and they show how a transfer system proceeds, taking basically a source word, and 'transferring' it. For example, the fact that 'de/des' are in most cases wrongly analysed as possessives when they are indeterminate articles proves that such a translation system actually processes sentences at a syntactic level, but without any proper semantic interpretation.

(F): Et ce n'est pas sérieux de chercher à comprendre pourquoi elles se donnent tant de mal pour se fabriquer *des* épines qui ne servent jamais à rien ?

(EMT): And this is not serious to try to understand why they give themselves so much pain to manufacture *itself/themselves of* thorns who serve never to anything ?

(EHT): And is it not a matter of consequence to to try to understand why the flowers go to so much trouble *to grow thorns* which are never of any use to them ?

6.7 Comparison of human and machine translation of *Le Petit Prince*

Despite the numerous criticisms developed above, and despite the range of errors, it seems that in some cases both human translators and software produced similar target sentences in terms of syntax (annexes 8.5 and 8.6). Most of the sentences concerned are rather short, but the convergence is sometimes striking

- (MT): 'I have serious reasons to believe that the planet from which the little prince came is the asteroid known as B612'
(HT): 'I have serious reason to believe that the planet from where came the little prince is the asteroid B612'.
- (MT): And that is how I made the acquaintance of the little prince.
(HT): And this is how I made the acquaintance of the small prince.
- (MT): 'In der Tat gab es auf dem Planeten des kleinen Prinzen wie auf allen Planeten gute Gewaechse und schlechte Gewaechse'
(HT) 'Und in der Tat, auf dem Planeten des kleinen Prinzen gab es wie auf allen Planeten guten Krauetern und schlechten Krauetern'.
- (MT): Der kleine Prinz stoerte meine Ueberlegungen von neuem.
(HT): Der kleine Prinz stoerte meine Ueberlegungen von neuem.

6.7.1 Units of translation

One of the conflicting tenets between human and machine translation concerns the 'unit of translation'¹⁸³, which is – allegedly – large in the case of human translation¹⁸⁴, but should be

¹⁸³ Although there is no single concept of a UT corresponding to a linguistic unit (Bennett, P., 1994: 13; Larose, 1989: 22-23), the definition of UT selected for this discussion is: « the smallest segment of the utterance where the cohesion of signs is such that they cannot be translated separately » (Vinay and Darbelnet, 1958:16).

¹⁸⁴ This view is not unanimous, and W. Haas, for example, advocates the smallest possible unit of translation, in the case of human translators as well: « The discipline of translation consists very largely in choosing the smallest possible unit that will admit of adequate matching. (...) (The translator) chooses what units to translate, and he chooses such units as correspond or can be made to correspond to one

kept as small as possible in machine translation¹⁸⁵. Surprisingly enough, one of the results of the study of the translations of *Le Petit Prince* is that the discrepancy – as far as the length of the units of translation is concerned – between human and machine translation is quite small. The software used for this study processes the input text sentence by sentence, and the published translations display the same pattern, with the exception of a few sentences, which are presented in the annexe 8.2.

The table 8.8 gives the number of sentences in each chapter for each version (French, English, German, Russian). Some variations in the numbers of sentences are simply due to different punctuation (eg a comma for a full stop). The figures provided indicate the arbitrary delimitation of ‘a sentence’ by a capital letter at the beginning and a full stop, an exclamation mark or a question mark at the end of the string of words. Moreover the colon was counted as delimiting a sentence when introducing a dialogue, while sentences proffered by the characters, and followed by a comma and ‘*he added*’, or a comma and ‘*-dit-il*’ were counted as part of the whole sentence. This counting system is purely arbitrary (but the same rule was applied to all versions of the text), but it matches the most common counting systems of translation software. Such a system is certainly arguable, since it does not necessarily correspond to the linguistic definition of ‘a sentence’¹⁸⁶, but it nevertheless gives interesting information about the ‘unit of translation’ as processed (consciously or unconsciously) by the translators.

6.7.2 Anaphora resolution

The machine translation presents a high percentage of wrong anaphora resolution, and there is hardly any pronoun correctly translated. The reasons for such errors are manifold : wrong assumptions about the gender of a pronoun in French (‘lui’ may refer to a masculine or a

another. He tries to keep the size of his translation units to a minimum. But he cannot, generally, avoid having to deal with units larger than the word. » (Haas, 1968: 107).

¹⁸⁵ « There are good reasons for keeping the UT (in the sense of translation atom) in MT as small – and hence as manageable – as possible. Adopting a larger unit may be less efficient, and is not guaranteed to improve translation quality. » (Bennett, 1994: 18)

¹⁸⁶ ‘Sentence: the largest unit of grammar, or the largest unit over which a rule of grammar can operate’, in the Oxford Concise Dictionary of Linguistics, Oxford University Press, 1997.

feminine noun/"l'" can refer to a feminine noun, "le" can be translated in English by 'him' or 'it'), erroneous analysis based on the ending of the past participle, as in the following example e).

The errors are similar in German, but with a lower frequency.

- a) (F): il *lui* donne pour nom un numéro ('lui' refers to the planet discovered, indicated a few sentences above)
(MT): he gives *him* for name a number.
(HT): when an astronomer discovers one of these, he does not give *it* a name, but only a number.
- b) (F): il *l'*appelle par exemple l'astéroïde 3251 ('l' is again the planet)
(MT): he/it calls *him* for example "the asteroid" 3251"
(HT): he might call *it*, for example, "Asteroid 325".
- c) (F): Alors seulement elles croient le connaître ('le' refers to 'a new friend')
(MT): then only they believe to know *it*.
(HT): Only from these figures do they think they have learned anything about *him*.
- (d) (F): Si j'essaie ici de *le* décrire, c'est afin de ne pas *l'*oublier ('le' and 'l' refer to the little prince)
(MT): if I try to describe *it* here, it is in order to not to forget *him* (the second pronoun is rightly translated, while the first one is wrong = inconsistency).
(HT): If I try to describe *him* here, it is to make sure that I shall not forget *him*.
- e) (F): Le petit prince ne renonçait jamais à une question, une fois qu'il *l'*avait posée ('l' = une question).
(MT): The small prince never gave up a question, once he had put *her*.
(HT): The little prince never let go of a question, once he had asked *it*.
- f) (F): Mon dessin numéro 1. *Il* était comme ça.
(MT): *Er* (= die Zeichnung) war so

(HT): Meine Zeichnung Nr. 1. So sah *sie* aus.

The software does not perform a complete analysis at the sentence level, but replaces the pronoun with a predetermined translation.

6.7.3 Reflexivation

a) (F): Quelque chose *s'était cassé* dans mon moteur.

(EMT): Something had broken *himself* in my motor.

(EHT): Something was broken in my engine.

(GMT): Etwas war in meinem Motor gebrochen.

(GHT): Etwas war an meinem Motor kaputtgegangen.

b) (F): Le premier soir je *me* suis donc endormi

(EMT): The first evening *me* I am lulled therefore.

(EHT): The first night, then, I went to sleep (on the sand).

(GMT): Den ersten Abend bin ich also (...) eingeschlafen.

(GHT): Am ersten Abend bin ich also im sand eingeschlafen.

c) (F): Mais je fus bien surpris de *voir s'illuminer* le visage de mon jeune juge.

(MT): I was surprised *to see to illuminate* the face of my young judge.

(HT): I was very surprised to see a light break over the face of my young judge.

(GMT): Aber ich wurde ueberrascht von das Gesicht meines jungen Richters *aufleuchten zu sehen*

(GHT): Und ich war hoechst ueberrascht, als ich das Gesicht meines jungen Kritigers *aufleuchten sah*.

The German translation software is clearly better conceived to tackle reflexivation. It seems that reflexive verbs are stored as such in the German data base, while the translation into English is obviously based on a word-for-word architecture. The software (as far as English is concerned) is not able to spot a reflexive verb when the pronoun is separated from the main verb.

6.8 Conclusion: human translation, machine translation, and the transfer/interlingua debate

The human translations of *Le Petit Prince* retained for this study support some of the approaches presented in chapter 3. Assuming that most translators (after the first stages of training) tend to translate 'automatically' (this term is here taken in the sense of 'by the use of mostly unconscious cognitive processes') through an interpretive processing of the ST, the translations as analysed above show that:

- translation is definitely a communicative process (as exemplified more particularly by the attempts to include the reader - use of informal 'you' in the translation, especially in Russian, even though the source sentences considered did not contain this feature);
- translation deals with texts as wholes. The translations of *Le Petit Prince*, whether in English, German or Russian are first and foremost translations of a text (which belongs to children's literature). The target texts belong to the respective text typologies in each of the target languages;
- there seems to be an equivalence between the source and target texts, not only in terms of text typology, but also in terms of socio-cultural settings. The respective target texts clearly address a similar readership in each of the target languages. Many sentences are expanded (through several repetitions) in English and Russian, hence following the didactic and linguistic requirements of the Russian fairy-tale. Similarly, the formal French 'vous' is often transformed into 'ty' in Russian – French 'tu' (children's readership). But the text is sometimes adapted to conform the expected political and social norms, as demonstrated by the dictator/sultan transformation in Russian.

These observations should nevertheless be tempered by the following remarks:

- in the case of this specific kind of texts, a sentence-by-sentence approach seems to be perfectly possible, even successful in many respects (hence contradicting one of the tenets of translation);

- there seems to be a systematic expansion of the source text by the translators, whatever the target language. This expansion usually takes the form of repetitions, or of the expansion of an anaphora (explanation process). An intermediate solution to the problem of anaphora resolution in machine translation would consist in the replacement of the pronoun with the complete antecedent. In the long term, the computation of the 'reference', for example through the use of a powerful artificial language could allow a more accurate analysis and translations of personal and relative pronouns.
- there seems to exist a convergence between syntactic structures of the source text and the respective target texts. Whether this convergence is due to a "source language contamination" (as would prove the differences between the German and the Russian translations for example), or whether it is due to the 'dynamic equivalence', or whether it expresses a correspondence between different syntaxes remains to be determined. The Applicative and Cognitive Grammar maintains that all natural languages (or phenotypes) are variations of the Phenotype Grammar. The convergences noticed during the translation operation should be analysed more specifically in order to determine if they actually express the 'Phenotype' language, or whether they express contaminations phenomena between source and target languages. Even in the case of a simple contamination, such results can be used as a possible basis for the development of machine translation software, more particularly in the case of sublanguages. Further similar analyses of human translation of other fairy-tales into several languages (preferably with different source and target languages) should permit the determination of convergent syntactic structures, hence the definition of a Genotype language specific to fairy-tales.

Another interesting result is the outcome of the comparison between the human translations and the machine outputs. The analysis of machine translation showed in particular that a transfer system is able to produce a target sentence close to the sentence produced by a human translator, as developed in the above paragraphs and as exemplified in the annexes 8.5 and 8.6. But the same analysis also showed the limits of such a system (unable to tackle locutions,

anaphora, deictic, tenses, and so on). Most of the similar human/machine sentences concerned are short, with few dependent clauses, but this feature is definitely very important for the development of machine translation. Machine translation usually fails to produce meaningful target sentences, but in some instances a series of syntactic operations performed on the source sentence seems sufficient to produce an acceptable target sentence. This observation remains to be correlated with other studies on different text types, but it tends to support the current computer approaches to machine translation. It is now necessary to assess the human/machine convergence, with respect to the size and composition of the sentences to be translated. How long a sentence should be to be translatable by a transfer system? What syntactic constructions are 'reasonably' translated by a transfer system? Once the limits of a transfer system are established for a specific sublanguage (in our case, fairy-tale), the next step for computer linguists is the determination of an interlingua system, and the comparison of the respective translations through the transfer system and through the interlingua system. If one follows the assumptions of Applicative and Cognitive Grammar, the comparison of the source text with its human translations in several languages should permit the determination of the different 'levels of interpretation' as well as the determination of the Genotype language. In that respect, the translation shifts presented in the above paragraphs are the expression of the deeper cognitive operations performed by translators. In our study, the analysis of the shifts of translation showed several tendencies:

- a tendency to expand (explain) in the process of translation. The context (previous sentence or paragraph, or background knowledge) clearly prompts translators to expand pronouns, or even nouns;
- a tendency to translate sentence-by-sentence, while taking into account the whole text when analysing and translating;
- a tendency to 'interpret' some parts of the text in a similar fashion, leading to similar syntactic and lexical choices in the target languages.

These tendencies seem to be independent of the target language considered, and appear to correspond to some cognitive processes (at least in the case of the translation of this particular sublanguage of fairy-tales).

In theory, cognitive invariants (as assumed by Applicative and Cognitive Grammars) carry the 'meaning' that is supposed to be transferred from a source text to a target text in the process of translation. The determination of all the systematic syntactic transformations (active/passive, and so on) is a first step towards the determination of the underlying 'universal' sign system, as assumed by Shaumyan (1977, 1990) and Desclés (1989, 1990, 1996). Nevertheless, such an analysis must take into account the following remarks:

- the text type determines the structure of the source and target texts, as well as the syntax and the lexicon,
- a systematic transfer operation expressed in one human translation does not necessarily entail a systematic operation at the level of the natural languages involved. It can also express a unique cognitive tendency of this particular human being (each human being has his/her own idiolect, determined by his/her culture, background, experience, linguistic abilities and so on);
- there is no such thing as an absolute, 100% accurate translation. Translation being an expression of human cognition, it is necessarily variable.

Comparison of the source and target texts is useful as far as applicative and cognitive grammar (ACG) is concerned, in the sense that different syntactic constructions in different target languages can express (at the natural language level) underlying cognitive invariants. The reflexive marker 'cja' in Russian is traditionally presented as the expression of a pronominal form, or as a passive construction. It can actually express many more notions in context. A statistical study of the occurrences of such a form, put in parallel with the source sentences it 'translates' would be extremely useful in designing new translation software, based on the articulation of the levels of interpretation, rather than on a straightforward transformation between source and target surface structures. But Applicative and Cognitive Grammar can only

be useful to design translation software aiming at rendering the 'content'. At the present stage of its development, it cannot account for the creative dimension of translation, nor for the 'dynamic equivalence' principle. It can nevertheless represent the first stage towards the development of new tools, such as machine interpreters (allowing two speakers who have no common language to communicate). One must distinguish between the form of the discourse, and the content. Human translation will always be necessary to treat the form, while machine translation should probably be developed - at least for the coming fifty years - with the aim of expressing the content of a text. With this distinction in mind, research on applicative and cognitive grammars, and on the articulation of levels of representation, is likely to be the next centre of focus in cognitive sciences. The results of all studies on sublanguages (whether technical aeronautical documentation or fairy-tales), which have already been very important for the development of transfer systems, now need to be applied to the development of the next generation of interlingua systems, in order precisely to improve the 'form of the content'. Interlingua approaches, such as the Applicative and Cognitive Grammar, tend towards the expression of the 'essence' of a sentence/paragraph/text. The determination of a text type, and the definition of all the features of a precise sublanguage have to be incorporated to any machine translation software. The shifts from 'vous' to 'tu' do not correspond to a linguistic transfer, but is the expression of the 'function equivalence' – a concept known by all translation scholars, but too often discarded by the computer linguists. 'Tu' is not a translation of 'vous' – at least not in lexicon databases, but in a specific sublanguage, for a specific text type, and for a specific target language (as the embodiment of a target culture), it becomes not only an acceptable translation, it is the proper – and expected – translation.

7 CONCLUSION

Two separate worlds share one word: translation. But scholars involved in translation studies, and more particularly in literary translation, have much trouble communicating with computer linguists. Debates in many conferences devoted to translation show too often how irreconcilable human translation and machine translation seem to be, machine translation being usually presented by translation experts as ‘alien to translation’. On the one hand, human translation is acknowledged as a creative process, considering the central role played by the translator (who takes into account the source and target languages and cultures when transforming a source text into an ‘equivalent’ target text). On the other hand, “machine translation considers all translational phenomena as susceptible of being logically described, schematised and formalized” (Gorlée, 1994:12). The lack of consensus among scholars of the two communities, as much as the limited advances in computing linguistics, and more particularly in machine translation (which is still in its infancy, if one compares this field with human translation), have been detrimental to the development of interdisciplinary studies in human and machine translation, more particularly as far as the application of machine translation to literary texts is concerned. Translation of literary texts deals with the determination of the linguistic, social and cultural factors that will contribute to the creation of a new entity (the target text) said to be in an equivalence relationship with the source text, while keeping the ‘meaning’ (or sense) of the source text. This ‘meaning transfer’ is done by the translator through a succession of complex cognitive operations, which need to be analysed and formalised, if one wishes to build a software programme able to ‘translate like a human being’. Such a research has not been attempted so far, considering the complexity of the human translation process, as well as the variety of the disciplines involved. Shortly after the dawn of machine translation, computer linguists became aware of the structural limits of their tools, and all machine translation products currently in use are devised to tackle (exclusively) technical texts, and are usually

based on a transfer system, matching source and target lexicons and syntactic structures, while avoiding the core issue of meaning transfer. In particular, literary texts have been excluded from machine translation, given the particular constraints that operate in the production of this type of text. But this very limitation of machine translation to non-literary texts prevents any real breakthrough in automatic translation, which must now tackle the issue of meaning (traditionally the object of study of semantics). As of today, no one would sensibly claim that fully automatic high quality translation is possible in the foreseeable future. The examples of the automatic translation of *The Little Prince* given in this dissertation prove that most current machine translation systems are still faced with serious difficulties such as anaphora resolution, disambiguation or tense analysis. Numerous teams of computer linguists all over the world are dedicated to the improvement of the existing systems, which are usually based on a transfer approach. In the fifties, computer linguists were already discussing the pros and cons of the direct and transfer systems on the one hand, and of the interlingual approach on the other hand. Today, the debate is still animating most conferences on machine translation. The practice (or is it rather the weight of tradition?) and the mechanical constraints of machine translation advocate the transfer approach, while theoretical studies on translation put forward philosophical arguments in favour of an interlingual system. Human beings have been in search of a 'universal' language for thousands of years, and the interlingua architecture touches the right emotional chord. Even practical and economical reasons tip the scale on the interlingua side: in principle, such a system requires only translation to and from the interlingua, while a transfer system necessitates translations between each single pair of languages considered. There are nowadays more and more advocates for the interlingual approach, and a few companies even designed interlingua systems (such as the Pivot system of NEC, or the Rosetta system of Philips). Some software (such as the Atlas I of Fujitsu) are a mixture of the transfer architecture and of the interlingual principle. Moreover, many computer linguists are increasingly aware of the too high emphasis put on the linguistic analysis in translation software, and some programmes now integrate knowledge databases in an attempt to mimic the actual translation process as performed by a human being. But many projects concerned with

fully automatic translation have been abandoned, or at least put aside, and software producers focus their attention and efforts on the development of translation tools (translation memories, computerised dictionaries, and spell checkers). Translation software as such still requests pre- and/or post-editing, if one wishes to get a 'sensible' output text, but as long as such operations will be needed, machine translation will fall short of its expectations, and human translators are unlikely to be ousted by machines. On the other hand, even human translators now admit that they are unable to tackle the ever increasing volume of translation required in the world. The admission of new members in the European Union is already creating serious trouble to the European institutions, unable to satisfy the translation requirements, considering the volume of texts and treaties, the number of language combinations, and the overall cost. Machine translation, whether appreciated or not, has simply become necessary, and requires fast improvement and development. A better understanding of the human translation process is urgent. As a contribution to this endeavour, this thesis offers significant insights into a specific interdisciplinary conjunction of activities (translation studies, representation of meaning and cognitive processes, machine translation, translation evaluation), and seeks to confirm or confound existing views on how the interaction could work. In particular, it offers some parameters for future research: size of translation units in human and machine translation, limitations of a transfer system in the case of a literary text belonging to a sublanguage, translation shifts in human translation, parallel syntactic structures in source and target texts, and formalisation of strategies used in the translation of a (literary) sublanguage. Considering the complexity of such an ambitious endeavour, the thesis does not seek to introduce any significant new directions, but represents a valuable contribution to Translation Studies and Computational Linguistics, by bringing these fields together, through the analysis of strengths and shortfalls of translation theories and of machine translation, in the light of the study of human and machine translation of a literary text. The thesis is therefore an initial mapping of an interdisciplinary focus, so far only seldom and partially tackled.

Research in translation studies and subsequent theories, such as the text-linguistic, the interpretative, or the functionalist approaches shed a new light on the complex cognitive

processes at stake in translation. Even though we are still unable to 'open the mind' of a translator to determine, evaluate and reproduce the operations performed, the study of the actual translation of a text in several languages give new insight into the 'transfer' operated by a translator between two texts, two linguistic systems, two contexts and two cultures. This research was initiated on the premise that translation was a series of operations performed on texts rather than on languages. The outcome of this study gives rise to a more complex picture of the problem at stake. The comparison of the source and target texts definitely places translation on a textual level, the target texts differing in many respects from the source text to fulfil their functions in the respective target cultures, at the time of the translation (as exemplified by the example of the 'sultan'/'dictator' in the Russian translation of *Le Petit Prince*). But these target texts are definitely related to the source texts, like an image of a person in a mirror. Most translation theories use the term 'equivalence', even though 'equivalence in translation' differs greatly from the mathematical concept of equivalence. The source and the target texts are definitely different, the target text carries the imprint ('memory') of the source text, but also the imprint of the translator, as cognitive agent, or interpreter. The attempts of computational linguists to reproduce the 'world knowledge' and insert it into the translation programmes they devise show the importance of the background knowledge necessary in translation. Good translators are said to be translators who know languages, but who also have an encyclopedic knowledge of *what* they are translating, and the best translators are often (and not surprisingly) experts in their domain. Many writers and poets have traditionally been translators, and highly technical and sensitive documents are often translated by engineers and technicians, who receive a training in translation, as an addition to their first qualifications. But knowledge is infinite and versatile, and databases are unlikely to contain all the ever-changing information processed by a human brain. Such as a complete reliance on the syntactic analysis of sentences in a source text failed to produce an appropriate translation of this source text, the belief that knowledge databases will be magic wands is likely to be disappointing, especially as long as the two processes (linguistic analysis on the one hand, background input on the other hand) will be separated. The human mind does not parse sentences, and then chooses the 'right

parse tree' depending on the context. The operations (analysis/interpretation) are intimately intertwined, and the analysis of translation supports the most recent theories in cognitive science about different 'levels of interpretation' that are connected to each other by a compilation process. Even though such theories remain hypothetical, they seem to be validated by several research projects (in cognitive sciences, but also in linguistics), and so far they have not been refuted. The purpose of the thesis is to identify how a more systematic formalisation of meaning representation in natural language might be achieved (using the intermediate levels of interpretation hypothesis), in order to improve machine translation, in particular of culture bound texts such as literary texts. The study presented in this dissertation supports the approach put forward by the applicative and cognitive grammar, and presents the theoretical advantages of the interlingua approach to machine translation, while assessing the practical limits of transfer systems. The thesis also offers translation scholars an insight into how computational linguists work within machine translation, while offering computational linguists an insight into human translation. Conference interpreters, when asked 'what is going in their head when they translate', sometimes confess that 'they see images', or 'a word comes to their mind' (not necessarily in the source or target languages concerned at that time, sometimes a third language seems to prompt them with the 'concept' they infer from the discourse). Such remarks point towards converging fields between computational linguistics (whose development owes much to the analysis of these 'images' or 'concepts') and translation studies (recently enriched by communicative and functionalist approaches). Computational linguists and translator scholars have much more in common than one might believe. One common field between translation studies and computational linguistics is semiotics, which is in particular devoted to the study of signs as meaning generators. This thesis seeks to demonstrate the extent to which machine translation can account for semiotic representation through comparison of machine translation and human translation. The multilingual focus of the study (translation from French into English, German, and Russian) allows a demonstration of representation of meaning and of translation strategies that are not necessarily language-pair specific.

The findings of this thesis are of general value for future research both in literary translation and computational linguistics, by bridging the gap between the two disciplines, and by discussing the validity of some highly valuable theoretical hypotheses (levels of representation, deep semiotic invariant on which languages are anchored, meaning representation) in the light of practice, through the study of human as well as machine translation of a literary text.

8 ANNEXES

8.1 Drawings in *Le Petit Prince*

Direct references to the drawings in the text as intersemiotic references

Number of drawings :

(French)	47 (16 black-and-white ; 31 coloured)
(English)	47 black-and-white
(German)	9 black-and-white
(Russian)	47 (16 black-and-white ; 31 coloured)

(F): (p.9) Voilà la copie du dessin.

(E): (p.5) Here is a copy of the drawing.

(G): *Drawing missing. No translation*

(R): (p.7) Vot kak eto bylo narisovano.

(F): (p.9) Mon dessin numéro 1. Il était comme ça :

(E): (p.5) My Drawing Number One. It looked like this:

(G): (p. 7) Meine Zeichnung Nr. 1. So sah sie aus:

(R): (p.7) Eto byl moj picunok N° 1. Vot cto ja naricoval.

(F): (p.10) Mon dessin numéro 2 était comme ça :

(E): (p.6) My Drawing Number Two looked like this:

(G): (p.8) Hier meine Zeichnung Nr. 2:

(R): (p.8) Vot moj ricunok N° 2.

(F): (p.12) Voilà le meilleur portrait que, plus tard, j'ai réussi à faire de lui.

(E): (p.8) Here you may see the best portrait that, later, I was able to make of him.

(G): (p.10) Hier das beste Portraet, das ich spaeter von ihm zuwege brachte.

(R): (p.10) Vot samij lucšij evo portret, kakoj mne posle udalos' narisovat'.

(F): (p.14) Je dessinai :

(E): (p.10) So then I made a drawing.

(G): (p.12) Ich zeichnete. (*Drawing missing*)

(R): (p.12) Ja narisoval.

(F): (p.14) Je refis donc encore mon dessin :

(E): (p. 10) So then I did my drawing over once more.

(G): (p. 12) Ich machte also meine Zeichnung noch einmal. (*Drawing missing*)

(R): (p.12) Ja opjat' narisoval po-drugomu.

(F): (p.14) Alors, faute de patience, comme j'avais hâte de commencer le démontage de mon moteur, je griffonnai ce dessin-ci.

(E): (p. 10) By this time my patience was exhausted, because I was in a hurry to start taking my engine apart. So I tossed off this drawing.

(G): (p.12) Mir ging die Geduld aus, es war hoechste Zeit, meinen Motor auszubauen, so kritzelte ich diese Zeichnung da zusammen (und knurrte dazu).

(R): (p. 12) Tut Ja poterjal terpenie - ved' mne nado bylo poskoree razobrat' motor - i natsarapal bot cto.

(F): (p.24) et, sur les indications du petit prince, j'ai dessiné cette planète-là. (...) Pourquoi n'y a-t-il pas, dans ce livre, d'autres dessins aussi grandioses que le dessin des baobabs ?

(E): (p.20) So, as the little prince described it to me, I have made a drawing of that planet.(...) "Why there are no other drawings in this book as magnificent and impressive as this drawing of the baobabs ?"

(G): (p.21) Und so habe ich denn diesen Planeten nach den Angaben des kleinen Prinzen gezeichnet.(...) Warum enthaelt dieses Buch nicht noch andere, ebenso grossartige Zeichnungen wie die Zeichnung von den Affenbroetbaeumen ?

(R): (p.22) Malenkij princ podrobno mne vse opisal, i ja narisoval etu planety.(...) nocemu v etoi knige net bolše takix vnušitel'nix ricunok, kak etot, s baobabami ?

(F): (p.95) Ça c'est, pour moi, le plus beau et le plus triste paysage du monde. C'est le même paysage que celui de la page précédente, mais je l'ai dessiné une fois encore, pour bien vous le montrer. C'est ici que le petit prince a apparu sur terre, puis disparu. Regardez attentivement ce paysage, afin d'être sûrs de le reconnaître, si vous voyagez un jour en Afrique, dans le désert.

(E): (p.91) This is, to me, the loveliest and saddest landscape in the world. It is the same as that on page 88, but I have drawn it again to impress it on your memory. It is here that the little prince appeared on Earth, and disappeared. Look at it carefully, so that you will be sure to recognize it in case you travel some day to the African desert.

(G): (p.73) Das ist für mich die schoenste und traurigste Landschaft der Welt. (*sentence suppressed - no drawing on the previous page*). Hier ist der kleine Prinz auf der Erde erschienen und wieder verschwunden. Schaut diese Landschaft genau an, damit ihr sie sicher wiedererkennt, wenn ihr eines Tages durch die afrikanische Wüste reist.

(R): (p.93) Eto, po-moemu, samoe kracivoe i samoe pecal'noe mesto na svete. Etot e ugolok pustyni narisovan i na pregygušcej stranice, no ja narisoval ešče paz, ctoby vy polučše ego pargljadeli. Zdec' Malen'kij princ v pervye pojavilsja na Zemle, a potom iscez. Vsmotrites' vnimatel'nej, ctoby nepremenno uznat' eto mesto, esli kogda-nibud' vy popadeete v Afriku, v pustyn'ju.

8.2 Compensation

Examples of compensation between sentences (merging, addition or suppression of elements).

E = English; G = German; R = Russian.

A- Compensation between sentences, with parts of speech being moved from one to the other sentence

<p>Compensation in English between the two sentences, the little prince being inserted in the first one, and referred to in the second with the anaphora.</p>	<p>Je me trompe un peu aussi sur la taille. Ici le petit prince est trop grand. Là il est trop petit.</p> <p>(E) : I make some errors, too, in the little prince's height : in one place he is too tall and in another too short.</p>
<p>Transposition between the two sentences.</p>	<p>"S'ils voyagent un jour, <i>me disait-il</i>, ça pourra leur servir. Il est quelquefois sans inconvénient de remettre à plus tard son travail.</p> <p>(E) : That would be very useful to them if they were to travel some day. Sometimes", <i>he added</i>, "there is no harm in putting off a piece of work until another day.</p>
<p>Inversion in English, and transformation of the subject (friends), the sentence being split into two independent phrases. Past tense transformed into a present tense, with emphasis on the means used (drawing).</p> <p>In Russian, 'j'ai tant travaillé ce dessin-là' is postponed to the second sentence, hence</p>	<p><i>C'est</i> pour avertir mes amis d'un danger qu'ils frôlaient depuis longtemps, comme moi-même, sans le connaître, que j'ai tant travaillé ce dessin-là. La leçon que je <i>donnais</i> en valait la peine.</p> <p>(E) :My friends, like myself, have been skirting this danger for a long time, without ever knowing it, <i>and so it is for them</i> that I have worked so hard</p>

<p>simplifying the 'c'est pour...que' (subordinate clause).</p>	<p>over this drawing. The lesson which <i>I pass on by this means</i> is worth all the trouble it has cost me.</p> <p>(R) : Ia xochu predupredit' moix druzei ob opasnosti, kotoria davno uzhe ix podsteregaet, a oni dazhe ne podozrevaiut o nei, kak ne podozreval prezhe u ia. Vot pochemu ia tak trudilcia nad etim risunkom, i mne ne zhal' potrachennogo truda.</p>
<p>Compensation between the two sentences, emphasis on the simultaneity between the two events (noon in the States, sunset in France).</p>	<p>En effet. Quand il est midi aux Etats-Unis, le soleil, tout le monde le sait, se couche sur la France. <i>Il suffirait de</i> pouvoir aller en France en une minute pour assister au coucher de soleil.</p> <p>(E) : Just so. Everybody knows than when it is noon in the United States the sun is setting over France. <i>If you could fly</i> to France in one minute, you could go straight into the sunset, <i>right from noon</i>.</p>

B- Splitting of sentences

<p>The French sentence is split into two in Russian, 'jusqu'à' is omitted and implied by 'and six years ago I had to'.</p>	<p>J'ai ainsi vécu seul, sans personne avec qui parler véritablement, jusqu'à une panne dans le désert du Sahara, il y a six ans.</p> <p>(R) : Tak ia zhil v odinochestve, u ne s kem mne bylo pogovorit' po dusham. I vot sheck' let tomu nazad prislos' mne sdelat' vynuzhdenuiu posadku</p>
----------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

	v Saxakhe.
The first sentence is split in English, and the fact that these two drawings are the only ones the author was able to draw is implied by 'the two pictures I had drawn so often'. The slight amendment actually reminds us of the preceding explanations about the author's failure at communicating with the grown-ups.	Comme je n'avais jamais dessiné un mouton, je refis, pour lui, l'un des deux seuls dessins dont j'étais capable. Celui du boa fermé. (E) : But I had never drawn a sheep. So I drew for him one of the two pictures I had drawn so often. It was that of the boa constrictor from the outside.
The splitting into two sentences emphasizes "the turkish astronomer" in English, although the stress in French is more on "only once".	Cet astéroïde n'a été aperçu qu'une fois au télescope, en 1909, par un astronome turc. (E) : This asteroid has only once been seen through the telescope. That was by a Turkish astronomer, in 1909.

C- Merging of sentences

The two sentences in French are joined together and « I had the occasion (to examine them from very near) » is added, hence putting emphasis.	J'ai beaucoup vécu chez les grandes personnes. Je les ai vues de très près. (G) : Ich bin viel mit Erwachsenen umgegangen und habe Gelegenheit gehabt, sie ganz aus der Naehе zu betrachten.
Merging : 'and' is added in the German version to replace the full stop separating the two sentences.	J'ai bien frotté mes yeux. J'ai bien regardé. (G) : Ich habe mir die Augen gerieben und genau

	hingeschaut.
--	--------------

D- Expansion (lexical or syntactic)

<p>« Frighten ? », taken from the previous question is repeated, adding strength to the sentence and underlining the surprise of the grown-ups.</p>	<p>Elles m'ont répondu : « Pourquoi un chapeau ferait-il peur ? »</p> <p>(E) : But they answered : « <i>Frighten ?</i> Why should any one <i>be frightened</i> by a hat ? »</p>
<p>Reinforcement of the statement of the author.</p> <p>What was hinted at in French is clearly expressed in English : grown-ups are unable to understand...</p>	<p>J'ai alors dessiné l'intérieur du serpent boa, afin que les grandes personnes puissent comprendre.</p> <p>(E) : But since the grown-ups were not able to understand it, I made another drawing : I drew the inside of the boa constrictor, so that the grown-ups could see it clearly.</p>
<p>« Je lançai » is expanded in English, and added to the preceding sentence in German. In Russian, it is translated by " I said to the child ". "Ça c'est la caisse" is translated in Russian by, this is a box for you".</p>	<p>Alors, faute de patience, comme j'avais hâte de commencer le démontage de mon moteur, je griffonnai ce dessin-ci. Et je lançai : - Ça c'est la caisse. Le mouton que tu veux est dedans.</p> <p>(E) : By this time my patience was exhausted, because I was in a hurry to start taking my engine apart. So I tossed off this drawing. An I threw out an explanation with it. « This is only his box. The sheep you asked for is inside. »</p> <p>(G) : Mir ging die Geduld aus, es war hoechste Zeit, meine, Motor auszubauen, so kritzelte ich diese Zeichnung da zusammen und knurrte</p>

	<p>dazu : »Das ist die Kiste. Das Schaf, das du willst, steckt da drin. »</p> <p>(R) : Tut ia poterial terpenie - ved' mne nado bylo poskoree raborat' motor- i nacarapal vot chto. ». I skazal malysu : -Vot tebe iashchik. A v njom sidit takoi barashek, kakovo tebe xochetsia.</p>
« he said » is added in Russian, and the action (he bent his head over the drawing) is expanded and translated as « bending the head and examining the drawing ».	<p>Il pencha la tête vers le dessin : - Pas si petit que ça...Tiens ! Il s'est endormi...</p> <p>(R) : Ne takoi uzh on malenkii...- skazal on, nakloniv golovu i razgliadivaia risunok. - Smotri-ka ! On usnul...</p>
Semantic adding in Russian (straight ahead, straight ahead, where the eyes see). The sentence is rendered more poetic in Russian, thanks to the alliteration (pria/glia)	<p>N'importe où. Droit devant lui</p> <p>(R) : -Malo li kuda ? Bse priamo, priamo, kuda gliaza gliadiat.</p>
Adding of "ved'"(since), and "there is very little place" for "it's so small".	<p>Ça ne fait rien, c'est tellement petit chez moi.</p> <p>(R) : Eto ne strashno, ved' u menia tam ochen' malo mesto.</p>
'alors' expanded into verbal phrase emphasizing the discovery, while the source sentence emphasizes rather the demonstration, as well as the concomitance.	<p>Il avait fait alors une grande démonstration de sa découverte à un Congrès International d'Astronomie.</p> <p>(E) : On making his discovery, the astronomer had presented it to the International Astronomical Congress, in a great demonstration.</p>
In English, the emphasis is put on the reason why he was not believed, by the expansion into a	<p>Mais personne ne l'avait cru à cause de son costume.</p>

<p>verbal phrase, and by the place (beginning of the sentence) of this phrase.</p> <p>Explanation and expansion both in English and Russian, by the mention of the Turkish costume, while in French, it is understood by the context (Turkish astronomer).</p>	<p>(E) : But he was in Turkish costume, and so nobody would believe what he said.</p> <p>(R) : No nikto iemu ne poveril, a vce potomu, chto on byl odet' po-tupetskii.</p>
<p>Emphasis in English on the elegance of the astronomer's dress, by the adding of another adjective. Expansion of 'être de son avis' with the introduction of the 'report' which his not mentioned in French.</p> <p>In Russian, the elegance of the astronomer is rendered by 'dressed according to the latest trend'.</p> <p>This interpretation gives some interesting information about the definition of 'elegance' for the Russian translator.</p>	<p>L'astronome refit sa démonstration en 1920, dans un habit très élégant. Et cette fois-ci tout le monde fut de son avis.</p> <p>(E) : So in 1920 the astronomer gave his demonstration all over again, dressed <i>with impressive style and elegance</i>. And this time everybody <i>accepted his report</i>.</p> <p>(R) : V 1920 godu tot astronom snova dolozhil o svoiem otkrytii. Na etot raz on byl odet <i>na poslednei mode</i>, i vce c nim soglasilis'.</p>
<p>Semantic expansion 'alors' rendered by 'only from these figures' (implicit in French, emphasized in English).</p>	<p>Alors seulement elles croient le connaître.</p> <p>(E) : <i>Only from these figures</i> do they think they have learned anything about him.</p>
<p>Systematic adding ('me') in English, Russian, and German, although it is not explicit in French, the sentence can be understood as a habit of the Little Prince, whoever he is addressing to.</p>	<p>Mon ami ne donnait jamais d'explications.</p> <p>(E) : My friend never explained anything <i>to me</i>.</p> <p>(G) : Mein Freund hat <i>mir</i> nie Erklaerung gegeben.</p> <p>(R) : Moi drug nikogda <i>mne</i> nichego ne obiasnial.</p>

<p>Adding in all target languages of 'the walls/plates' of the boxes, which is only implicit in French.</p>	<p>Mais moi, malheureusement, je ne sais pas voir les moutons <i>à travers les caisses</i>.</p> <p>(E) : But I, alas, do not know how to see sheep <i>through the walls of boxes</i>.</p> <p>(G) : Aber ich bin leider nicht imstande, <i>durch die Kistenbretter</i> hindurch Schafe zu sehen.</p> <p>(R) : No ia, k sozhaleniiu, ne umeiu uvidet' barashka <i>skvoz' smenki iashchika</i>.</p>
<p>Semantic expansion in English and Russian by the introduction of 'the little prince's' and the pronouns referring to him.</p>	<p>Chaque jour j'apprenais quelque chose sur la planète, sur le départ, sur le voyage.</p> <p>(E) : As each day passed I would learn, in our talk, something about <i>the little prince's</i> planet, <i>his</i> departure from it, <i>his</i> journey.</p> <p>(R) : Kazhdyj den' ia uznaval chto-nibud' novoe o <i>ego</i> planete, o tom, kak <i>on</i> pokinul i kak <i>stranstvoval</i>.</p>
<p>Adding of 'And what good would it do to tell them that?' in English (in Russian, 'si vous leur dites' is repeated), emphasizing the lack of understanding from grown-ups.</p>	<p>Ainsi, si vous leur dites : "La preuve que le petit prince a existé c'est qu'il était ravissant, qu'il riait, et qu'il voulait un mouton. Quand on veut un mouton, c'est la preuve qu'on existe", elles hausseront les épaules et vous traiteront d'enfant !</p> <p>(E) : Just so, you might say to them : "The proof that the little prince existed is that he was charming, that he laughed, and that he was looking for a sheep. If anybody wants a sheep, that is a proof that he exists." <i>And what good would it do to tell them that ?</i> They would shrug</p>

	their shoulders, and treat you like a child.
Adding of 'albernen' in German ('not serious') emphasizing the uselessness of figures, this message of the author being rendered by constant repetitions and references to figures. In Russian this emphasis is rendered by the translation of 'numéros' into 'numbers and figures' (although French has another word 'nombre', which is different (semantically) from 'numéros').	<p>Mais bien sûr, nous qui comprenons la vie, nous nous moquons bien des numéros.</p> <p>(G) : Wir freilich, die wir wissen, was das Leben eigentlich ist, wir machen uns nur lustig ueber die albernen Zahlen.</p> <p>(R) : No my, kto ponimaet, chto takoie zhizn', my, konechno, smeiomsia nad <i>nomerami i tsiframi</i>.</p>
Semantic expansion with the adding of 'give an air of truth to the story'. Active construction in Russian, putting more emphasis on the role of those who understand life, emphasis reinforced by "spazu uvideli by" (they would see immediately).	<p>Pour ceux qui comprennent la vie, ça aurait eu l'air beaucoup plus vrai.</p> <p>(E) : To those who understand life, that would <i>have given a much greater air of truth to my story</i>.</p> <p>(R) : Te, kto ponimaet, chto takoe zhizn', spazu uvideli by, chto eto vce chistaia pravda.</p>
"lire" is translated by "to take". The translator might have missed the difference in French between the two expressions ("prendre à la légère" being a common phrasing), or he might have consciously used that reference to give more weight to the remark.	<p>Car je n'aime pas qu'on lise mon livre à la légère.</p> <p>(G) : Denn ich moechte nicht, dass man mein Buch leicht nimmt.</p>
'and if I forget him' is implicit in French, thanks to the two previous sentences. The explicit adding of the phrase might indicate the wish to make the message very clear to the young audience, or the interpretation performed by the translator.	<p>C'est triste d'oublier un ami. Tout le monde n'a pas eu un ami. Et je puis devenir comme les grandes personnes qui ne s'intéressent plus qu'aux chiffres.</p> <p>(E) : To forget a friend is sad. Not every one has had a friend. <i>And if I forget him</i>, I may become like the grown-ups who are no longer interested in</p>

	anything but figures.
Adding in German, that emphasizes the length of time, hence the ageing.	J'ai dû vieillir. (G) : Ich musste ja <i>im Laufe der Zeit</i> aelter werden.
Adding in English ('as I learned'), due perhaps to the recurrent 'j'apppris que...' in the French text. This adding might be absolutely unconscious. Semantic expansion in English ('the planet where the little prince lived').	Et en effet, sur la planète du petit prince, il y avait comme sur toutes les planètes, de bonnes herbes et de mauvaises herbes. (E) : Indeed, <i>as I learned</i> , there were on the planet <i>where the little prince lived</i> as on all planets - good plants and bad plants.
Adding of 'plainly', attempt to emphasize the warning.	Je dis : "Enfants! Faites attention aux baobabs!" (E) : "Children," I say <i>plainly</i> , "watch out for the baobabs!"
Nearly lyrical interpretation (through expansion) in English of the rather simple French sentence. In Russian, use of three terms to describe this 'urgency'.	Quand j'ai dessiné les baobabs, <i>j'ai été animé par le sentiment de l'urgence</i> . (E) : When I made the drawings of the baobabs <i>I was carried beyond myself by the inspiring force of urgent necessity</i> . (R) : A kogda ia risoval baobaby, menia vdoxnovlialo soznanie, chto <i>eto strashno vazhno i neotlozhno</i> .
Adding: two words ('sad' and 'monotonous') to translate 'mélancolique'.	Ah! Petit prince, j'ai compris, peu à peu, ainsi, ta petite vie <i>mélancolique</i> . (R) : O Malen'kii prints ! Ponemnogu ia ponial takzhe, kak <i>pechal'na i odonoobrazna</i> byla tvoia zhizn'.
Adding of repetition in English and Russian,	Et tu regardais le crépuscule chaque fois que tu le

<p>hence emphasising the importance of the sunset.</p>	<p>désirais.</p> <p>(E) : You can see <i>the day end and the twilight falling</i> whenever you like...</p> <p>(R) : I ty <i>snova i snova</i> smotrel na zakatnoe nebo, stoilo tol'ko zakhote't'.</p>
<p>Splitting of the sentence, and repetition, emphasis on this special sunset time.</p>	<p>- Attendre quoi ? - Attendre que le soleil se couche.</p> <p>(E) : "Wait ? for what ?" "<i>For the sunset</i>. We must wait <i>until it is time</i>."</p>
<p>Expansion in Russian ("<i>my heart is compressed with pain</i>" for "I have suffered too much grief"), and reintroduction of the little prince ("these memories" become "my little friend").</p>	<p>J'éprouve tant de chagrin à évoquer ces souvenirs.</p> <p>Il y a six ans déjà que mon ami s'en est allé avec son mouton.</p> <p>(R) : Serdtse moe bol'no czhimaetcia, kogda ia vspominaiu moego malen'kovo druga, i nelegko o nim govorit'. Proshlo uzhe shest' let s tekh por, kaka on vmeste so svoim barashkom pokinul menia.</p>

E- Adaptation to children's readership

<p>In Russian, "l'une d'elles" is translated by "such a small planet" (importance of repetitions in a text targeted at children).</p>	<p>Quand un astronome découvre l'une d'elles, il lui donne pour nom un numéro. Il l'appelle par exemple : "l'astéroïde 3251".</p> <p>(R) : Kogda astronom otkryvaet takuiu planetku, on daet ei ne imia, a prosto nomer. Naprimer : asteroid 3251.</p>
---------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<p>The formal 'vous' (you-'thou') is rendered in Russian by 'tu', although Russian has both forms (unlike English). This transformation emphasises the targeted readership - children, whom should be addressed in this way. The French 'vous' used by St Exupéry is closer to 'one'.</p>	<p>Quand <i>vous</i> leur parlez d'un nouvel ami, elles ne vous questionnent jamais sur l'essentiel.</p> <p>(R) : Kogda <i>rasskazyyvaesh'</i> im, to u <i>tebja</i> nojavilis' novyi drug, oni nikogda ne sprosiat o samom glavnom.</p>
<p>Once again, the 'vous' is transformed into 'tu' - targeted children's readership.</p>	<p>Si vous dites aux grandes personnes...</p> <p>(R) : Kogda govorish' vzroslym...</p>
<p>In Russian, the verb 'collect' is rendered by 'catch', another way to aim at a children's readership.</p>	<p>Est-ce qu'il collectionne les papillons ?</p> <p>(R) : Lovit li on babochek ?</p>
<p>The repetition in French ('de bonnes <i>herbes</i> et de mauvaises <i>herbes</i>'), characteristic of children's literature (didactic reasons) is lost in Russian.</p>	<p>Et en effet, sur la planète du petit prince, il y avait comme sur toutes les planètes, de bonnes herbes et de mauvaises herbes.</p> <p>(R) : Na planete Malen'kogo printsa, kak i na liuboi drugoi planete, rastut <i>travy poleznyie i vrednie</i>.</p>
<p>Active form in Russian ('the baoboabs threaten') for a passive in French, giving life to the baobabs (children's literature) and emphasizing the danger.</p>	<p>Mais le danger des baobabs est si peu connu...</p> <p>(R) : No malo kto znaet, chem groziat baobaby...</p>
<p>'se demander' translated in English by 'to ask s.o.'. This can express the wish to include the reader (important feature of children's literature).</p>	<p>Vous <i>vous</i> demanderez peut-être: Pourquoi n'y a-t-il pas, dans ce livre, d'autres dessins aussi grandioses que le dessin des baoboabs ? La réponse est bien simple : j'ai essayé mais je n'ai</p>

	<p>pas pu réussir.</p> <p>(E) : Perhaps you will ask <i>me</i>, "Why there are no other drawings in this book as magnificent and impressive as the drawing of the baobabs ? The reply is simple. I have tried. But <i>with the others</i> I have not been successful.</p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

F- Lexical or syntactic syntactic modifications entailing a change in meaning

<p>Both in English and Russian, the '8 days' are replaced by a week, '8 jours' having precisely this meaning in French. The German version kept the precise '8 Tage'.</p> <p>"It was a question of life or death for me " is translated in Russian by "I had to repair the motor or to die". The Russian expression is more literal, and adds a repetition in the text, making it more accessible to a younger readership.</p>	<p>C'était pour moi une question de vie ou de mort.</p> <p>J'avais à peine de l'eau à boire pour huit jours.</p> <p>(E) : It was a question of life or death for me : I had scarcely enough drinking water to last a week.</p> <p>(R) : Ia dolzhen by ispravit' motor ili pogibnut'.</p> <p>Vody u menia edva khvatilo by na nedelju.</p>
<p>Emphasis in Russian by the elliptic nominal phrase ('<i>What a people these grown-ups !</i>)</p>	<p>Les grandes personnes sont comme ça.</p> <p>(R) : Uzh takoi narod eti vzroslye !</p>
<p>Particularizing in German and Russian (church tower for the church itself), while in English the image is altogether different, the baobabs being compared to castles (different connotative</p>	<p>Je fis remarquer au petit prince que les baobabs ne sont pas des arbustes, mais <i>des arbres grands comme des églises</i> et que, même s'il emportait avec lui tout un troupeau d'éléphants, ce troupeau</p>

<p>meaning).</p>	<p>ne viendrait pas à bout d'un seul baobab.</p> <p>(E) : I pointed out to the little prince that baobabs were not little bushes, but, on the contrary, <i>trees as big as castles</i>; and that even if he took a whole herd of elephants away with him, the herd would not eat up one single baobab.</p> <p>(G) : Ich erklarte dem kleinen Prinzen ausführlich, dass Affebrotbaeume doch keine Stauden sind, sondern <i>kirchturmhohe Baeume</i>, und selbst wenn er eine ganze Herde Elefanten mitnaehme, wuerde diese Herde nicht mit einem einzigen Affenbrotbaum fertig werden.</p> <p>(R) : Ia vozprazil, chto baobaby ne kusty, a <i>ogromnye derev'ia, vyshinoi s kolokol'niu</i>, i ecli dazhe on privedet tseloe ctado slonov, im ne c"ect' i odnogo baobaba.</p>
<p>In German, use of an anaphora (reference to context).</p> <p>In Russian, the little prince becomes the subject and agent.</p>	<p>L'idée du troupeau d'éléphants fit rire le petit prince.</p> <p>(G) : Der Einfall mit den Elefanten brachte <i>ihn</i> zum Lachen.</p> <p>(R) : Uslykhav pro slonov, <i>Malen'kii prints zasmelialc'ia</i>.</p>
<p>Emphasis on the 'costume', both in English and Russian. Moreover, the Russian translation explicitly calls the turkish dictator a 'sultan', this word making reference to the political organisation of Turkey at the beginning of the</p>	<p>Heureusement pour la réputation de l'astéroïde B612 un <i>dictateur</i> turc imposa à son peuple, sous peine de mort, de s'habiller à l'euro péenne.</p> <p>(E) : Fortunately, however, for the reputation of Asteroïde B-612, a Turkish <i>dictator</i> made a law</p>

<p>century (the years 1920s witnessing the fall of the ottoman empire), as well as to a painting describing the refusal of the Cosaks to surrender to the Turks. The expression 'sultan' also gives more exoticism to the narrative.</p>	<p>that his subjects, under pain of death, should change to European costume.</p> <p>(R) : K scast'iu dlia reputatsii asteroida B-612, tureckii <i>sultan</i> velei svoim poddannym pod strakhom smerti nosit' ievropeiskoie plat'e.</p>
<p>As in chapter 4, there is an expansion of 'la planète du petit prince' in the English translation. It seems that the translator refuses to use the possessive case. It might be a reference to the talk between the little prince and the businessman (no one really 'possesses' any planet). In Russian, emphasis is put on 'terrible' by the use of two different words (role of repetition, especially in children's literature).</p>	<p>Or il y avait des graines <u>terribles</u> <i>sur la planète du petit prince</i>...c'étaient les graines de baobab.</p> <p>(E) : Now there were some terrible seeds <i>on the planet that was the home of the little prince</i> ; and these were the seeds of the baobab.</p> <p>(R) : I vot na planete Malen'kogo printsa est' <u>uzhasnye, zlovrednie</u> semena... Eto semena baobabov.</p>
<p>43 is purposefully (44 repeated two sentences below) translated by 44 in English. Adding of 'you said to me'.</p>	<p>Un jour, j'ai vu le soleil se coucher <i>quarante-trois</i> fois !</p> <p>(E) : "One day", <i>you said to me</i>, "I saw the sunset <i>forty-four</i> times!"</p>

G- Deictic and anaphora resolution : pronouns, or expansions giving the exact reference

<p>« Ça » refers to the grass (question from the little prince) and is omitted but explicitated in English and Russian. In English, the grass is explicitly</p>	<p>Ça suffira sûrement. Je t'ai donné un tout petit mouton.</p> <p>(E) : « There will surely be enough grass for</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------

<p>mentioned, as well as the sheep. In Russian, the sheep only (use of a pronoun) is mentioned.</p>	<p>him », I said. It is a very small sheep that I have given you.</p> <p>(R) : Iemu xvatit. Ia tebe daiu sovsem malen'kovo barashka.</p>
<p>Expansion of "en" ("on this subject"). The vagueness in French is transformed into an exact reference.</p>	<p>Je m'efforçai donc d'en savoir plus long.</p> <p>(E) : I made a great effort, therefore, to find out more on this subject.</p>
<p>Anaphora resolution of 'ça' through an inversion and the use of the pronoun 'they' in English.</p>	<p>Les baobabs, avant de grandir, <i>ça</i> commence par être petit.</p> <p>(E) : <i>Before they grow so big, the baobabs start out by being little.</i></p> <p>(R) : Baobaby sperva, <i>poka ne vyrastut</i>, byvaiut sovsem malen'kie.</p>
<p>In Russian, 'ici' and 'là' are expanded and explicited as 'drawings' (reference to the context).</p>	<p>Je me trompe un peu aussi sur la taille. Ici le petit prince est trop grand. Là il est trop petit.</p> <p>(R) : Ia delaiu oshibki i v roste: na odnom ricunke prinz u menia vyshel chereschur bol'shoi, na drugom chereschur malen'kii.</p>
<p>"En" is expanded into "from me".</p>	<p>J'éprouve tant de chagrin à évoquer ces souvenirs.</p> <p>Il y a six ans déjà que mon ami s'en est allé avec son mouton.</p> <p>(E) : I have suffered too much grief in setting down these memories. Six years have already passed since my friend went away <i>from me</i>.</p>

H- Modifications due to the differences between the two linguistics systems

<p>Semantic adding ("d'où viens-tu en volant") necessary in Russian, on the ground of the complex system of verbs of movement. In French, the means used to come (or go) is implied by the context, the verb itself being semantically vague, while in Russian there is a different verb for each kind of movement (on foot, by car, on bicycle, by boat...)</p>	<p>D'où viens-tu mon petit bonhomme ? (R) : Otkuda zhe ty priletel, malysh ?</p>
<p>The noun phrase ("après un silence méditatif") is expressed in Russian by a verbal phrase ("il se tut dans une méditation") and the chronological order of the two events is established by the adverb 'potom' - thereafter.</p>	<p>Il me répondit après un silence méditatif : (R) : On pomolchal v pazdum'e, potom skazal :</p>
<p>In Russian, "I made <i>again</i> a very important discovery" for "I learned a <i>second</i> important fact".</p>	<p>J'avais ainsi appris une seconde chose très importante. (R) : Tak ia sdelal eshche odno vazhnoe otkrytie.</p>
<p>Nominalization in English of 'to succeed'.</p>	<p>Mais je ne suis pas tout à fait sûr de réussir. (E) : But I am not at all sure of success.</p>
<p>In English, German and Russian, the dative 'me' is expressed through a complete verbal phrase (= 'il croyait <i>que je...</i>')</p>	<p>Il <i>me</i> croyait peut-être semblable à lui. (E) : He thought, perhaps, <i>that I was</i> like him. (G) : Er glaubte wahrscheinlich, <i>ich sei</i> wie er. (R) : Mozhet byt', on dumal, <i>chtoija</i> takoi zhe, kak on.</p>

<p>Verbalization in Russian of 'departure' and 'journey'.</p>	<p>Chaque jour j'apprenais quelque chose sur la planète, sur le départ, sur le voyage.</p> <p>(R) : Kazhdyj den' ia uznaval chto-nibud' novoe o ego planete, o tom, kak on pokinul i kak <i>stranstvoval</i>.</p>
<p>Both in English and Russian, the infinitive form 'avant de grandir' is translated by a verbal phrase equivalent to 'avant qu'ils ne grandissent'.</p>	<p>Les baobabs, avant de grandir, <i>ça</i> commence par être petit.</p> <p>(E) : <i>Before they grow so big, the baobabs start out by being little.</i></p> <p>(R) : Baobaby sperva, <i>poka ne vyrastut</i>, byvaiut sovsem malen'kie.</p>
<p>'faire des portraits ressemblants' is translated in Russian by 'give the likeness', the 'portraits' being omitted and implied.</p>	<p>J'essaierai, bien sûr, de faire des portraits le plus ressemblants possible.</p> <p>(R) : Konechno, ia postaraius' peredat' skhodstvo kak mozhno luchshe.</p>
<p>In Russian, the ageing process continues (use of a present tense).</p>	<p>J'ai dû vieillir.</p> <p>(R) : Naverno, <i>ia stareiu</i>.</p>
<p>Both English and German require two different words for 'brindilles', and necessitate a distribution of the noun phrases. Note : the Russian sentence follows the French structure (two subordinate clauses of one noun).</p>	<p>S'il s'agit d'une <i>brindille</i> de radis ou de rosier, on peut la laisser pousser comme elle veut.</p> <p>(E) : If it is only a <i>sprout</i> of radish or the <i>sprig</i> of a rose-bush, one would let it grow wherever it might wish.</p> <p>(G) : Wenn es sich um einen Radieschen - oder Rosentrieb handelt, kann man ihn waschen lassen, wie er will.</p>
<p>Indirect discourse translated by a direct discourse in English, with the subsequent necessary changes</p>	<p>Et un jour <i>il me conseilla de m'appliquer</i> à réussir un beau dessin, pour bien faire entrer ça dans la</p>

(I/you)	<p>tête des enfants <i>de chez moi</i>.</p> <p>(E) : And one day <i>he said to me</i> : "You ought to make a beautiful drawing, so that the children <i>where you live</i> can see exactly how all this is."</p>
Impersonal negative formulation in Russian ('one cannot escape...') for an affirmation in French (as well as in English and German).	<p>Mais, s'il s'agit des baobabs, c'est toujours une catastrophe.</p> <p>(R) : No kogda rech' udet o baobabax, bedy ne minovat'.</p>
In Russian, same inversion (negation for an affirmation as above).	(R) : On ne vypolol vovremia tri kustika.
Noun phrase translated in Russian by a verbal phrase ('how the sun sets').	<p>- J'aime bien les couchers de soleil. Allons voir <i>un coucher de soleil</i>.</p> <p>(R) : Ja ochen' liubliu zakat. Poidem posmotrim, <i>kak zoxodit solntse</i>.</p>
Verbal phrase in English for a noun phrase in French, addition of 'quiet pleasure' to translate 'douceur'. Splitting into two sentences in Russian, and introduction of a verb (to admire'), similar to the 'look' added in English.	<p>Tu n'avais eu longtemps pour distraction que la douceur des couchers de soleil.</p> <p>(E) : For a long time you have found your only entertainment in the <i>quiet pleasure of looking at the sunset</i>.</p> <p>(R) : Dolgoe vremia u tebia bylo lish' odno razvlechenie : ty ljubovalcia zakatom.</p>
Nominalization of "à raconter" in German.	<p>J'éprouve tant de chagrin à évoquer ces souvenirs.</p> <p>Il y a six ans déjà que mon ami s'en est allé avec son mouton.</p> <p>(G) : Ich empfinde so viel Kummer beim Erzaehlen dieser Erinnerungen. (...)</p>

I- Similarities between the English and the Russian translations

<p>Both the English and Russian versions render "j'avais pu être intrigué" by "ma curiosité a été éveillée". Such a similarity might be explained by the fact that the Russian translator used the English translation. Otherwise, one might wonder about such a convergence of semantic interpretation.</p>	<p>Vous imaginez combien j'avais pu être intrigué par cette demi-confiance sur 'les autres planètes'.</p> <p>(E) : You can imagine how my curiosity was aroused by this half-confidence about the "other planets"</p> <p>(R) : Mozhetе sebe predstavit', kak razgorelos' moe liubopytstvo ot etovo polupriznaniia o "drugix planetax".</p>
<p>Both English and Russian versions present a totally similar semantic interpretation for "il lui donne pour nom un numéro", and the syntax of both sentences is identical.</p>	<p>Quand un astronome découvre l'une d'elles, il lui donne pour nom un numéro. Il l'appelle par exemple : "l'astéroïde 3251".</p> <p>(E) : When an astronomer discovers one of these he does not give it a name, but only a number.(...)</p> <p>(R) : Kogda astronom otkryvaet takuiu planetku, on daet ei ne imia, a prosto nomer. Naprimer : asteroid 3251.</p>
<p>Use of a verbal phrase both in English and Russian for 'on the indications'. Noun phrase in German.</p>	<p>Et, sur les indications du petit prince, j'ai dessiné cette planète-là.</p> <p>(E) : So, as the little prince described it to me, I have made a drawing of that planet.</p> <p>(R) : Malen'kii prints podrobno mne vce opisal, i ia narisoval etu planetu.</p>

8.3 Machine translation error analysis –English

Type of error	Automatically translated target text (chapter)
Words untranslated (absent from the dictionary), they appear as such (but highlighted) in the translated text.	stylographe (2) - décoiffée (8) - sottte/enrhumée (9) - démodent (15) - pensum (17) - assoiras (21) - loco (22) - assoit /rayon (24) - margelle (25) -
Tense analysis : identical forms in French for different tenses (present, and past historic for example), unrecognized tenses (past historic)	when I finally <i>succeed</i> in speaking, I <i>tell</i> to him (2) - He/it <i>blushes</i> , took then (7) - he/it <i>cannot</i> say anything besides (7) - On all that ? tells the small prince. - On all that... <i>replies</i> the king (10)
Unrecognized locutions	to my tower (1) - the big people (1 and al.) - of the first squint (1) - if he/it pleases you (2) - to the rising of the day (2) - an all small sheep (2)
Tense agreement or analysis	I drew the inside of the boa then, so that the big people <i>can</i> understand.(1)
Ungrammatical constructions - preposition use	it is tiresome <i>of</i> always and always <i>to</i> give them explanation (1) - it is necessary to tell <i>to</i> them (4) - he/it saw me, my hammer <i>to</i> the hand, and the <i>black fingers of grease</i> (...) (7)
Unsolved anaphora analysis	he/it was like that(1) - I wanted to know if she/it was indeed understanding(1) -But always it/she answered me (1) - She/it said (2) - and he/it repeated me - he/it answered me(2) - make another <i>of it</i> (2) -but <i>he/it</i> didn't answer me (3) - retiring my sheep of his/her/its pocket, he/it immersed <i>himself/itself</i>

	<p>in the contemplation of his/her/its treasure (3)- I endeavored therefore of <i>in</i> to know longer (3) - but where want you that <i>he/it</i> goes? - <i>he/it</i> calls him for example "the asteroid" 3251" (4) - <i>he/it</i> had made a big demonstration of his/her/its discovery then (4) - there <i>he/it</i> is too small (4) - I didn't understand why <i>he/it</i> was so important that sheep ate bushes -</p>
<p>Untranslated possessives due to unsolved anaphora resolution</p>	<p>I got to his/her/its range(1) - I immediately glimpsed a gleam, in the mystery of <i>his/her/its</i> presence (3) - retiring my sheep of <i>his/her/its</i> pocket, <i>he/it</i> immersed himself/itself in the contemplation of <i>his/her/its</i> treasure (3) - it is that <i>his/her/its</i> planet of origin was hardly bigger than a house (4) - <i>he/it</i> had made a big demonstration of <i>his/her/its</i> discovery then (4) - but no one had believed him because of <i>his/her/its</i> costume (4) - what is the sound of <i>his/her/its</i> voice ? (4) - how much does his/her/itsfather win? (4) - I also hesitate on the color of his/her/its costume (4)</p>
<p>Erroneous anaphora resolution</p>	<p>he gives him for name a number (4) - <i>he/it</i> calls <i>him</i> for example "the asteroid" 3251" (4) - then only they believe to know it (4) - if I try to describe <i>it</i> here, it is in order ti not to forget him (4) - the small prince never gave up a question, once he had put <i>her</i> (7) - and that makes <i>it</i> inflate pride (7) - I took <i>it</i> in arms, I rocked <i>it</i> (7)</p>

Subject and object pronoun	But, of course, <i>us</i> that understand life, we mock good of numbers ! (4) - once upon a time there was a small prince who hardly lived in a planet bigger than <i>he</i> (4) - but <i>me</i> , unfortunately, I don't know how to see sheep through cases (4)
Relative pronoun	without anyone with <i>that</i> to speak truly(2) - once upon a time there was a small prince who hardly lived in a planet bigger than he, and <i>that</i> had need of a friend (4) - for those <i>that</i> understand life that would have seemed a lot trueer (4) - I made the small prince <i>whom</i> baobabs are not bushes notice (5) - and this is not serious to try to understand why they give themselves so much pain to manufacture itself/themselves of thorns <i>who</i> serve never to anything ? (7)
Miscellaneous grammar mistakes	to know an as reasonable man (1) - is six years old(2) - I didn't have with me nor mechanic, nor <i>temporary</i> (2) - the best portrait that I succeeded in making him (2) - with eyes all circles of astonishment (2) - I remembered <i>whereas</i> I had especially studied geography...(2) - I want a sheep that <i>quick</i> a long time (2) - and I trust to teach to him that I flew (3) -un très joli <i>éclat de rire</i> (3) - what planet are you? (3) - my friend had a new burst to laugh (3) - When you speak them of a new friend (4) - if I try to describe it here, it is in order <i>to</i> not <i>to</i> forget

	him (4) - that made me a few shame (7)
Reflexive verbs	something had broken <i>himself</i> in my motor (2) - The first evening <i>me</i> I am lulled therefore(2) - I was surprised to see to illuminate the face of my young judge (2)
Idiomatic phrasing of 'bien'	I was <i>well</i> more isolated (2) - I rubbed my eyes <i>well</i> (2) - I watched <i>well</i> (2) - you see <i>well</i> (2) - you cannot come from far good (3) - what is <i>well</i> , with the case that you gave me (3) - nous nous moquons <i>bien</i> des numéros (4) - is it <i>well</i> true, this is not, that sheep eat bushes (5) - Unfortunately France is <i>well</i> too distant (6) - I mocked <i>good</i> of my hammer, my bolt, the thirst and the death (7)
Idiomatic use of 'faire'	(...if my drawing) made them fear (1) What do you make here? (2) - that doesn't make anything (2, 4) - make another of it (2)
'faire faire'	The idea of the elephant herd had the small prince laughed (5) - and that makes it inflate pride (7)
'devoir'	I should have chosen another profession therefore(1) - I should have aged (4)
'verb+verb' construction	I was surprised <i>to see to illuminate</i> the face of my young judge (2) -
Miscellaneous idiomatic phrasing	he/it leaned <i>the head</i> toward the drawing (2) - he/it nodded the head mildly(3) - to dress to the European (4) - what age does he/it have ?(4) - they will make higher shoulders and wil call you

	child! (4) - because I don't like that one reads my book <i>to the light</i> (4) - then I grope <i>as here and like that</i> , somehow (4) - at random of reflections (5) -
Impersonal passive constructions	him I was necessary a long time to understand...(3)
'Qu'est-ce que...?'	that is that that thing?(3)
Prepositions	I took of my pocket a sheet of paper (2) - retiring my sheep of his/her/its pocket (3) - that will serve him of house (3) - to the telescope (4) - This time everybody was his/her/its opinion (4) - with geraniums to windows (4) - ...who lived in a planet (4) - he/it believed me maybe similar in him (4) -
Idiomatic use of 'droit'	right before him (3) - right before oneself (3)
'ne...que'	This asteroid has not been seen that once (4)
'autre...que'	...when one never made any other tentatives that the one of a boa closed and the one of an open boa (4)
'comme' wrongly translated as a comparative	As it is pretty ! (4)
Word order changing the meaning of the sentence	Once upon a time there was a small prince who <i>hardly</i> lived in a planet bigger than he (4)
Comparatives	That would have seemed a lot <i>trueer</i> (4) - and I can become <i>as</i> the big people who are only interested in numbers (4) - but of the big trees as churches (5) - this more are not serious and more important that additions of a thick Mr red ? (7)
Superlatives	I will try, of course, to make some possible the

	most alike portraits (4)
'il y a ' - expression of time	My friend left with his/her/its sheep <i>for</i> six years (4) - has the millions of years that flowers manufacture thorns (7) - the millions of years ago that sheep eat nevertheless flowers (7)
Conditional	So same he/it carried away with him all one herd of elephants (5)
'ça'	Baobabs, before growing, that starts with being small (5)
'un peu'	And <i>a few</i> later (6) - that made me <i>a few</i> shame (7)
Erroneous analysis of 'de' taken as a possessive, hence inserted before the noun phrase, or translated as 'of'	He/it was sufficient you <i>to pull your some step chair</i> (6) - he/it shook to the wind <i>of</i> hairs quite gilt (7) - the small prince was now all pale <i>of</i> anger (7) -
Misplaced adverb, changing the meaning of the target sentence	The fifth day, thanks to the sheep this secret of the life of the small prince had been revealed <i>always</i> (7)
'à quoi'	Then thorns, to what do they serve ? (7)
Unrecognized past tense due to the adverb separating the verb and the auxiliary in the source sentence.	He/it never <i>has</i> anything <i>makes</i> another one that of additions (7)
Wrong analysis of 'de/des' indefinite article	and this is not serious to try to understand why they give themselves so much pain to manufacture itself/themselves <i>of</i> thorns who serve never to anything ? (7)

8.4 Machine translation error analysis -German

Type of error	Automatically translated target text (chapter)
Words untranslated (absent from the dictionary), they appear as such (but highlighted) in the translated text.	Geschichten <i>Vécues</i> (1) - <i>Parce</i> das (1,2,...) - <i>décoiffée</i> (8) - <i>enrhumée</i> (9) - <i>businessman</i> (13, 14,...) - <i>démodent</i> (15) - <i>pensum</i> (17) - <i>assoiras</i> (21) - <i>margelle</i> (25)
Words wrongly translated	Die einen <i>Falben</i> schluckte (1)
Wrong case due to a wrong source sentence analysis	Dies ist eine Disziplin-Frage, sagte mehr mir spaet <i>den kleinen Prinzen</i> (5)
Tense analysis : identical forms in French for different tenses (present, and past historic for example), unrecognized tenses (past historic/imperative), wrong subjunctive in German	<i>Zeichnet</i> mir ein Schaf! (2) - Mein Freund lachelt nett, mit Nachsicht (2) - ich wuerde gern anfangen (4) - ich wuerde gern sagen (4)
Unrecognized locutions	Von Seite zu lassen (1) - vom ersten kurze Blick (1) - von sehr nahe (1) - Tote unter Androhung (3) - nach der Art von den Erzaehlung von Feen (4) - haette das die Luft wirklicher gehabt (4) - ...dass man mein Buch an der Leichten liest (4) - so gut, dass Uebel (4)
Wrong case agreement	Ich habe dann <i>den Inneren</i> der Boa gezeichnet (1)
Preposition use	Ich sprach mit ihm <i>von</i> Bruecke, Golf, Politik und Bindern (1) - und werden Sie von Kind (4) - Jeden Tad lernte ich etwas <i>auf</i> dem Planeten, <i>auf</i> der Abfahrt, <i>auf</i> der Reise (5)
Unsolved anaphora analysis	Ohne es/sie/ihn zu kauen (1) - Dies is ganz und

	gar so das ich ihr/ihm wollte! (2)
Untranslated possessives due to unsolved anaphora resolution	Ich stellte mich in seine/ihre Reichweite (1) - im Geheimnis seiner/ihrer Anwesenheit (3) - er vertiefte sich in der Betrachtung seines/ihrer Schatzes (3)
Erroneous anaphora resolution	Er (= die Zeichnung) war so (1) - Er stellte eine Boa-Schlange dar - Ich wollte wissen, ob <i>sie</i> (er in the previous sentence) wirklich verstaendig waere (1) - Er (= das Maennchen) hatte die erscheinung eines verlouerten Kindes in nichts in der Mitte der Wueste...(2)
Relative pronoun	In einem Buch auf dem Urwald, <i>der</i> "Geschichten Vécues" hiess (1) - Das ist das beste Portraet, <i>der</i> spaeter ist es mir gelungen, von ihm zu machen (2) -
Subordinate clauses	Vergesst nicht <i>nur</i> ich befand mich an tausend Tausenden aller bewohnten Gegend (2)
Miscellaneous grammar mistakes	Dies ist ermuedend, fuer die Kinder, <i>vor</i> ihnen immer und immer von den Erklaerungen zu geben (1) - und so machte ich die Kenntis <i>aus dem kleinen Prinzen</i> (2)
Erroneous analysis of the indefinite article 'de/des'	Von ihnen <i>von</i> den Erklaerungen zu geben (1) Dann ausscheidend mein Schaf seiner/ihrer Tasche...(3) - dass ich einen Malkasten gekauft habe und <i>von</i> den Stiften (4)
Idiomatic phrasing of 'bien'	Ich habe <i>gut</i> meine augen abgerieben (2)
'de' in negative phrase ('ne...pas...de')	Ich will nicht <i>von</i> einem Elefant in einer Boa

	(2)
Miscellaneous idiomatic phrasing	Haengst !(2) - Ja, ich <i>machte</i> bescheiden (3) - Ist dies sehr wahr, ist dies nicht, dass...?(5)
Impersonal constructions	Glaubst du das er viel Kraut an diesem Schaf muss ? (2) - Es brauchte mich lange, um zu verstehen, woher er kam (3) - <i>er</i> war einmal (4) - Ich verstand nicht, warum <i>er/es so wichtig war</i> , wie die Schafe die Straeucher assen (5) - Sie schlafen im Geheimnis der Erde, <i>bis er Phantasie am einen von ihnen nimmt</i> , aufzuwachen (5)
'Qu'est-ce que...?'	Das isst, dass diese Sache ? (3)
'parce que'	<i>Parce das</i> bei mir dies ganz klein is (2) - <i>parce, den</i> die Trinker dopplet sehen (3) - <i>parce, den</i> die Blumen Eintagsfliege sind (3) -
Prepositions	In einem Buch <i>auf dem</i> Urwald (1) - ich habe dann <i>auf</i> den Abenteuern des Dschungels nachgedacht (1)
'Faire faire' (translation in German depends on the verb)	Ich <i>liess</i> am kleinen Prinzen <i>bemerkten...</i> (5)
'ne...jamais'	Der kleine Prinz, der mir viele Fragen stellte, schien <i>nie</i> meine zu hoeren <i>nicht</i> (3)-
'ne... plus ... que'	Die nicht mehr als an den zahlen interessieren (4)
'ça'	Die Affenbrotbaeume, bevor zuzunehmen, beginnt <i>das</i> damit, klein zu sein (5)
Word order changing the meaning of the sentence	Ich war in der Mitte des Ozeanes viel isolierter als ein Floss (2)
Comparatives	Aber meine Zeichnung is natuerlich viel weniger

	begeistert, <i>dass</i> das Modell (2) - Dies ist eine Disziplin-Frage, sagte <i>mehr</i> mir <i>spat</i> den kleinen Prinzen (5)
'quand/lorsque'	<i>Wenn</i> ich sechs Jahre alt war... (1)
'pendant'	Waehrend der sechs Monate ihrer Verdauung (1)

8.5 Human and machine identical translations (in terms of syntax) - English

Source sentences	Human/ <i>Machine</i> translations
Voilà la copie du dessin.	Here is a copy of the drawing. <i>There is a copy of the drawing.</i>
J'avais été découragé par l'insuccès de mon dessin numéro 1 et de mon dessin numéro 2.	I had been disheartened by the failure of my Drawing Number One and of my Drawing Number Two. I had been discouraged by the failure of my drawing number one and my drawing number two.
C'était pour moi une question de vie ou de mort.	It was a question of life or death for me. <i>It was for me a matter of life and death.</i>
Dessine-moi un mouton !	Draw me a sheep ! <i>Draw me a sheep !</i>
Et c'est ainsi que je fis la connaissance du petit prince.	And that is how I made the acquaintance of the little prince. <i>And this is how I made the acquaintance of the small prince.</i>
Ce n'est pas une chose. Ça vole. C'est un avion. C'est mon avion.	That is not an object. It flies. It is an aeroplane. It is my aeroplane. <i>This is not a thing. That flies. It is a plane. It is my plane.</i>
Puis il ajouta :	Then he added : <i>Then he/it added :</i>

<p>J'entrevis aussitôt une lueur dans le mystère de sa présence, et j'interrogeais brusquement :</p> <p>- Tu viens donc d'une autre planète ?</p>	<p>[At that moment] I caught a gleam of light in the [impenetrable] mystery of his presence ; and I demanded, abruptly :</p> <p>- Do you come from another planet ?</p> <p><i>I [immediately] glimpsed a gleam, in the mystery of his/her/its presence, and I interrogated suddenly :</i></p> <p><i>-Do you [therefore] come from another planet ?</i></p>
<p>Puis, sortant mon mouton de sa poche, il se plongea dans la contemplation de son trésor.</p>	<p>Then, taking my sheep out of his pocket, he buried himself in the contemplation of his treasure.</p> <p><i>Then, retiring my sheep of his/her/its pocket, he/it immersed himself/itself in the contemplation of his/her/its treasure.</i></p>
<p>Il me répondit après un silence méditatif :</p>	<p>After a reflective silence he answered :</p> <p><i>He/it answered me after a meditative silence :</i></p>
<p>J'ai de sérieuses raisons de croire que la planète d'où venait le petit prince est l'astéroïde B612.</p>	<p>I have serious reason to believe that the planet from which the little prince came is the asteroid known as B612.</p> <p><i>I have serious reasons to believe that the planet from wher came the small prince is the asteroid B612.</i></p>
<p>Est-ce qu'il collectionne les papillons ?</p>	<p>Does he collect butterflies ?</p> <p><i>Does he/it collect butterflies?</i></p>
<p>Combien pèse-t-il ?</p>	<p>How much does he weigh ?</p> <p><i>How much does he/it weigh?</i></p>
<p>La preuve que le petit prince a existé c'est qu'il</p>	<p>The proof that the little prince existed is that he</p>

<p>était ravissant, qu'il riait, et qu'il voulait un mouton.</p>	<p>was charming, that he laughed, and that he was looking for a sheep.</p> <p><i>The proof that the small prince existed is that he was charming, that he laughed, and that he wanted a sheep.</i></p>
<p>Elles sont comme ça.</p>	<p>They are like that.</p> <p><i>They are like that.</i></p>
<p>- Oui, c'est bien vrai. – Ah ! Je suis content.</p>	<p>'Yes, that is true.' 'Ah ! I am glad.'</p> <p><i>- Yes, it is true. – Ah ! I am happy.</i></p>
<p>Mais le petit prince ajouta</p>	<p>But the little prince added :</p> <p><i>But the small prince added:</i></p>
<p>Mais les graines sont invisibles.</p>	<p>But seeds are invisible.</p> <p><i>But seeds are invisible.</i></p>
<p>Et si la planète est trop petite, et si les baobabs sont trop nombreux, ils la font éclater.</p>	<p>And if the planet is too small, and [the] baobabs are too many, they split it in pieces...</p> <p><i>And if the planet is too small, and [if] baobabs are too numerous, they make it explode.</i></p>
<p>« C'est une question de discipline, » me disait plus tard le petit prince.</p>	<p>'It is a question of discipline,' the little prince said to me later on.</p> <p><i>« It is a question of discipline, » told me later the small prince.</i></p>
<p>C'est un travail très ennuyeux, mais très facile.</p>	<p>It is a very tedious work, [the little prince added], but very easy.</p> <p><i>It is a very boring work, but very easy.</i></p>
<p>Un mouton, s'il mange les arbustes, il mange aussi les fleurs ?</p>	<p>A sheep – if it eats little bushes, does it eat flowers, too ?</p> <p>A sheep – if he/it eats bushes, does he/it also eat</p>

	flowers ?
Même les fleurs qui ont des épines ?	Even flowers that have horns ? <i>Even flowers that have horns?</i>
Oui. Même les fleurs qui ont des épines.	Yes, even flowers that have horns. <i>Yes. Even flowers that have horns.</i>
Je ne te crois pas ! Les fleurs sont faibles. Elles sont naïves. Elles se rassurent comme elles peuvent.	I don't believe you ! Flowers are weak creatures. They are naive. They reassure themselves as [best] they can. <i>I don't believe you! Flowers are weak. They are naive. They reassure themselves as they can.</i>
Le petit prince déranga de nouveau mes réflexions :	Again the little prince disturbed my thoughts : <i>The small prince disturbed my reflections again :</i>
Il me regarda stupéfait.	He stared at me, thunderstruck. <i>He/it looked at me amazed.</i>
Il était vraiment très irrité.	He was really very angry. <i>He/it was indeed very angry.</i>
Ma fleur est là quelque part...	Somewhere my flower is there... <i>My flower is there somewhere...</i>
La nuit était tombée.	The night had fallen. <i>The night had fallen.</i>
La fleur que tu aimes n'est pas en danger...	The flower that you love is not in danger... <i>The flower that you like is not in danger...</i>
Elle choisissait avec soin ses couleurs. Elle s'habillait lentement, elle ajustait un à un ses pétales.	She chose her colours with [the greatest] care. She dressed [herself] slowly. She adjusted her petals one by one. <i>She/it chose with care his/her/its colors. She/it dressed slowly, she/it adjusted his/her/its petals</i>

	<i>one by one.</i>
Et je suis née en même temps que le soleil...	And I was born at the same time as the sun... <i>And I was born at the same time as the sun...</i>
J'allais le chercher mais vous me parliez !	I was [just] going to look for it [when] you spoke to me... <i>I was going to look for it [but] you spoke to me!</i>
Les fleurs sont si contradictoires !	Flowers are so inconsistent ! <i>Flowers are so contradictory !</i>
Mais j'étais trop jeune pour savoir l'aimer.	But I was too young to know how to love [her]... <i>But I was too young to know how to like [him].</i>
Mais, comme il disait : « On ne sait jamais ! »	But, as he said, 'One never knows !' <i>But, as he/it said : 'One never knows!'</i>
S'ils sont bien ramonés, les volcans brûlent doucement et régulièrement, sans éruptions.	If they are well cleaned out, volcanoes burn slowly and steadily, without [any] eruptions. <i>If they are swept well, volcanoes burn mildly and regularly, without eruptions.</i>
Les éruptions volcaniques sont comme des feux de cheminée.	Volcanic eruptions are [like] fires in a chimney. <i>[The] volcanic eruptions are [as] the fires of chimney.</i>
Evidemment sur notre terre nous sommes beaucoup trop petits pour ramoner nos volcans.	On our earth we are obviously much too small to clean out our volcanoes. <i>On our earth we are evidently a lot too small to sweep our volcanoes.</i>
Le petit prince arracha aussi, avec un peu de mélancolie, les dernières pousses de baobab.	The little prince also pulled up, with a certain sense of dejection, the last shoots of baobabs. <i>The small princed also pulled, with a little melancholy, the last shoots of baobabs.</i>

Mais tous ces travaux familiers lui parurent, ce matin-là, extrêmement doux.	But on this last morning all these familiar tasks seemed very precious to him. <i>But all these familiar works appeared to him, that morning, extremely soft.</i>
La fleur toussa.	The flower coughed. <i>The flower coughed.</i>
Il fut surpris par l'absence de reproches.	He was surprised by this absence of reproaches. <i>He/It was surprised by the absence of reproaches.</i>
Il ne comprenait pas cette douceur calme.	He did not understand this quiet sweetness. <i>He/it didn't understand this quiet sweetness.</i>
Je suis une fleur.	I am a flower. <i>I am a flower.</i>
J'ai mes griffes.	I have my claws. <i>I have my claws.</i>
Et elle montrait naïvement ses quatre épines.	And, naïvely, she showed her four claws. <i>And she/it showed his/her/its four thorns ingeniously.</i>
Puis elle ajouta	Then she added: <i>Then she added:</i>
Le premier était habité par un roi.	The first [of them] was inhabited by a king. <i>The first was lived by a king.</i>
Ah ! Voilà un sujet, s'écria le roi quand il aperçut le petit prince.	Ah ! Here is a subject, exclaimed the king, when he saw the little prince coming. <i>Ah ! There is a topic, exclaimed the king, when he/it saw the small prince.</i>
Il resta donc debout et, comme il était fatigué, il bâilla.	[So] he remained standing [upright], and, since he was tired, he yawned.

	<i>He/it remained standing [therefore], and, as he/it was tired , he/it yawned.</i>
Il est contraire à l'étiquette de bâiller en présence d'un roi.	It is contrary to etiquette to yawn in the presence of a king. <i>He/it contrary to the label to yawn in presence of a king.</i>
(...) je t'ordonne de bâiller.	(...) I order you to yawn. <i>(...) I order you to yawn.</i>
Bâille encore. C'est un ordre.	Yawn again! It is an order. <i>Yawn again; It is an order.</i>
Ça m'intimide...Je ne peux plus...	That frightens me...I cannot, any more... <i>That intimidates me, I am not able to anymore.</i>
Il bredouillait un peu et paraissait vexé.	He spluttered a little, and seemed vexed. <i>He/it mumbled a little and appeared vexed.</i>
C'était un monarque absolu.	[He] was an absolute monarch. <i>[It] was an absolute monarch.</i>
Ce serait ma faute.	It would be my fault. <i>It would be my mistake.</i>
Puis-je m'asseoir ?	May I sit down ? <i>Can I sit down?</i>
Mais le petit prince s'étonnait. La planète était minuscule.	But the little prince was wondering. The planet was tiny. <i>But the small prince wondered. The planet was minuscule.</i>
Sire...Sur quoi réglez-vous ?	Sire – over what do you rule ? <i>Lord...On what do you reign?</i>
Elles obéissent aussitôt. Je ne tolère pas	They obey instantly. I do not permit

l'indiscipline.	insubordination. <i>They immediately obey. I don't tolerate the insubordination.</i>
Je voudrais voir un coucher de soleil...Faites-moi plaisir.	I should like to see a sunset...Do me that kindness... <i>I would like to see a sunset...Make me pleasure...</i>
L'autorité repose d'abord sur la raison.	[Accepted] authority rests first [of all] on reason. <i>The authority first rests on [the] reason.</i>
J'ai le droit d'exiger l'obéissance parce que mes ordres sont raisonnables.	I have the right to require obedience because my orders are reasonable. <i>I have the right to require the obedience because my orders are reasonable.</i>
Le petit prince bâilla. Il regrettait son coucher de soleil manqué.	The little prince yawned. He was regretting his lost sunset. <i>The small prince yawned. He/it regretted his/her/its unsuccessful sunset.</i>
Tu le condamneras à mort de temps en temps. Ainsi sa vie dépendra de ta justice.	From time to time you will condemn him to death. Thus his life will depend on your justice. <i>You will condemn him from time to time to death. So his/her/its life will depend on your justice.</i>
Si votre majesté désirait être obéie ponctuellement, elle pourrait me donner un ordre raisonnable.	If your Majesty wishes to be promptly obeyed, he said, he should be able to give me a reasonable order. <i>If your Majesty wanted to be obeyed punctually, it could give me a reasonable order.</i>
Il me semble que les conditions sont favorables...	It seems to me that conditions are favourable... He/it seems to me that conditions are favorable...

Je te fais mon ambassadeur.	I make you my ambassador. <i>I make you my ambassador.</i>
La seconde planète était habitée par un vaniteux.	The second planet was inhabited by a conceited man. <i>The second planet was lived by a conceited man.</i>
Car, pour les vaniteux, les autres hommes sont des admirateurs.	For, to conceited men, all other men are admirers. <i>Because, for conceited persons the other men are admirers.</i>
Après cinq minutes d'exercice le petit prince se fatigua de la monotonie du jeu.	After five minutes of this exercise the little prince grew tired of the game's monotony. <i>After five minutes of exercise the small prince got tired of the monotony of the game.</i>
Mais le vaniteux ne l'entendit pas.	But the conceited man did not hear him. <i>But the conceited person didn't hear him.</i>
Est-ce que tu m'admires vraiment beaucoup ?	Do you really admire me very much ? <i>Do you admire me really a lot?</i>
Les grandes personnes sont décidément bien bizarres.	The grown-ups are certainly very odd. <i>The big people are decidedly very odd.</i>
La planète suivante était habitée par un buveur.	The next planet was inhabited by a tippler. <i>The following planet was lived by a drinker.</i>
Je bois, répondit le buveur, d'un air lugubre.	I am drinking, replied the tippler, with a lugubrious air. <i>I drink, answered the drinker, with a lugubrious air.</i>
Trois et deux font cinq.	Three and two make five. <i>Three and two make five.</i>
Millions de ces petites choses que l'on voit	"Millions of those little objects," he said, "which

quelquefois dans le ciel.	one sometimes sees in the sky". <i>Millions of these small things that one sometimes sees in the sky.</i>
Rien. Je les possède.	Nothing. I own them. <i>Nothing. I possess them.</i>
Les rois ne possèdent pas. Ils "règnent sur".	Kings do not own. They reign over. <i>Kings don't possess. They "reign on".</i>
Je les gère. Je les compte et je les recompte, dit le businessman. C'est difficile.	I administer them, replied the businessman. I count them and recount them. It is difficult. <i>I manage them. I count them and I recount them, said the businessman. It is difficult.</i>
Le petit prince n'était pas satisfait encore.	The small prince was still not satisfied. <i>The small prince was not yet satisfied.</i>
Ça veut dire que j'écris sur un petit papier le nombre de mes étoiles.	That means that I write the number of my stars on a little paper. <i>That means that I write on a small paper the number of my stars.</i>
Et c'est tout ?	And that is all? <i>And that is all?</i>
C'est amusant, pensa le petit prince. C'est assez poétique.	"It is entertaining", thought the little prince. It is rather poetic. <i>It is funny, thought the small prince. It is poetic enough.</i>
Car je ramonne aussi celui qui est éteint. On ne sait jamais.	For I also clean the one that is extinct; one never knows. <i>Because I also sweep the one that is extinguished. One never knows.</i>

La cinquième planète était très curieuse.	The fifth planet was very strange. <i>The fifth planet was very curious.</i>
C'était la plus petite de toutes.	It was the smallest of all. <i>It was smallest of all.</i>
C'est une occupation très jolie. C'est véritablement utile puisque c'est joli.	That is a beautiful occupation. And since it is beautiful, it is truly useful. <i>It is a very pretty occupation. It is truly useful since it is pretty.</i>
Ce n'est pas drôle du tout, dit l'allumeur.	"It is not funny at all!", said the lamplighter. <i>This is not at all funny, says the lighter.</i>
Voilà un explorateur! S'écria-t-il, quand il aperçut le petit prince.	"Here is an explorer!" he exclaimed to himself when he saw the little prince coming. <i>There is an explorer! He/it exclaimed, when he/it saw the little prince.</i>
Le petit prince s'assit sur la table et souffla un peu. Il avait déjà tant voyagé!	The little prince sat down on the table and panted a little. He had already travelled so much and so far! <i>The little prince sat down on the table and blew a little. He had already traveled so much!</i>
Ça c'est bien intéressant, dit le petit prince.	"That is very interesting", said the little prince. <i>"That is well interesting", says the little prince.</i>
Et il jeta un coup d'œil autour de lui sur la planète du géographe.	And he cast a look around at the planet of the geographer. <i>And he/it glanced around him at the geographer's planet.</i>
Ce n'est pas le géographe qui va faire le compte	It is not the geographer who goes out to count the

<p>des villes, des fleuves, des montagnes, des mers, des océans et des déserts.</p>	<p>towns, the rivers, the mountains, the seas, the oceans, and the deserts.</p> <p><i>This is not the geographer who is going to make the account of cities, streams, mountains, seas, oceans, and deserts.</i></p>
<p>Le géographe est trop important pour flâner. Il ne quitte pas son bureau. Mais il y reçoit les explorateurs.</p>	<p>The geographer is much too important to go loafing about. He does not leave his desk. But he receives the explorers in his study.</p> <p><i>The geographer is too important to stroll. He/it doesn't leave his/her/its office. But he/it receives explorers there.</i></p>
<p>Parce que les ivrognes voient double. Alors le géographe noterait deux montagnes là où il n'y en a qu'une seule.</p>	<p>Because intoxicated men see double. Then the geographer would note down two mountains in a place where there was only one.</p> <p><i>Because drunkards see double. Then the geographer would note two mountains in a place where there is only one of it alone.</i></p>
<p>Je connais quelqu'un, dit le petit prince, qui serait mauvais explorateur.</p>	<p>I know someone, said the little prince, who would make a bad explorer.</p> <p><i>I know someone, said the little prince, who would be bad explorer.</i></p>
<p>Mais toi, tu viens de loin!</p>	<p>But you - you come from far away!</p> <p><i>But you, you come from afar!</i></p>
<p>Et le géographe, ayant ouvert son registre, tailla son crayon.</p>	<p>And, having opened his big register, the geographer sharpened his pencil.</p> <p><i>And the geographer, having opened his/her/its register, carved his/her/its pencil.</i></p>

<p>J'ai trois volcans. Deux volcans en activité, et un volcan éteint. Mais on ne sait jamais.</p>	<p>I have three volcanoes. (Two volcanoes are active and the other is extinct.) But one never knows.</p> <p><i>I have three volcanoes. (Two volcanos in activity and a volcano extinguishes). But one never knows.</i></p>
<p>- On ne sait jamais, dit le géographe.</p> <p>- J'ai aussi une fleur.</p> <p>- Nous ne notons pas les fleurs, dit le géographe.</p>	<p>"One never knows", said the geographer.</p> <p>"I have also a flower"</p> <p>"We do not record flowers", said the geographer.</p> <p><i>"One never knows", said the geographer.</i></p> <p><i>"I also have a flower"</i></p> <p><i>"We don't note flowers", said the geographer.</i></p>
<p>Il est très rare qu'une montagne change de place. Il est très rare qu'un océan se vide de son eau.</p>	<p>It is very rarely that a mountain changes its position. It is very rarely that an ocean empties itself of its water.</p> <p><i>He/it is very rare that a mountain changes room.</i></p> <p><i>He/it is very rare that an ocean empties himself of his/her/its water.</i></p>
<p>Elle (= la montagne) ne change pas.</p>	<p>It does not change.</p> <p><i>She/it doesn't change.</i></p>
<p>La septième planète fut donc la terre. La Terre n'est pas une planète quelconque !</p>	<p>So then the seventh planet was the Earth. The Earth is not just an ordinary planet!</p> <p><i>The seventh planet was therefore the Earth. The Earth is not an any planet!</i></p>
<p>Pour vous donner une idée des dimensions de la Terre, je vous dirai qu'avant l'invention de l'électricité on y devait entretenir, sur l'ensemble des six continents, une véritable armée de quatre cent soixante-deux mille cinq cent onze allumeurs</p>	<p>To give you an idea of the size of the Earth, I will tell you that before the invention of electricity it was necessary to maintain, over the whole of the six continents, a veritable army of 462,511 lamplighters for the street lamps.</p>

de réverbères.	<i>To give you an idea of measurements of the Earth, I will tell you that before the invention of one had to maintain there, on the set of the six continents, a real army of four hundred sixty-two thousand five hundred eleven lighters of lampposts.</i>
Les mouvements de cette armée étaient réglés comme ceux d'un ballet d'opéra.	The movements of this army could be regulated like those of the ballet in the opera. <i>Movements of this army were adjusted like those of an opera ballet.</i>
D'abord venait le tour es allumeurs de réverbères de Nouvelle-Zélande et d'Australie.	First would come the turn of the lamplighters of New Zealand and Australia. <i>First came the tour of the lighters of lampposts of New Zealand and Australia.</i>

8.6 Human and machine identical translations (in terms of syntax) - German

Source sentences	Human/Machine translations
Lorsque j'avais six ans j'ai vu, une fois, une magnifique image, dans un livre sur la Forêt Vierge qui s'appelait « Histoires Vécues ».	Als ich sechs Jahre alt war, sah ich einmal in einem Buch ueber den Urwald, das « Erlebte Geschichten » hiess, ein praechtiges Bild. <i>Wenn ich sechs Jahre alt war, habe ich gesehen, ein Mal, ein herrliches Bild, in einem Buch auf dem Urwald der « Geschichten [Vécues] » hiess.</i>
Mon dessin n°1.	Meine Zeichnung Nr 1. <i>Meine Zeichnung Nummer 1.</i>
Mon dessin ne représentait pas un chapeau. Il représentait un serpent boa qui digérait un éléphant.	Meine Zeichnung stellte aber keinen Hut dar. Sie stellte eine Riesenschlange, dar, die einen Elefanten verdaut. <i>Meine Zeichnung stellte keinen Hut dar. Er stellte eine boa-Schlange, dar, die einen Elefanten verdaute.</i>
Quelque chose s'était cassé dans mon moteur.	Etwas an meinem Motor war kaputtgegangen. <i>Etwas war in meinem Motor gebrochen.</i>
C'était pour moi une question de vie ou de mort.	Es war fuer mich eine Frage [auf] Leben und Tod. <i>Dies war eine Frage [von] Leben oder Totem fuer mich.</i>
Mais... Qu'est-ce que tu fais là ?	Aber... Was machst [denn] du da ? <i>Aber... Was machst du dort ?</i>
Et il me répéta alors, tout doucement, comme une chose très sérieuse :	Da wiederholte es ganz sanft, wie eine sehr ernsthafte Sache : <i>Und er wiederholte mich dann ganz sanft, wie</i>

	<i>eine sehr ernste Sache :</i>
Dessine-moi un mouton...	Zeichne mir ein Schaf... <i>Zeichne mir ein Schaf...</i>
Quand le mystère est trop impressionnant, on n'ose pas désobéir.	Wenn das Geheimnis zu eindrucksvoll ist, wagt man nicht zu widerstehen. <i>Wenn das Geheimnis zu beeindruckend ist, wagt man nicht, nicht zu gehorchen.</i>
Ça ne fait rien. Dessine-moi un mouton.	Das macht nichts. Zeichne mir ein Schaf. <i>Das macht nichts. Zeichne mir ein Schaf.</i>
J'ai besoin d'un mouton. Dessine-moi un mouton. Alors j'ai dessiné.	Ich brauche ein Schaf. Zeichne mir ein Schaf. Also habe ich gezeichnet. <i>Ich brauche ein Schaf. Zeichne mir ein Schaf.</i> <i>Dann habe ich gezeichnet.</i>
Non ! Celui-là est déjà très malade. Fais-en un autre.	Nein ! Das ist schon sehr krank. Mach ein anderes. <i>Nein ! Jener ist schon sehr krank. Mache noch einen anderen.</i>
Mon ami sourit gentiment, avec indulgence.	Mein Freund laechelte artig und mit Nachsicht. <i>Mein Freund laechelt nett, mit Nachsicht.</i>
Je refis encore mon dessin.	Ich machte also meine Zeichnung noch einmal. <i>Ich machte noch einmal also noch meine Zeichnung.</i>
Mais il fut refusé, comme les précédents.	Aber sie wurde ebenso abgelehnt wie die vorigen. <i>Aber er wurde abgelehnt, wie die Praezedenzfaelle.</i>
Celui-là est trop vieux. Je veux un mouton qui vive longtemps.	Das ist schon zu alt. Ich will ein Schaf, das lange lebt.

	<i>Jener ist zu alt. Ich will ein Schaf, das lange lebt.</i>
Je t'ai donné un tout petit mouton.	Ich habe dir ein ganz kleines Schaf geschenkt. <i>Ich habe dir ein ganz kleines Schaf gegeben.</i>
Il pencha la tête vers le dessin.	Er neigte den Kopf ueber die Zeichnung. <i>Er neigte den Kopf gegen die Zeichnung.</i>
Il s'est endormi.	Es ist eingeschlafen. <i>Er ist eingeschlafen.</i>
Et c'est ainsi que je fis la connaissance du petit prince.	So machte ich die Bekanntschaft des kleinen Prinzen. <i>Und so machte ich die Kenntnis aus dem kleinem Prinzen.</i>
Ainsi, quand il aperçut mon avion pour la première fois (je ne dessinerai pas mon avion, c'est un dessin beaucoup trop compliqué pour moi), il me demanda :	So fragte er, als er zum erstenmal mein Flugzeug sah (ich werde mein Flugzeug nicht zeichnen, das ist eine viel zu komplizierte Sache fuer mich) : <i>So als er mein Flugzeug zum ersten Mal erblickte, werde ich mein Flugzeug nicht zeichnen, dies ist sehr zu viel eine Zeichnung fuer mich kompliziert, er fragte [mich] :</i>
Ce n'est pas une chose. Ça vole. C'est un avion. C'est mon avion.	Das ist kein Ding. Das fliegt. Das ist ein Flugzeug. Es ist mein Flugzeug. <i>Dies ist keine Sache. Das fliegt. Dies ist ein Flugzeug. Dies ist mein Flugzeug.</i>
Alors il s'écria : - Comment ! tu es tombé du ciel ?	Da rief er : -Wie !Du bist von Himmel gefallen ? <i>Dann rief er aus : - Wie ! Bist du von Himmel gefallen?</i>
Alors, toi aussi tu viens du ciel ! De quelle planète es-tu ?	Also auch du kommst von Himmel ! Von welchem Planeten bist du denn ?

	<i>Dann kommst du auch von Himmel! Von welchem Planeten bist du?</i>
Tu viens donc d'une autre planète ?	Du kommst also von einem anderen Planeten ? <i>Kommst du also von einem anderen Planeten?</i>
Mais il ne me répondit pas.	Aber er antwortete nicht. <i>Aber er antwortete mich nicht.</i>
Et il s'enfonça dans une rêverie qui dura longtemps.	Und er versankt in eine Traeumerei, die lange dauerte. <i>Und er sprengte in einer Traeumerei auf, die lange dauerte.</i>
D'où viens-tu , mon petit bonhomme ?	Woher kommst du, mein kleines Kerlchen ? <i>Woher kommst du, mein kleines Maennchen?</i>
Où veux-tu emporter mon mouton ?	Wohin willst du mein Schaf mitnehmen ? <i>Wohin willst du mein Schaf mitnehmen?</i>
Il me répondit après un silence méditatif :	Er antwortete nach einem nachdenklichen Schweigen : <i>Er antwortete mich nach einer meditativen Stille :</i>
La proposition parut choquer le petit prince :	Dieser Vorschlag schien den kleinen Prinzen zu kraenken: <i>Der Vorschlag schien den kleinen Prinzen zu schocken :</i>
Alors le petit prince remarqua gravement :	Da versetzte der kleine Prinz ernsthaft: <i>Dann bemerkte der kleine Prinz ernsthaft:</i>
Droit devant soi on ne peut pas aller bien loin...	Geradeus kann man nicht sehr weit gehen... <i>Gerade vor einen kann man nicht sehr weit gehen...</i>
J'avais ainsi appris une seconde chose très	Ich hatte eine zweite sehr wichtige Sache

importante :	erfahren : <i>Ich hatte so eine zweite sehr wichtige Sache gelernt:</i>
Quand un astronome découvre l'une d'elles, il lui donne pour nom un numéro.	Wenn ein Astronom einen von ihnen entdeckt, gibt er ihm [statt des Namens] eine Nummer. <i>Wenn ein Astronom den einen von ihnen entdeckt, gibt er ihr/ihm eine Nummer [fuer Namen].</i>
J'ai de sérieuses raisons de croire que la planète d'où venait le petit prince est l'astéroïde B612.	Ich habe ernsthafte Gruende zu glauben, dass der Planet, von [dem] der kleine Prinz kam, der Asteroid B612 ist. <i>Ich habe ernste Gruende, zu glauben, dass der Planet, von [wo] der kleine Prinz kam, der Asteroid B612 ist.</i>
Il avait fait alors une grande démonstration de sa découverte à un congrès International d'Astronomie.	Er hatte damals beim internationalen Astronomenkongress einen grossen Vortrag ueber seine Entdeckung gehalten. <i>Er hatte dann eine grosse Demonstration aus seiner/ihrer Entdeckung an einem Internationalen Astronomie-Kongress gemacht.</i>
Les grandes personnes sont comme ça.	Die grossen Leute sind so. <i>Die grossen Personen sind so.</i>
Quand vous leur parlez d'un nouvel ami, elles ne vous questionnent jamais sur l'essentiel.	Wenn [ihr] ihnen von einem neuen Freund erzahlt, befragen sie [euch] nie ueber das Wesentliche. <i>Wenn [Sie] mit ihnen von einem neuen Freund sprechen, befragen sie [Sie] nie ueber die Hauptsache.</i>

Est-ce qu'ils collectionne les papillons ?	Sammelt er Schmetterlinge? <i>Sammelt er die Schmetterling?</i>
Elles vous demandent : Quel âge a-t-il ? Combien a-t-il de frères ? Combien pèse-t-il ? Combien gagne son père ?	Sie fragen [euch] : Wie alt ist er? Wieviel Brueder hat er? Wieviel wiegt er? Wieviel verdient sein Vater? <i>Sie fragen [Sie] : Wie alt ist er? Wieviel Brueder hat er? Wieviel wiegt er? Wieviel gewinnt sein/ihr Vater?</i>
Si j'essaie ici de le décrire, c'est afin de ne pas l'oublier .	Wenn ich hier versuche, ihn zu beschreiben, [so tue ich das], um ihn nicht zu vergessen. <i>Wenn ich hier versuche, es/sie/ihn zu beschreiben, [ist dies], um ihn nicht zu vergessen.</i>
C'est triste d'oublier un ami.	Es ist traurig, einen Freund zu vergessen. <i>Es ist traurig, einen Freund zu vergessen.</i>
Je me trompe aussi un peu sur la taille. Ici le petit prince est trop grand. Là il est trop petit.	Ich irre mich auch mitunter in den Massen. Da ist der kleine Prinz zu gross und da ist er zu klein. <i>Ich irre mich so ein bisschen auf der Groesse. Hier ist der kleine Prinz zu gross. Dort ist er zu klein.</i>
Chaque jour j'apprenais quelque chose sur la planète, sur le départ, sur le voyage.	Jeden Tag erfuhr ich etwas [Neues] ueber den Planeten, ueber die Abreise [und] ueber die Fahrt. <i>Jeden Tag lernte ich etwas auf dem Planeten, auf der Abfahrt, auf der Reise.</i>
Mais le petit prince ajouta : - Par conséquent ils mangent aussi les baobabs ?	Aber der kleine Prinz fuegte hinzu: - Dann fressen sie doch auch Affenbrotbaeume? <i>Aber der kleine Prinz fugte hinzu: - Essen sie infolgedessen auch die Affenbrotbaeume ?</i>

<p>Mais pourquoi veux-tu que tes moutons mangent les petits baobabs ?</p>	<p>Aber warum willst du, dass deine Schafe die kleinen Affenbrotbaeume fressen? <i>Aber warum willst du, dass deine Schafe die kleinen Affenbrotbaeume essen?</i></p>
<p>Et en effet sur la planète du petit prince, il y avait comme sur toutes les planètes, de bonnes herbes et de mauvaises herbes.</p>	<p>In der Tat gab es auf dem Planeten des kleinen Prinzen wie auf allen Planeten gute Gewaechse und schlechte Gewaechse. <i>Und in der Tat, auf dem Planeten des kleinen Prinzen gab es wie auf allen Planeten guten Kraeutern und schlechten Kraeutern.</i></p>
<p>Mais les graines sont invisibles.</p>	<p>Aber die Samen sind unsichtbar. <i>Aber die Kerne sind unsichtbar.</i></p>
<p>Or il y avait des graines terribles sur la planète du petit prince...c'étaient les graines de baobabs.</p>	<p>Auf dem Planeten des kleinen Prinzen gab es fuerchterliche Samen...und das waren die Samen der Affenbrotbaeume. <i>Aber es gab fuerchterliche Kerne auf dem Planeten des kleinen Prinzen ...dies waren die Kerne von Affenbrotbaeumen.</i></p>
<p>C'est un travail très enuyeux, mais très facile.</p>	<p>Das ist eine zwar langweilige, aber leichte Arbeit. <i>Dies ist eine sehr langweilige, aber sehr leichte Arbeit.</i></p>
<p>Il avait négligé trois arbustes.</p>	<p>Er hatte drei Straecher uebersehen. <i>Er hatte drei Straechervernachlaessigt.</i></p>
<p>Et, sur les indications du petit prince, j'ai dessiné cette planète-là.</p>	<p>Und habe ich denn diesen Planeten nach den Angaben des kleinen Prinzen gezeichnet. <i>Und auf den Hinweisen des kleinen Prinzen habe ich diesen Planeten gezeichnet.</i></p>

Je n'aime guère prendre le ton d'un moraliste.	Ich nehme nicht gerne den Tonfall eines Moralisten an. <i>Ich nehme nicht gern kaum den Ton eines Moralisten.</i>
La réponse est bien simple.	Die Antwort ist sehr einfach. <i>Die Antwort ist sehr einfach.</i>
Ah ! Petit prince, j'ai compris, peu à peu, ainsi, ta petite vie mélancolique.	Ach, kleiner Prinz, so nach und nach habe ich dein kleines schwermuetiges Leben verstanden. <i>Ach ! Kleiner Prinz, ich habe nach und nach so verstanden dein kleines melancholisches Leben.</i>
Et un peu plus tard tu ajoutais : - Tu sais...quand on est tellement triste on aime les couchers de soleil...	Und ein wenig spaeter fuegtest du hinzu : - Du weisst doch, wenn man recht traurig ist, lieb man die Sonnenuntergaenge. <i>Und ein bisschen spaeter fuegtest du hinzu : - Du weisst...wenn man dermassentraurig ist, liebt man die Sonne-Untergaenge.</i>
Mais le petit prince ne répondit pas.	Aber der kleine Prinz antwortete nicht. <i>Aber der kleine Prinz entsprach nicht.</i>
Même les fleurs qui ont des épines ?	Auch die Blumen, die Dornen haben? <i>Selbst die Blumen, die Dornen haben?</i>
Je ne le savais pas.	Ich wusste es nicht. <i>Ich wusste es nicht.</i>
Le petit prince ne renonçait jamais à une question, une fois qu'il l'avait posée.	Der kleine Prinz verzichtete niemals auf eine Frage, wenn er sie einmal gestellt hatte. <i>Der kleine Prinz verzichtete nie auf eine Frage, sobald er sie gestellt hatte.</i>
Les fleurs sont faibles. Elles sont naïves. Elles se	Die Blumen sind schwach. Sie sind arglos. Sie

rassurent comme elles peuvent.	schuetzen sich, wie sie koennen. <i>Die Blumen sind schwach. Sie sind naiv. Sie beruhigen sich, als sie koennen.</i>
Le petit prince déranga de nouveau mes réflexions.	Der kleine Prinz stoerte meine Ueberlegungen von neuem. <i>Der kleine Prinz stoerte meine Ueberlegungen von neuem.</i>
Et tu crois, toi, que les fleurs...	Und du glaubst, dass die Blumen... <i>Und du glaubst, [du,] dass die Blumen...</i>
Mais non ! Mais non ! Je ne crois rien ! J'ai répondu n'importe quoi.	Aber nein! Aber nein! Ich glaube nichts! Ich habe irgend etwas dahergeredet. <i>Aber nein! Aber nein! Ich glaube nichts! Ich habe irgendetwas geantwortet.</i>
Il me regarda stupéfait.	Er schaute mich verdutzt an. <i>Er betrachtete mich verdutzt.</i>
Tu parles comme les grandes personnes !	Du sprichst ja wie die grossen Leute! <i>Du sprichst wie die grossen Personen!</i>
Mais, impitoyable, il ajouta :	Er aber fuegte unbarmherzig hinzu : <i>Aber, erbarmungslos, er fuegte hinzu :</i>
Il était vraiment très irrité.	Er war wirklich sehr aufgebracht. <i>Er war wirklich sehr erregt.</i>
Il n'a jamais regardé une étoile.	Er har nie einen einen Stern eingeschaut. <i>Er har nie einen einen Stern betrachtet.</i>
Il n'a jamais aimé personne.	Er hat nie jemanden geliebt. <i>Er hat nie jemanden geliebt.</i>
Et toute la journée il répète comme toi : "Je suis un homme sérieux! Je suis un homme sérieux !"	Und den ganzen Tag wiederholt er wie du: Ich bin ein ernsthafter Mann! Ich bin ein ernsthafter

	<p>Mann!</p> <p><i>Und den ganzen Tag wiederholt er wie du: Ich bin ein ernster Mann! Ich bin ein ernster Mann!</i></p>
Mais, ce n'est pas un homme, c'est un champignon!	<p>Aber das ist kein Mensch, das ist ein Schwamm.</p> <p><i>Aber dies ist kein Mann, dies ist ein Pilz!</i></p>
Le petit prince était maintenant tout pâle de colère.	<p>Der kleine Prinz war jetzt ganz blass vor Zorn.</p> <p><i>Der kleine Prinz war jetzt vor Argern ganz blass.</i></p>
Il y a des millions d'années que les fleurs fabriquent des épines.	<p>Es sind nun Millionen Jahre, dass die Blumen Dornen hervorbringen.</p> <p><i>Es gibt Millionen Jahre, dass die Blumen Dornen anfertigen.</i></p>
Il y a des millions d'années que les moutons mangent quand même les fleurs.	<p>Es sind Millionen Jahre, dass die Schafe trotzdem die Blumen fressen.</p> <p><i>Es gibt Millionen Jahre, dass die Schafe dennoch die Blumen essen.</i></p>
Il se dit: "Ma fleur est là, quelque part..."	<p>Er sagt sich: Meine Blume ist da oben, irgendwo...</p> <p><i>Er sagt sich: Meine Blume ist dort, irgendwo...</i></p>
Il ne put rien dire de plus.	<p>Er konnte nichts mehr sagen.</p> <p><i>Er konnte nichts sagen (von) mehr.</i></p>
Il éclata brisquement en sanglots.	<p>Er brach ploetzlich in Schluchzen aus.</p> <p><i>Er explodierte ploetzlich in Schlurchzern.</i></p>
La nuit était tombée.	<p>Die Nacht war hereingebrochen.</p> <p><i>Die Nacht war gefallen.</i></p>
J'avais lâché mes outils.	<p>Ich hatte mein Werkzeug weggelegt.</p> <p><i>Ich hatte meine Werkzeuge fallengelassen.</i></p>
Je le berçais.	<p>Ich wiegte ihn.</p>

	<i>Ich wiegte es/sie/ ihn.</i>
La fleur que tu aimes n'est pas en danger...	Die Blume, die du liebst, ist nicht in Gefahr... <i>Die Blume, die du liebst, ist nicht in Gefahr...</i>
Je lui dessinerai une muselière à ton mouton...	Ich werde ihm einen Maulkorb zeichnen, deinem Schaf... <i>Ich werde ihr/ihm ein Maulkorb zeichnen, an deinem Schaf...</i>
Je te dessinerai une armure pour ta fleur...	Ich werde ihm einen Zaun für deine Blume zeichnen... <i>Ich werde dir eine Ruestung fuer deine Blume zeichnen...</i>
C'est tellement mystérieux, le pays des larmes.	Es ist so geheimnisvoll, das Land der Traenen. <i>Dies ist dermassen mysterios, das Land der Traenen.</i>
Elles apparaissaient un matin dans l'herbe, et puis elles s'éteignaient le soir.	Sie leuchteten eines Morgens im grase auf und erloschen am Abend. <i>Sie erschienen eines Morgens im Kraut, und dann erloschen sie den Abend.</i>
Ça pouvait être un nouveau genre de baobab.	Das konnte eine neue Art Affenbrotbaum sein. <i>Das konnte eine neue Affenbrotbaum- Art sein.</i>
Mais l'arbuste cessa vite de croître, et commença de préparer une fleur.	Aber der Strauch hoerte bald auf zu wachsen und begann, eine Bluete anzusetzen. <i>Aber der Strauch hoerte schnell auf, zu wachsen, und fing an, eine Blume vorzubereiten.</i>
Elle choisissait avec soin ses couleurs. Elle s'habillait lentement, elle ajustait un à un ses pétales.	Sie waelhte ihre Farben mit Sorgfalt, sie zog sich langsam an, sie ordnete ihre Bluetenblaetter eins nach dem andern.

	<i>Sie waelhte seine/ihre Farben mit Sorge. Sie zog sich langsam an, sie justierte einer nach dem anderen seine/ihre Bluetenblaetter.</i>
Sa toilette mystérieuse avait duré des jours et des jours.	Ihre geheimnisvolle Toilette hatte also Tage und Tage gedauert. <i>Seine/Ihre mysterioese Toilette hatte also Tage und die Tage gedauert.</i>
Et puis voici qu'un matin, justement à l'heure du lever du soleil, elle s'était montrée.	Und dann, eines Morgens, gerade zur Stunde des Sonnenaufganges, hatte sie sich enthuellt. <i>Und dann jetzt einen Morgen, eben an der Stunde des Sonnenaufganges, hatte sie sich gezeigt.</i>

8.7 Barcelona Machine Translation Technology (by courtesy of Globalink)

- **Barcelona Architecture**

The Barcelona technology follows the transfer system model. The fundamental components of Barcelona are:

1. Dictionaries

For each language pair Barcelona uses two dictionaries, one for each language, independent from each other. The dictionaries contain translations, morphological data and other elements that connect them with the morphology tables and the rule. Terms can be added or modified within the dictionaries.

2. Morphology Tables

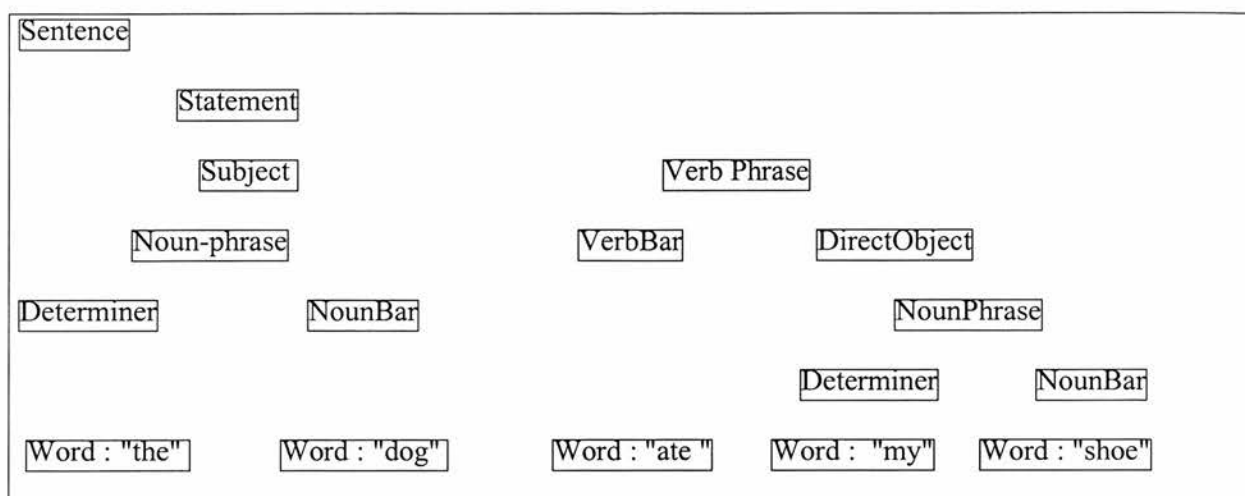
Morphology tables contain inflections for any valid part of speech. For example, morphology tables allow for proper declension of nouns in German and proper conjugation of verbs.

3. Rule system

Barcelona's rule system is a proprietary programming language that creates, annotates and reorders syntactical structures from the source text into the target text.

Barcelona gathers information about a sentence by building a tree that corresponds to the structure of the sentence. The middle nodes of the tree correspond to things like noun phrases and prepositional phrases. Below is a conceptual diagram of a tree for the sentence,

« *The dog ate my shoe.* »



The purpose of the Barcelona Rule Language is to build these trees. Each rule in the rule system contains a description of a small part of the tree, called the *Pattern*. Each rule also describes modifications that will be made to the tree, called the *Production*. Barcelona builds trees by repeatedly comparing the current tree with the pattern part of each rule; if a rule's pattern matches part of the tree, then Barcelona modifies the tree according to the production.

4. Engine

The core of the MT system, the engine, is responsible for the application of linguistic rules and for handling the information used in the analysis of the source language and the production of translations in the target language.

The Barcelona Translation Process

1. Word Analysis Stage

When the engine begins a translation, it identifies the input data (a string of characters that represent the source sentence). This sentence is in its raw form, exactly as created by the user. The engine scans the string, looking for contractions so that to expand to their de-contracted form. For example, the word *can't* would be re-written internally as *can not*.

After processing contractions, the program begins looking up the words in the dictionary. If the program can not find a word in the dictionary, Barcelona checks whether the word is a compound of several words or affixes. The words, along with their dictionary information are placed in a tree structure, for processing by the linguistic rules.

2. Surface Analysis and Disambiguation Stage

The first set of rules that act on the tree are the Surface Rules, whose main job is to discover the actual part of speech for any words that can have more than one part of speech. The Disambiguation Rules perform this task. Once the Disambiguation Rules have finished their task, each word in the tree is reduced to a single part of speech possibility. However, because Barcelona cannot always determine what part of speech a word has in a surface level analysis, it is sometimes necessary to create multiple trees. Each tree is a possible interpretation of the sentence, and each word of each tree has a single, known part of speech. The engine continues analysing all these trees in parallel.

3. Simple Phrase Analysis Stage

Once each part of speech usage is made unambiguous, Barcelona begins assembling the simple phrase structures of the language, like noun phrases and adverb phrases. At this stage Barcelona does not yet process verbs or the more complicated phrases like preposition phrases. The rules that do this are the Phrase Reduction Rules. At this point, many lexicon rules are executed, which can make changes to the source tree that are not limited as surface analysis rules but also do not need to work with a complete tree.

4. Full Source Analysis Stage

Now Barcelona begins building the complete analysis tree. The engine combines phrases of the sentence into phrases that are more complicated and Barcelona builds verb-phrases, statements, and relative and subordinate clauses. To handle the special problem of preposition attachment (whether to a verb or to a preceding noun), Barcelona will often create two or more trees that differ in the way they attach the preposition. These rules are the Statement Reduction Rules, and are arguably the most complicated rules in the system. Once the Statement Reduction Rules have finished, the engine goes to the lexicon and executes any Source to Source Lexicon Rules,

which can make final changes to the tree while it is still in its native language. The preceding rules have possibly many different trees, from which the engine now needs to choose the single best tree. Barcelona uses several heuristics and rules to arrive at its choice, which is called the Final Analysis Tree. The last thing that the engine does to the Final Analysis Tree is to traverse it, setting various attributes. For instance, this final traversal would look at the word *ate* and set a flag that says that this word is in the preterite tense.

5. Target Transfer Stage

Now that the engine has built the source language tree, Barcelona begins to transfer its knowledge into the target language. As a first step, the engine simply transfers many attributes from the source language to the target language. Object information attached to the source verb is used to perform Verb Phrase level transformations. Verbal Object information is stored as special rules in the lexicon and the engine uses these rules to describe how the target verb phrase should look versus the source verb phrase. The lexicon also provides many rules at this stage, which can cause certain words or phrases to have a specific translation.

6. Target Rewrite Stage

At this point, the engine has a tree that contains the right words, but the wrong structure. Routines are required to change the order of adjectives and nouns, or move the verbs to the end of the statements. Target Rewrite Rules perform the transformations that make the tree into a target language tree.

In addition to simple transformations, the engine also performs some more complicated transformations at this stage. For instance, English uses adverbs in many places where romance languages use subordinate clauses – in English one says « normally » where in Spanish one says « It is normal that. » Complex transformations like this can be done by Target Rewrite Rules.

7. Final Stage

The purpose of the final stage is to output the sentence. However, the engine still needs to inflect the words to make them agree; nouns and adjectives must agree in number and gender, and verbs must agree with their subjects. The engine traverses the tree, telling each word to **agree with** other words or phrases. Now, a part of the engine called the morphology system

uses the agreement information to inflect all words to agree according to the rules of that particular language. The final step in translation is to perform any contractions and convert the words back into a string.

8.8 Number of sentences in the human translations (indicating a tendency to translate sentence-by-sentence)

	FRENCH	ENGLISH	GERMAN	RUSSIAN
Chapter 1	38	38	36	37
Chapter 2	73	74	70	70
Chapter 3	48	50	50	48
Chapter 4	63	64	60	60
Chapter 5	52	53	51	52
Chapter 6	22	24	22	13
Chapter 7	79	87	81	80
Chapter 8	61	69	60	67
Chapter 9	52	51	50	51
Chapter 10	111	118	106	108
Chapter 11	29	30	27	27
Chapter 12	12	13	12	13
Chapter 13	106	104	99	104
Chapter 14	74	75	72	77
Chapter 15	87	84	86	91
Chapter 16	17	13	17	17
Chapter 17	54	58	53	61
Chapter 18	14	14	13	14
Chapter 19	16	17	15	16
Chapter 20	19	20	18	19
Chapter 21	144	141	145	147
Chapter 22	23	23	23	26
Chapter 23	11	11	11	10

Chapter 24	57	59	57	60
Chapter 25	77	81	78	78
Chapter 26	171	180	176	171
Chapter 27	32	30	34	34
Chapter 28	8	9	6	7
TOTAL	1550	1590	1528	1558

9 BIBLIOGRAPHY

- Adorni, G. & L. Massone. 1987. 'Toward a language-independent generator of sentences', *Applied Artificial Intelligence*, 1:53-75
- Alexieva, B. 1993. 'A cognitive approach to translation equivalence', in Zlateva (ed.), 101-109
- Allen, J. 1987. *Natural Language Understanding* (Menlo Park, California: The Benjamin/Cummings Publishing Company)
- Alpac. 1966. *Languages and Machines: Computers in Translation and Linguistics*, a report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Research Council (Washington, DC: National Academy of Sciences)
- Anis, J. 1994. 'Ordinateurs et traduction: survol d'un demi-siècle', *Langage*, 116: 111-122
- Arnold D. 1990. 'Text Typology and Machine Translation', in Mayorcas (ed.), 73-79
- Arnold, D., L. Balken, L. Humphreys, S. Meijer, & L. Sadler (eds). 1994 *Machine Translation. An Introductory Guide* (London: NCC Blackwell)
- Arsac, J. 1984. 'L'informatique et le sens', in Dreyfus, *Intelligence artificielle, mythes et limites*, 411-421
- Austermuehl, F. 2001. *Electronic Tools for Translators* (Manchester: St Jerome)
- Backus, J. 1978. 'Can programming be liberated from the von Neumann style?', *Comm. ACM*, 21(8): 411-18
- Baker M. (ed.) 2001. *Routledge Encyclopedia of Translation Studies* (London: Routledge)
- Ballard, M. 1998. 'Idiomatisme et traduction', *La Tribune Internationale des Langues Vivantes*, 23: 24-29
- Bar-Hillel, Y. 1962. 'The Future of Machine Translation', *Times Literary Supplement* (London: Times Newspaper, April 20th, 1962)
- Bara, B.G. & G. Guida 1984. *Computational Models of Natural Language Processing* (Amsterdam: Elsevier Science Publishers B.V.)

- Barkhudarov, L. 1993. 'The problem of the Unit of Translation', in Zlateva (ed.), 39-46
- Barthes, R. 1957/1987. *Mythologies* (Paris: Seuil)
- Barthes, R. 1983 *A Barthes Reader*, ed. Susan Sontag. (New York: Hill & Wang)
- Barnstone, W. 1994. 'Translation Theory with a Semiotic Slant', *Semiotica*, 102-1/2: 89-100
- Barr, A. & E.A. Feigenbaum, 1981 *The Handbook of Artificial Intelligence, Vol. I* (Los Altos (Cal.) : William Kaufmann)
- Beaugrande, R.A. de & W.U. Dressler 1981 *Introduction to Text Linguistics* (London: Longman)
- Bell, R.T., 1991. *Translation and Translating: Theory and Practice* (New York: Longman, Inc.)
- Bennett, P. 1994. 'The Translation Unit in Human and Machine', *Babel*, 40: 12-20
- Bigelow, J. 1978. 'Semantics of Thinking, Speaking and Translation', in Guenther & Guenther-Reutter (eds.), 109-135
- Biskri, I. & J.P. Desclés 1996. 'La grammaire catégorielle combinatoire applicative', *TAL*: 37/2: 71-84
- Blanche, R. & J. Dubucs, J. 1996. *La logique et son histoire* (Paris: Armand Colin/Masson)
- Blatt, A., & K. H. Freigang, K. D. Schmitz, G. Thome 1985 *Computer und Übersetzen. Eine Einführung* (Hildesheim: Georg Olms Verlag)
- Blühndorn, H. 1998. 'The relation between Pragmatics, Semantics and Grammar and the Notion of Linguistic Coherence', *S- European Journal for Semiotic Studies*, 10(1,2): 25-71
- Blum-Kulka, S. 2000. 'Shifts of cohesion and coherence in translation', in Venuti, 2000: 298-313.
- Boitet C., & H. Blanchon 1994. 'Promesses et problèmes de la "TAO pour tous" après Lidia 1, une première maquette', *Langage*, 116: 20-47
- Bolinger, D. 1965. 'The Atomization of Meaning', *Language*, vol 41/4: 557-573
- Bouton, C. 1984. *La linguistique appliquée* (Paris: PUF)
- Brame, M. 1976. *Conjectures and Refutations in Syntax and Semantics* (New York & Amsterdam: North Holland Publishing Company)
- Brower, R. (ed.) 1959. *On Translation* (Cambridge, Massachusetts: Harvard University Press)

- Brown, K., & J. Miller (eds.) 1996. *Concise Encyclopedia of Syntactic Theories* (Cambridge, UK: Cambridge University Press)
- Bühler, K., 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache* (Jena: G. Fischer)
- Butler F., & R. Rotert (eds.) 1984. *Reflections on Literature for Children* (USA: Library Professional Publications)
- Capestany, E. 1982. *The Dialectic of the Little Prince* (USA: University Press of America)
- Cappelli, A., & G. Ferrari, L. Moretti, I. Prodanov 1984. 'A Framework for Integrating Syntax and Semantics'. In Bara, B.G. & G. Guida, G., 33-57
- Carbonell et al. 1981. *Steps Towards Knowledge-Based Machine Translation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 3
- Carnap, R., 1956. *Meaning and Necessity, 2nd edition with Supplements* (Chicago: University of Chicago Press.
- Catford, J.C. 1965. *A Linguistic Theory of Translation: An Essay in Applied Linguistics* (London: Oxford University Press)
- Catford, J.C. 2000. 'Translation Shifts', in Venuti (ed.), 141-147
- Cavazza, M. & P. Zweigenbaum 1992. 'Comprehension automatique du langage naturel par construction de modèles', *Techniques et sciences informatiques*, 11/4: 119-138
- Cavazza, M. & P. Zweigenbaum 1994. 'Semantic analysis and in-depth understanding of technical texts', *Applied Artificial Intelligence*, 8(3): 425-453
- Chambreuil, M. & A. Ben Gharbia, P. Gamallo Otero 1998. 'Variations sur la Compositionnalité Montaguienne', *Traitement Automatique du Langage*, 39(1): 35-65
- Chandler, D. 2002. *Semiotics: The Basics* (London: Routledge)
- Chazal, G., & M-N. Terrasse 1996. *Philosophie du langage et informatique* (Paris: Editions Hermès)
- Chernyakhovska, L. 1993. 'Sense and its Expression through Language', in Zlateva (ed.), 110-121
- Chesterman, A. 1997. *Memes of Translation* (Amsterdam/Philadelphia: John Benjamins Publishing Company)

- Chomsky, N. 1965. *Aspects of the Theory of Syntax* (Cambridge, MA: The MIT Press)
- Chomsky, N. 1969. *Structures syntaxiques* (Paris: Editions du Seuil)
- Chomsky, N. 1981. *Réflexions sur le langage* (Paris: Flammarion)
- Chuquet, H. & M. Paillard 1989. *Approche linguistique des problèmes de traduction* (Paris: Ophrys)
- Comité de l'Association des Amis d'Antoine de Saint-Exupéry (ed.) 1980. *Cahiers Saint-Exupéry* (Paris: Gallimard)
- Church, A. 1941. *The Calculi of Lambda Conversion* (USA: Princeton University Press)
- Church, A. 1951. 'A Formulation of the Logic of Sense and Denotation'. In P. Henke, H. Kallen, & S. Langer (eds.), *Structure, Method and Meaning: Essays in Honor of H.M. Sheffer* (New York)
- Cooper, R. 1975. *Montague's Semantic Theory and Transformational Grammar*. PhD Dissertation (Amherst: University of Massachusetts, University of Massachusetts Graduate Linguistics Student Association)
- Cooper, R. 1978. *Montague's Theory of Translation and Transformational Syntax*. (London: Duckworth)
- Cooper, R. 1980. 'Montague's syntax'. In Moravcsik, E. & J.R. Wirth, 19-43
- Cooper, R. & T. Parsons 1976. 'Montague Grammar, Generative Semantics and Interpretive Semantics', in Partee, 311-362
- de Cornulier B. 1985. *Effets de sens* (Paris: Les Editions de Minuit)
- Couturat, L. 1903/0961. *Opuscules et fragments inédits de Leibniz. Extraits des manuscrits de la Bibliothèque royale de Hanovre* (Paris: Félix Alcan; Hildesheim: Georg Olms)
- Cresswell, M.J. 1976. 'The Semantics of Degree', in Partee, 261-291
- Cresswell, M.J., 1985. *Structured Meanings* (Cambridge, MA: MIT Press)
- Curry, H. & R. Feys 1958. *Combinatory Logic, vol.1* (Amsterdam: North Holland)
- Daddesio, T.C. 1995. *On Minds and Symbols* (Berlin/New York: Mouton de Gruyter)
- Daelemans, W. 1998. 'Toward an Exemplar-Based Computational Model for Cognitive Grammar', in van der Auwera et al. (eds.), 73-82

- Davis, S. & M. Mithum, M. (ed.) 1979. *Linguistics, Philosophy and Montague Grammar* (Austin, Texas: University of Texas Press)
- Delisle, J. 1980. *L'Analyse du discours comme méthode de traduction* (Ottawa: University of Ottawa Press)
- Delisle, J. 1988. *Translation: an Interpretive Approach* (Ottawa: University of Ottawa Press)
- Derrida, J. 1980/1985. 'Des Tours de Babel', in J. Graham (ed.) *Difference in Translation* (Ithaca: Cornell University Press), 165-207
- Desclés, J.P. 1989. 'Intermediate representations in the cognitive sciences', *Semiotica* 77-1/3: 121-135
- Desclés, J.P. 1990. *Langages applicatifs, langues naturelles et cognition* (Paris: Hermès)
- Desclés, J.P. 1994a. 'Réflexions sur les grammaires cognitives', *Modèles linguistiques*, vol. 29, XV, 1: 69-98
- Desclés, J.P. 1994b. 'Relations casuelles et schèmes sémantico-cognitifs', *Langage*, 13: 113-125
- Desclés, J.P. 1996. 'Cognition, compilation et langage', in Chazal, G., & M-N. Terrasse: 103-145
- Desclés, J.P., Z. Guentcheva, & S. Shaumyan 1986. 'Theoretical Analysis of Reflexivization in the Framework of Applicative Grammar', *Linguisticae Investigationes*, X/1:1-65
- Desclés, J.P., V. Flageul, C. Kekenbosch, J-M. Meunier & J-F. Richard 1998 'Sémantique cognitive de l'action', *Langage*, 132: 28-47
- Dowty, D.R. 1979. *Word Meaning and Montague Grammar* (Dordrecht, Boston, and London: D. Reidel Publishing Company)
- Dowty, D.R., R.E. Wall, & S. Peters 1981. *Introduction to Montague Semantics* (Dordrecht, Boston, and London: D. Reidel Publishing Company)
- Drewerman, E. 1992 *L'essentiel est invisible* (Paris: Les Editions du Cerf)
- Dreyfus, H.L. 1984. *Intelligence artificielle, mythes et limites* (Paris: Flammarion)
- Eco, U. 1976. *A Theory of Semiotics* (Bloomington/London: Indiana University Press)
- Eco, U. 1984. *Semiotics and the Philosophy of Language* (London: MacMillan)
- Eco, U. 1988. *Le signe* (Bruxelles: Editions Labor)

- Eco, U. 1992a. *La production des signes* (Paris: Librairie Générale Française)
- Eco, U. 1992b. *Les limites de l'interprétation* (Paris: Bernard Grasset)
- Eco, U. 1994. *La recherche de la langue parfaite* (Paris: Seuil)
- Edelman, G.M. & G. Tononi, 2000. *A Universe of Consciousness* (New York: Basic Books)
- Enkvist, N.E. 1978. 'Contrastive Text Linguistics and Translation', in Grahs, L., G. Korlén, & B. Malmberg (ed.), Berne: Peter Lang, 169-187
- Even-Zohar, I. 1971. 'Mavo le-teorya shel hartigum ha-sifrut' [Introduction to a Theory of Literary translation]. Unpublished PhD, Tel Aviv: Tel Aviv University.
- Even-Zohar, I. 1978a. 'The Position of Translated Literature Within the Literary Polysystem', in Holmes, Lambert and van der Broeck (ed.)
- Even-Zohar, I. 1978b. Papers in Historical Poetics (Papers on Poetics and Semiotics 8), Tel Aviv: Porter Institute for Poetics and Semiotics.
- Everaert, G. & T. van Steenberghe, T. 1995. 'Meaning Representation and Processing', *La Tribune Internationale des Langues Vivantes*, 17/3: 30-33
- Fawcett, P. 1997. *Translation and Language* (Manchester, UK: St Jerome Publishing)
- Fay-Varnier, C., C. Fouqueré, G. Prigent, & P. Zweigenbaum, P. 1991. 'Modules syntaxiques des systèmes d'analyse du français', *Techniques et Sciences Informatiques*, 10/6: 403-425
- Fillmore, C. 1968. The Case for Case. In E. Bach and R.T. Harnms (eds), *Universals in Linguistic Theory*, 1-88. New-York: Holt, Rhinehart and Winston.
- Fillmore, C. 1981. 'Ideal Readers and Real Readers', *Proceedings of the 32nd Georgetown Round Table on Language and Linguistics* (Washington, D.C.: Georgetown University Press)
- Fitch, F.B. 1974. *Elements of Combinatory Logic* (USA: Yale University Press)
- Fodor, J.A. 1975. *The Language of Thought* (Cambridge, MA: Harvard University Press)
- Fodor, J.A. 1983. *Modularity of Mind* (Cambridge, MA: MIT Press)
- Fodor, J.A. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, MA: Bradford Books, MIT Press)
- Fontanille, J. 1998. *Sémiotique du discours* (Limoges, France: PULIM)

- Frawley, W. 1992. *Linguistic Semantics* (Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers)
- Frege, G. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*, 100, 25-50
- Frege, G. 1953. *The Foundations of Arithmetic* (Oxford: Basil Blackwell)
- Frege, G. 1971. *Ecrits logiques et philosophiques* (Paris: Editions du Seuil)
- Friedman, J. & Warren, D.S. 1978. A Parsing Method for Montague Grammars. *Linguistics and Philosophy* 2, 347-372
- Gaerdenfors, P. 1993. 'The Emergence of Meaning', *Linguistics and Philosophy* , 16: 285-309
- Gagnon, L. 1984. 'Webs of Concern', in Butler F. & Rotert R. (eds.), 66-71
- Galembert, L. 2001. La grandeur du Petit Prince (DEA de Lettres Modernes)
- Gambier, Y. & Tommola J. (eds) 1993. *Translation and Knowledge*. Proceedings of the Fourth Scandinavian Symposium on Translation Theory, Turku, 4.-6.6.1992 (University of Turku - Centre for Translation and Interpreting)
- Gambier, Yves and J. Tommola (eds) (1993) *Translation and Knowledge: Proceedings of the 1992 Scandinavian Symposium on Translation Theory* (Turku: Centre for Translation and Interpreting)
- Gayral, F. 1998. 'Créativité du sens en contexte et hypothèse de compositionnalité', *T.A.L.*, 39/1: 67-98
- Gazdar, G. 1982. 'Phrase Structure Grammar', in Jacobson P. & G. Pullum (eds), *The Nature of Syntactic Representation*, 131-186 (Dordrecht: Reidel)
- Gile, D. 1994. 'Opening up in Interpretation Studies', in Snell-Hornby, M. et al. (eds.), 149-157
- Ginisti, J.P. 1997. *La logique combinatoire* (Paris: PUF)
- Gladkii, A.V. 1970. 'Leçons de linguistique mathématique', *Documents de Linguistique Quantitative*, vol.5&6 (Paris: Dunod)
- Godard-Wendling, B., F. Ildefonse, J-C. Pariente & I. Rosier-Catach 1998. 'Penser le principe de compositionnalité: éléments de réflexion historiques et épistémologiques', *T.A.L.*, 39/1: 9-34

- Godden, K., 1981. *Montague Grammar and Machine Translation between Thai and English*. Ph.D. dissertation, University of Kansas.
- Gordon, P.C., R. Hendrick, K. Ledoux, & C.L. Yang, 1999. 'Processing of Reference and the Structure of Language: an Analysis of Complex Noun Phrases', *Language and Cognitive Processes*, 14(4): 353-379
- Gorlée, D.L. 1994. *Semiotics and the Problem of Translation* (Amsterdam/Atlanta GA: Rodopi)
- Gosselin, L. 1996. *Sémantique de la temporalité en français. Un modèle calculatoire et cognitif de la temporalité en français* (Louvain, Belgique: Duculot)
- Graehs, L., G. Korlén, & B. Malmberg (ed.) 1977. *Theory and Practice of Translation* (Bern: Peter Lang)
- Gross, M. 1994. 'Dictionnaires électroniques et traduction automatique', *Langage*, 116: 48-58
- Guenther, F., & M. Guenther-Reutter, (ed.) 1978. *Meaning and Translation. Philosophical and Linguistic Approaches* (London: Duckworth)
- Guiraud, P. 1983. *La grammaire* (Paris: Seuil)
- Gutt, E.-A. 2000. *Translation and Relevance. Cognition and Context* (Manchester: St Jerome Publishing)
- Haas, W. 1968. 'The Theory of Translation', in Parkinson (ed.), 86-108
- Hacker, P.M.S. 1986. *Wittgenstein: Meaning and Understanding* (Chicago, IL: University of Chicago Press)
- Hagège, C. 1982. *La structure des langues* (Paris: PUF)
- Hagège, C. 1985. *L'homme de paroles. Contribution linguistique aux sciences humaines* (Paris: Fayard)
- Hall-Partee, B. 1978. *Fundamentals of Mathematics for Linguistics* (Dordrecht, Holland: D. Reidel Publishing Company)
- Harris, Z.S. 1954. 'Distributional Structure', *Word*, 10: 146-62
- Harris, Z.S. 1965. 'Transformational Theory', *Language*, 41/4: 363-401
- Harris, Z.S. 1968. *The Mathematics of Language* (Dordrecht: Reidel)
- Hatim, B. & I. Mason 1990. *Discourse and the Translator* (England: Longman)

- Hatim, B. & I. Mason 1997. *The Translator as Communicator* (London and New York: Routledge)
- Hermans, T. 1999. *Translation in Systems* (Manchester: St Jerome Publishing)
- Hervey, S. 1982. *Semiotic Perspectives* (London: George Allen & Unwin)
- Hervey, S., & I. Higgins 1992. *Thinking Translation* (London/New York: Routledge)
- Hewson, L., & Martin 1991. *Redefining Translation, the Variational Approach* (London: Routledge)
- Hilton, J. 1993. 'Some Semiotic Reflections on the Future of Artificial Intelligence', in *Applied Artificial Intelligence*, p.1-14
- Hindley, J.R., & J.P. Seldin, 1986. *Introduction to Combinators and Lambda-Calculus* (Cambridge: Cambridge University Press)
- Hirtle, W. 1994. 'Meaning and Referent: For a Linguistic Approach', *Word*, 45/2: 103-117
- Hjelmslev, L. 1943/1969. *Prolegomena to a Theory of Language* (Madison: University of Wisconsin Press)
- Holmes, J.S. 1988. *Translated! Papers on Literary Translation and Translation Studies* (Amsterdam: Rodopi)
- Holmes, J.S., J. Lambert & R. ven der Broeck (eds). 1978. *Literature and Translation: New Perspectives in Literature Studies* (Leuven: Acco)
- House, J. 1977/1981. *A Model for Translation Quality Assessment* (Tübingen, Germany: Narr)
- House, J. & S. Blum-Kulka (eds.) 1986. *Interlingual and Intercultural Communication: Discourse and Cognition in Translation and Second-Language Acquisition Studies* (Tübingen: Gunter Narr)
- Husson, W. 1994. 'A Wittgensteinian critique of the Encoding-Decoding Model of Communication', *Semiotica*, 98-1/2: 49-70
- Hutchins, J. & H. Somers (1992) *An Introduction to Machine Translation* (London: Academic Press Ltd)
- Israeli, A. Speakers' Attitudes, Goals and Aspectual Choices in WH-Questions, *Le Langage et l'Homme*, XXXIII, 1: 55-77

- Jackendoff, R. 1988. 'Conceptual Semantics', in U. Eco, M. Santambrogio & P. Violi (éds), *Meaning and Mental Representations* (Bloomington / Indianapolis: Indiana University Press), 81-97.
- Jacobson, P. 1999. 'Towards a Variable-Free Semantics', *Linguistics and Philosophy*, 22: 117-184
- Jakobson, R. 1959. 'On linguistic aspects of translation', in Brower (ed.), 232-9
- Jakobson, R. 1963. *Essais de linguistique générale* (Paris: Editions de Minuit)
- Johnson, D. & L.S. Moss, 1994. 'Grammar Formalisms Viewed as Evolving Algebras', *Linguistics and Philosophy*, 17: 537-560
- Johnson, M. 1994. 'Two Ways of Formalizing Grammars', *Linguistics and Philosophy*, 17: 221-248
- Johnson-Laird, P.N. 1984. 'Semantic Primitives or Meaning Postulates: Mental Models or Propositional Representations ?', In Bara, B.G., & G. Guida, 227-245
- Kaiser-Cook, M. 1994. 'Translatorial Expertise - a Cross-Cultural Phenomenon from an Interdisciplinary Perspective', in Snell-Hornby, M. et al. (eds.), 135-139
- Kamp, H. 1978. The adequacy of Translation between Formal and Natural Languages. In Guentner, F. & M. Guentner-Reutter, 275-306
- Katz, J. 1972. *Semantic Theory* (New York: Harper & Row Publishers)
- Katz, J. 1978. 'Affability and Translation', in Guentner, F. & M. Guentner-Reutter *Meaning and Translation*, 191-234
- Katz, J. & J.A. Fodor. 1963. 'The Structure of a Semantic Theory', *Language*, 39: 170-210
- Katz, J.A. & P.M. Postal. 1964. *An integrated Theory of Linguistic Descriptions* (Cambridge, MA: MIT Press).
- Kayser, D. 1985. 'Des machines qui comprennent notre langue', *La Recherche*, 170: 1198-1212
- Kayser, D. 1990. 'Adéquation et inadéquation de la logique au traitement automatique des langues', *Modèles Linguistiques*, XII: 119-136
- Keenan, E. (ed) 1975. *Formal Semantics of Natural Language* (Cambridge: Cambridge University Press)

- Kempson, R. M. 1977. *Semantic Theory* (Cambridge: Cambridge University Press)
- Kirk, R. 1986. *Translation Determined* (Oxford: Clarendon Press)
- Kittredge, R.I. 1987. 'The Significance of Sublanguage for Automatic Translation', in Nirenburg (ed), 59-67
- Kristeva, J. 1981. *Le langage, cet inconnu* (Paris: Editions du Seuil)
- Ladmiral, R. 1994. 'Le traducteur et l'ordinateur', *Langage*, 116: 5-19
- Lakoff, G., Johnson, M. 1980. *Metaphors We Live By* (Chicago/London: The University of Chicago Press)
- Lakoff, G. 1988. 'Cognitive Semantics', in U. Eco, M. Santambrogio & P. Violi (éds), *Meaning and Mental Representations* (Bloomington / Indianapolis: Indiana University Press), 119-154
- Lambrecht, K. 1994. *Information Structure and Sentence Form: Topics, Focus and the Mental Representations of Discourse Referents* (Cambridge: Cambridge University Press)
- Landsbergen, S. P. J. 1976. 'Syntax and Formal Semantics of English in PHLIQA1'. In L. Steeds (ed.), *Advances in natural Language Processing* (Antwerp: University of Antwerp)
- Landsbergen, S. P. J. 1981. 'Adaptation of Montague Grammar to the Requirements of Parsing'. In J.A.G. Groenendijk et al. (eds.), *Formal Methods in the Study of Language*, Part 2. MC Tract 136 (Amsterdam: Mathematical Center)
- Landsbergen, S. P. J. 1985. 'Isomorphic Grammars and their Use in the Rosetta Translation System'. In M. King (ed.) *Machine Translation Today* (Edinburgh: Edinburgh University Press)
- Landsbergen, S. P. J. 2003. Montague Grammar and Machine Translation, in S. Nirenburg, H. Somers & Y Wilks, 239-254.
- Langacker, R.W. 1987. *Foundations of Cognitive Grammar, vol.1: Theoretical Prerequisites* (Stanford, Ca: Stanford University Press)
- Langacker, R.W. 1991a. *Foundations of Cognitive Grammar, vol.2.* (Stanford, Ca: Stanford University Press)
- Langacker, R.W. 1991b. *Concept, Image, and Symbol. The Cognitive Basis of Grammar* (Berlin & New York: Mouton de Gruyter)

- Lappin S. (ed.) 1997. *The Handbook of Contemporary Semantic Theory* (Oxford: Blackwell Publishers)
- Largeault, J. 1993. *La logique* (Paris: PUF)
- Larose, R. 1989. *Théories contemporaines de la traduction* (Sillery, Québec: Presses de l'Université du Québec)
- Larson, M. 1984. *Meaning-based Translation: a Guide to Cross-Language Equivalence* (Lanham, New York and London: University Press of America)
- Lawendowski, B.P. 1978. 'On Semiotic aspects of Translation', in T. Sebeok (ed.) *Sight, Sound, Sense* (Bloomington, In. : Indiana University Press)
- Le Hir Y., 1954 *Fantaisie et mystique dans 'Le Petit Prince' de Saint-Exupéry* (Paris: Nizet)
- Le Ny, J.F. 1995 'Mental lexicon and machine lexicon: Which properties are shared by machine and mental word representations ? Which are not ?' In Saint-Dizier, P. & E. Viegas, 50-67
- Lederer, M. 1994. *La traduction aujourd'hui. Le modèle interprétatif* (Paris: Hachette F.L.E)
- Lehrberger, J. & L. Bourbeau 1988. *Machine Translation* (Amsterdam/Philadelphia: John Benjamins Publishing Co)
- Levy, J. 2000. 'Translation as a Decision Process', in Venuti (ed.), 148-159
- Lobatchev, B. 1998. 'Le côté sémiotique de l'aspect: théorie d'obstacles et de distance', *La Tribune Internationale des Langues Vivantes* , 23: 16-23
- Locke, W.N. & A.D. Booth 1955. *Machine Translation of Languages* (Cambridge, Massachusetts: the MIT Press)
- Lörscher, W. 1991a. *Translation Performance, Translation Process and Translation Strategies: A Psycholinguistic Investigation* (Tuebingen: Gunter Narr)
- Lörscher, W. 1991b. 'Thinking-aloud as a Method for Collecting Data on Translation Processes', in Tirkkonen-Condit (ed.), 67-77
- Lörscher, W. 1993. 'Translation Process Analysis', in Gambier Y. & J. Tammola (eds), 195-211
- Lörscher, W. 1996. 'A Psycholinguistic Analysis of Translation Processes, *Meta* 41: 26-32
- Loffler-Laurian A.-M. 1996. *La traduction automatique* (Villeneuve d'Ascq: Presses Universitaires du Septentrion)

- Lotman, Y.M. [1974] 1977a. 'The Dynamic Model of Semiotic Systems', *Semiotica*, 21: 193-210
- Lotman, Y.M. 1977b. 'Problems in the Typology of Texts', in Lucid, D.P.(ed.), 119-124
- Lotman, Y.M. 1977c. 'The Structure of the Narrative Text', in Lucid, D.P.(ed.), 193-197
- Lotman, Y.M. 1977d. 'Problems in the Typology of Culture', in Lucid, D.P.(ed.), 213-221
- Lotman, Y.M. & Pjatigorskij. 1977. 'Text and Function', in Lucid, D.P.(ed.), 125-135
- Lotman, Y.M. 1990. *Universe of the Mind. A Semiotic Theory of Culture* (London: I.B. Tauris & Co. Ltd)
- Lucid, D.P.(ed.) 1977. *Soviet Semiotics* (USA: The Johns Hopkins University Press)
- Ludlow, P. (ed.) 1997. *Readings in the Philosophy of Language* (Cambridge, the MIT Press)
- Ludskanov, A. 1975. 'A Semiotic Approach to the Theory of Translation', in *Language Sciences*, 35: 5-8
- Lyons, J. 1968. *Theoretical Linguistics* (Cambridge: Cambridge University Press)
- Lyons, J. 1981a. *Language and Linguistics: An Introduction* (Cambridge: Cambridge University Press)
- Lyons, J. 1981b. *Language, Meaning and Context* (Great Britain: Fontana Paperbacks)
- Lyons, J. 1991. *Natural Language and Universal Grammar. Essays in Linguistic Theory, vol.1.* (Cambridge: Cambridge University Press)
- Lytinen, S.L. 1987. 'Integrating Syntax and Semantics', in Nirenburg (ed.), 302-16
- Malone, J.L. 1988. *The Science of Linguistics in the Art of Translation: Some Tools from Linguistics for the Analysis and Practice of Translation* (Albany: State University of New York Press)
- Martin, R. 1983. *Pour une logique du sens* (Paris: PUF)
- Martinet, A. 1989. *Fonction et dynamique des langues* (Paris: A. Colin)
- Matthews, P. 1997. *Concise Oxford Dictionary of Linguistics* (Oxford Oxford University Press)
- Mayorcas, P. (ed.) 1988. *Translating and the Computer 10* (London: ASLIB)
- McCawley, J.D. 1979. 'Helpful Hints to the Ordinary Working Montague Grammarian', in Davis, S. & M. Mithum, 103-123

- McCrorry D. 1990. 'Le Petit Prince: Toward an Interpretation of the Text', *Francophonie*, 1990/2: 39-45
- McDonald, R. 1979. 'The Problem of Machine Translation', *Trends in Linguistics*, 11: 91-145
- Melby, A. K. 1995. *The Possibility of Language* (Amsterdam and Philadelphia: John Benjamins)
- Meunier, J.-G. 1989. 'Artificial Intelligence and Sign Theory', *Semiotica* 77, 1/3: 43-63
- Meunier, J.-G. 1990. 'Semiotic Primitives and Symbolic Machines', in Sebeok T.A. (ed), 403-463.
- Miller, P. & H. Torris (eds.) 1990. *Formalismes syntaxiques pour le traitement automatique du langage naturel* (Paris: Hermès)
- Millican, P.J.R. & A. Clark (eds.) 1996. *Machines and Thought, the Legacy of Alan Turing* (Oxford: Clarendon Press)
- Milner, J.-C. 1982. *Ordres et raisons de langue* (Paris: Editions du Seuil)
- Milner, J.-C. 1995. *Introduction à une science du langage* (Paris: Editions du Seuil)
- Minsky, M. 1975. 'A framework for representing knowledge', in Winston P.H. (ed.), *The Psychology of Computer Vision* (New York: McGraw-Hill)
- Montague, R. 1973. The Proper Treatment of Quantification in Ordinary English. Reprinted in Thomason, ed., 247-270.
- Moravcsik, E. & J.R. Wirth 1980. *Syntax and Semantics* (New York: Academic Press, Inc.)
- Morris, C.W. 1938/1970. *Foundations of the Theory of Signs* (Chicago: Chicago University Press)
- Mott, P. 1996. 'A Grammar-Based Approach to Common Sense Reasoning', in Millican, P.J.R. & A. Clark (eds.), 233-254
- Mounin, G. 1963. *Les problèmes théoriques de la traduction* (Paris: Gallimard)
- Mounin, G. 1997. *La sémantique* (Paris: Editions Seghers)
- Mulder J.W.F. & S.G.J. Hervey 1975. 'Language as a System of Systems', *La Linguistique*, 11/2: 3-22.

- Mulder J.W.F. 1994. 'Written and Spoken Languages as Separate Semiotic Systems', *Semiotica*, 101-1/2: 41-72.
- Nabokov, V. 1941. 'The Art of Translation', *New Republic* 105: 160-2.
- Naro, G. & J. Rey 1998. 'Prépositions et temporalité', *Le Langage et l'Homme*, XXXIII/1: 109-119.
- Neubert A. 1994: 'Competence in Translation: a Complex Skill, how to Study and how to Teach it', in Snell-Hornby, M. et al. (eds.), 411-420.
- Neubert, A. & G.M. Shreve 1992. *Translation as Text* (USA: The Kent State University Press)
- Newmark, P. 1982. *Approaches to Translation* (Oxford: Pergamon Press)
- Newmark, P. 1988. *A Textbook of Translation* (New Jersey, USA: Prentice Hall)
- Newmark, P. 1991. *About Translation* (UK: Multilingual Matters Ltd)
- Nida, E.A. 1964. *Toward a Science of Translating* (Leiden: Brill)
- Nida, E.A. 1975. *Componential Analysis of Meaning* (The Hague: Mouton & Co. N.V. Publishers)
- Nida, E.A. 1982. *Translating Meaning* (San Dimas, California: English Language Institute)
- Nirenburg, S. (ed.) 1987. *Machine Translation. Theoretical and Methodological Issues* (Cambridge: Cambridge University Press)
- Nirenburg, S., H. Somers & Y. Wilks, 2003. *Readings in Machine Translation* (Cambridge, Massachusetts: MIT)
- Noble, H.M. 1988. *Natural Language Processing* (Oxford: Blackwell Scientific Publications)
- Nord, C. 1991. *Text Analysis in Translation* (Amsterdam and Atlanta: Rodopi)
- Nord, C. 1997. *Translating as a Purposeful Activity* (Manchester: St Jerome Publishing)
- Ogden, C. K. & , I. A. Richards. 1923. 'The Meaning of Meaning.' 8th Ed. (New York, Harcourt, Brace & World, Inc.)
- Opolska-Kokoszka, B. 1987. 'Pour un modèle interdisciplinaire de la traduction' *Babel*, 33/1: 18-22
- Paillard, D. & F. Fici Giusti 1998. 'L'inchoation en russe: entre préverbes et auxiliaires', *Le Langage et l'Homme*, XXXIII, n°1: 79-93

- Parkinson, G.H.R. 1968. *The Theory of Meaning* (London: Oxford University Press)
- Parret, H. & J. Bouveresse (ed.) 1981. *Meaning and Understanding* (Berlin: Walter de Gruyter)
- Partee, B. (ed.) 1976. *Montague Grammar* (New York: Academic Press)
- Peirce, C.S. 1931-1966. *Collected Papers of Charles Sanders Peirce*, Charles Hartshorne, Paul Weiss, and Arthur W. Burks (eds.) 8 vols. (Cambridge, MA: Belknap Press, Harvard University Press) (In-text references are to CP, followed by volume and paragraph numbers)
- de Penanros, H. 1998. 'Entre lexique, syntaxe et aspect: le préverbe *pri* en russe contemporain', *Le Langage et l'Homme*, XXXIII/1, 95-107
- Pergnier, M. 1993. *Les fondements sociolinguistiques de la traduction* (Lille: Presses Universitaires de Lille)
- Picken, K. (ed.) [1983] 1989. *The Translator's Handbook* (London: Aslib)
- Ping, K. 1996. 'A Sociosemiotic Approach to Meaning in Translation', *Babel*, 42/2: 74-83
- Piotrovskii, R.G. '1998. Automatic Text Processing: Theory and Practice at the End of the 20th Century', *Automatic Documentation and Mathematical Linguistics*, 32/3: 12-27
- Pitrat, J. 1985. *Textes, ordinateurs et compréhension* (Paris: Editions Eyrolles)
- Prieto, L. 1975. *Essais de linguistique et sémiologie générale* (Paris: Editions de Minuit)
- Propp, V. 1965. *Morphologie du conte* (Paris: Seuil)
- Puurtinen, T.1994. 'Dynamic Style as a Parameter of Acceptability in Translated Children's Books, Snell-Hornby, M. et al. (eds.), 83-90
- Puurtinen, T. 1998. 'Syntax, Readability and Ideology in Children's Literature', *Meta*, LXIII, 1998 (4): 524-533
- Quine, W. V. 1959. 'Meaning and Translation', in Brower, R. (ed.) *On Translation*, 148-172
- Quine, W. V. 1960. *Word and Object* (Cambridge, MA: MIT Press)
- Quine, W.V. 1977. *Le mot et la chose* (Paris: Flammarion)
- Raskin, V. 'Linguistics and Natural Language Processing', in Nirenburg S. (ed.), 42-58
- Reifler, E. 1955. 'The Mechanical Determination of Meaning'. In Locke, W.N., Booth, A.D. *Machine Translation of Languages*, 136-164.

- Reiss, K. 2000. 'Type, kind and Individuality of Text: Decision Making in Translation', in Venuti, L. (ed.), 160-171.
- Renaud, F. 1996. *Sémantique du temps et lambda-calcul* (Paris: PUF)
- Revzin, I. I. 1977. 'Language as a Sign System and the Game of Chess', in Lucid, D.P.(ed.), 87-92
- Rich, E. & K. Knight 1991. *Artificial intelligence* (New York: Mc Graw-Hill Inc)
- Ricoeur, P. 1976. *Interpretation Theory: Discourse and the Surplus of Meaning* (USA: The Texas Christian University Press)
- Riffaterre, M. 1985. 'Transposing Presuppositions on the Semiotics of Literary Translation', *Texte*, 4: 99-109
- Robinson, D. 1991. *The Translator's Turn* (Baltimore and London: The John Hopkins University Press)
- Rumelhart, D.E., J. Mc Clelland, and the PDP Research Group. 1986. *Parallel Distributed Processing, Exploration in the Microstructure of Cognition, vol. 1 and 2.* (Cambridge, MA: The MIT Press.
- Sabah, G. 1990. 'Caramel: un système multi-experts pour le traitement automatique des langues', *Modèles Linguistiques*, 95-118
- Sabah, G. 1997. 'Le sens dans les traitements automatiques des langues', *Traitement Automatique des Langues*, 38/2: 1-133
- Sager, J. 1988. 'Ten Years of Machine Translation Design and Application: from FAHQT to Realism', In Mayorcas (ed.), 3-10
- Sager, J. 1994. *Language Engineering and Translation. Consequences of Automation* (Amsterdam and Philadelphia: John Benjamins)
- Sager, J., 1997. 'Text Types and Translation', in Trosborg (ed.), 25-41.
- Saint-Dizier, P. & E. Viegas 1995. *Computational Lexical Semantics* (Cambridge: Cambridge University Press)
- Sapir, E. (1921) *Language* (New York: Harcourt, Brace 1 World)
- Sarukkai, S. 2001. 'Mathematics, Language and Translation', *Meta*, XLVI, 4: 664-674.

- Saussure, F. 1916/1995. *Cours de linguistique générale* (Paris: Payot)
- Schank, R.C. 1972. 'Conceptual Dependency: a Theory of Natural Language Understanding', *Cognitive Psychology*, 1972/3: 552-631
- Schank, R.C. 1975. *Conceptual Information Processing* (Amsterdam/Oxford: North-Holland Publishing Company)
- Schank, R.C. & R.P. Abelson 1977. *Scripts, Plans, Goals, and Understanding* (Hillsdale NJ: Lawrence Erlbaum)
- Schogt, H. 1985. 'Semantic Theory and Translation Theory', *Textes*, 4: 151-159
- Schogt, H. 1988. *Linguistics, Literary Analysis, and Literary Translation* (Toronto, buffalo and London: University of Toronto Press)
- Schönfinkel, M. 1924. 'Ueber die Bausteine der Mathematischen Logik', *Math. Annalen*, 305-15
- Schulte, R. & J. Biguenet, J. (ed.) 1992. *Theories of Translation, an Anthology of Essays from Dryden to Derrida* (USA: The University of Chicago Press)
- Searle, J. 1980. Minds, Brains, and Programs, *The Behavioral and Brain Sciences*, vol.3: 417-459.
- Searle, J. 1990. L'esprit est-il un programme d'ordinateur ? , *Pour la Science*, 149.
- Sebeok T.A. & J.U. Sebeok (eds) 1990. *Recent Developments in Theory and History, The Semiotic Web* (Berlin: Mouton de Gruyter)
- Sebeok, T.A. (ed.) 1986. *Encyclopedic Dictionary of Semiotics* (Berlin: Mouton de Gruyter)
- Seguinot, C. 1996. 'Some Thoughts about Think-Aloud Protocols', *Target*, 8: 75-95
- Seleskovitch, D. & M. Lederer, M. 1984. *Interpréter pour traduire* (Paris: Publications de la Sorbonne)
- Seleskovitch, D. 1975. *Langage, langues et mémoire* (Paris: Minard)
- Sharples, M., D. Hogg, C. Hutchinson, S. Torrance, & D. Young 1989. *Computers and Thought* (USA: Massachusetts Institute of Technology)
- Shaumyan, S. 1971. *Principles of Structural Linguistics* (The Hague: Mouton)
- Shaumyan, S. 1977. *Applicative Grammar as a Semantic Theory of Natural Language* (Edinburgh: Edinburgh University Press)

- Shaumyan, S. 1987. *A Semiotic Theory of Natural Language* (Bloomington: Indiana university Press)
- Shaumyan, S. 1990. 'Computer Languages as an Instrument of Cognition', *Modèles Linguistiques*, XII-1: 137-167
- Shavit, Z. 1986. *Poetics of Children Literature* (Athens, Georgia: The University of Georgia Press)
- Shuttleworth, M. & M. Cowie, 1997. *Dictionary of Translation Studies* (Manchester, UK: St Jerome)
- Shveitzer, A.D., 1973. *Perevod i lingvistika* (Moscow: Voenizdat)
- Shveitzer, A.D., 1988. *Teoriya Perevoda* (Moscow: Nauka)
- Shwartz, S. C. 1987. *Applied Natural Language Processing* (USA: Petrocelli Books Inc)
- Siegel, M. 1976. 'Capturing the Russian Adjective', in Partee (ed.), 293-309
- Simon H. A., 1996. 'Machine as Mind', in Millican & Clark (ed.), 81-102
- Sless, D. 1986. *In Search of Semiotics* (London: Croom Helm)
- Sloman, A. 1996. 'Beyond Turing Equivalence' , in Millican & Clark (ed.), 179-219
- Snell, B.M. (ed.) 1979. *Translating and the Computer* (Amsterdam, North-Holland Publishing Company)
- Snell-Hornby, M. 1988. *Translation Studies: An Integrated Approach* (Amsterdam and Philadelphia: John Benjamins)
- Snell-Hornby, M, F. Poehhacker, & K. Kaindl. (ed.) 1994. *Translation Studies, an Interdiscipline* (Amsterdam/Philadelphia: John Benjamins)
- Somers, H.L. 1998. 'Machine Translation, Methodology', in Baker (ed.), 1998: 143-149.
- Steedman, M. 2000. *The Syntactic Process* (Cambridge, Ma, and London: The MIT Press)
- Steiner, G. 1974/1992/1998. *After Babel* (Oxford/New York: Oxford University Press)
- Strohner, H. 1990. *Textverstehen. Kognitive und Kommunikative Grundlagen der Sprachverarbeitung* (Opladen: Westdeutscher Verlag)
- Sutherland, R.D. 1970. *Language and Lewis Carroll* (The Hague/Paris: Mouton)

- Tabakowska, E. 1993. *Cognitive Linguistics and Poetics of Translation* (Tubingen: Gunter Narr Verlag)
- Tamba-Mecz, I. 1981. *Le sens figuré* (Paris: PUF)
- Tamba-Mecz, I. 1988. *La sémantique* (Paris: PUF)
- Tarski, A. 1933. 'The concept of Truth in formalized languages', in Tarski, 1983: 152-278
- Tarski, A. 1983. *Logic, Semantics, Metamathematics* (Indianapolis: Hackett)
- Tarski, A. 1995. *Introduction to Logic and to the Methodology of Deductive Science* (New York: Dover Publications Inc.)
- Taylor, K. 1998. *Truth and Meaning. An Introduction to the Philosophy of Language* (Oxford: Blackwell Publishers)
- Taylor, K. A. 1995. 'Meaning Reference and Cognitive Significance', *Mind and Language*, 10-1/2: 129-180
- Thomason, R.H. (ed.) 1974. *Formal Philosophy. Selected Papers of Richard Montague* (New Haven & London, Yale University Press)
- Thurmair, G. 1991. 'Recent Developments in Machine Translation', *Computers and the Humanities*, 25: 115-128
- Tirkkonen-Condit, S. (ed.) *Empirical Research in Translation and Intercultural Studies: Selected Papers of the TRANS-SIF Seminar, Savonlinna 1988* (Tübingen: Gunter Garr)
- Todorov, T. 1977. *Théories du symbole* (Paris: Editions du Seuil)
- Toury, G. 1980. *In Search of a Theory of Translation* (Tel Aviv: The Porter Institute for Poetics and Semiotics)
- Toury, G. 1986. 'Monitoring Discourse Transfer': A Test-Case for a Developmental Model of Translation', in House & Blum-Kulka (ed.), 79-94
- Toury, G. 1995. *Descriptive Translation Studies and Beyond* (Amsterdam/Philadelphia: John Benjamins Publishing Company)
- Toussaint, B. 1978. *Qu'est-ce que la sémiologie ?* (Toulouse: Privat)
- Trosborg A. (ed.) 1997. *Text Typology and Translation* (Amsterdam/Philadelphia: John Benjamins)

- Trost, H. 1991. 'Recognition and Generation of Word Forms for Natural Language Understanding Systems: Integrating Two-Level Morphology and Feature Unification', *Applied Artificial Intelligence*, 5(4): 411-457
- Tucker A.B. 1987. 'Current Strategies in Machine Translation Research and Development', in Nirenburg (ed.), 22-41
- Uspenskij, B. 1977 'Semiotics of Art', in Lucid, D. (ed.), *Soviet Semiotics* (Baltimore & London: Johns Hopkins University Press), 171-173
- van Benthem J. 1986. *Essays in Logical Semantics* (Holland: D.Reidel Publishing Company)
- van der Auwera et al. (eds.) 1998. *English as a Human Language* (Muenchen: LINCOM)
- Vandervecken, D. 1990. *Meaning and Speech Acts, Vol. 1* (Cambridge: Cambridge University Press)
- Vandervecken, D. 1991. *Meaning and Speech Acts, Vol. 2* (Cambridge: Cambridge University Press)
- Venuti, L.(ed.) 2000. *The Translation Studies Reader* (London and New York: Routledge)
- Vermeer, H. J. 1978. 'Ein Rahmen für eine allgemeine Translationstheorie', *Lebende Sprachen*, 9(1): 93-117;
- Vermeer, H. J. 1987 'What does it mean to translate?', *Indian Journal of Applied Linguistics* 13(2): 25-33
- Vermeer, H.J. 1994. 'Translation Today: Old and New Problems', in Snell-Hornby, M. et al. (eds.), 3-15
- Vignaux, G. 1991. *Les sciences cognitives* (Paris: Editions La Découverte)
- Vinay, J.P. & J. Darbelnet, 1958. *Stylistique comparée du français et de l'anglais* (Paris: Didier)
- Vinay, J.P.& J. Darbelnet, 2000. 'A methodology for translation', in Venuti (ed.), 84-93
- Von Neumann, J. 1996. *L'ordinateur et le cerveau* (Paris: Champs Flammarion)
- Wagner, P. 1998. *La machine en logique* (Paris: PUF)
- Wchrli, E. 1997. *L'analyse syntaxique des langues naturelles. Problèmes et méthodes* (Paris: Masson)

- Weizenbaum, I. 1976. *Computer Power and Human Reason: From Judgment to Calculation* (USA: W.H. Freeman)
- Wierzbicka, A. 1988. *The Semantics of Grammar* (Amsterdam/Philadelphia: John Benjamins Publishing Company)
- Wilks, Y. 1979. 'Machine Translation and Artificial Intelligence', in SNELL, B.M. (ed.), 27-43
- Winograd T. 1972. *Understanding Natural Language*, Edinburgh: Academic Press.
- Wittgenstein, L. 1953. *Philosophical Investigations* (Oxford: Blackwell; New York: Barnes and Noble)
- Wittgenstein, L. 1961. *Tractatus logico-philosophicus & Investigations philosophiques* (Paris: Gallimard)
- Wittgenstein, L. 1967. *Zettel*, Anscombe, G. E. M. & G. H. von Wright (eds.) Trans. G. E. M. Anscombe (Oxford: Basil Blackwell) (In-text references are to Zettel followed by paragraph number)
- Woods, W.A. 1970. 'Transition Network Grammars for Natural Language Analysis'. In *Communication of ACM* n° 13 (10).
- Zadrozny, W. 1994. 'From Compositional to Systematic Semantics', *Linguistics and Philosophy*, 17: 329-342
- Zaslowski, D. 1981. On the Semantic Structure of "Meaning" and "Understanding". In Parrett, H., Bouveresse, J. (ed.), 61-75
- Zhu, C. 1999. 'UT Once More: The Sentence as the Key Functional Unit of Translation', *Meta*, 44 (3): 429-447
- Zimmermann, T.E. 1999. 'Meaning Postulates and the Model-Theoretic Approach to Natural Language Semantics', *Language and Philosophy*, vol. 22: 529-561
- Zlateva, P. (ed.) 1993. *Translation as Social Action: Russian and Bulgarian Perspectives* (London: Routledge)