



**UNIVERSITÉ
DE GENÈVE**

Archive ouverte UNIGE

<https://archive-ouverte.unige.ch>

Master

2012

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Langage contrôlé et traduction automatique : une application pratique sur
un texte de type touristique / Post-édition: le cas italien

Chiodi, Céline

How to cite

CHIODI, Céline. Langage contrôlé et traduction automatique : une application pratique sur un texte de type touristique / Post-édition: le cas italien. Master, 2012.

This publication URL: <https://archive-ouverte.unige.ch/unige:21650>

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

Céline Chiodi

Alexandra Dupont

Langage Contrôlé et Traduction Automatique :
une application pratique sur un texte de type
touristique

Post-Édition : le cas italien (Céline Chiodi)

Systran vs. Lucy LT : une évaluation comparative
(Alexandra Dupont)

Mémoire présenté à la Faculté de Traduction et d'Interprétation en vue de
l'obtention du Master en traduction - mention technologies de la traduction.

Directeur : Prof. Pierrette Bouillon

Juré : Marianne Starlander

Université de Genève

Janvier 2012

Remerciements

Nous voudrions remercier les personnes suivantes pour leur aide précieuse tout au long de ce travail de mémoire :

- Pierrette Bouillon, pour ses conseils et ses encouragements tout au long de notre master en TIM et pour nous avoir fourni le cadre de notre mémoire.
- Marianne Starlander, pour sa disponibilité, ainsi que pour ses révisions.
- Johanna Gerlach, pour avoir eu la patience de nous former au système Lucy LT et pour sa participation active.
- Les experts italiens pour leur avis et leurs commentaires précieux.
- Les experts anglais pour leur avis et leurs commentaires tout aussi précieux.
- Nadia Monaco, avec qui nous avons partagé l'espace de travail, les doutes et Systran.
- Céline Pellissier, Héloïse Laureau et Maïlys Fèche pour nous avoir soutenu et supporté tout au long du master, et avec qui nous avons partagés fous-rires et pains au double chocolats réconfortants.
- Nos familles et amis pour nous avoir soutenu jusqu'au bout.

Et enfin...

- Alexandra Dupont voudrait remercier Céline Chiodi pour avoir été à ses côtés dès le premier TP et pour avoir partagé les moments de doutes et les moments d'euphories tout au long du master et pendant le mémoire.
- Céline Chiodi voudrait remercier Alexandra Dupont pour avoir partagé ces quelques pages de désespoir, de fous-rires nerveux et d'émerveillement face au résultat final.

TABLE DES MATIÈRES

<u>I. INTRODUCTION</u>	9
<u>II. TRADUCTION AUTOMATIQUE ET LANGAGES CONTRÔLÉS</u>	13
2.1 INTRODUCTION	13
2.2 LA TRADUCTION AUTOMATIQUE (TA)	13
2.2.1 INTRODUCTION	13
2.2.2 LA TRADUCTION AUTOMATIQUE	15
2.2.3 UN BREF HISTORIQUE	19
2.2.4 AVANTAGES ET INCONVÉNIENTS DE LA TA	21
2.3 LE LANGAGE CONTRÔLÉ	24
2.3.1 INTRODUCTION	24
2.3.2 LE LC	25
2.3.3 HISTORIQUE	26
2.3.4 STRUCTURE D'UN LANGAGE CONTRÔLÉ	30
2.3.5 AVANTAGES ET DÉSAVANTAGES DU LANGAGE CONTRÔLÉ	34
2.4 CONCLUSION	37
<u>III. LES OUTILS ET LE CORPUS</u>	39
3.1 INTRODUCTION	39
3.2 PRÉSENTATION DES OUTILS UTILISÉS	39
3.2.1 SYSTÈMES DE TRADUCTION	39
3.2.1.1 Systran	39
3.2.1.2 Lucy LT	44
3.2.2 OUTIL DE VÉRIFICATION	46
3.2.2.1 Acrolinx IQ	46
3.2.3 OUTIL POUR LE CORPUS	48
3.2.3.1 Wordsmith Tools	48
3.3 LE CORPUS	49
3.3.1 INTRODUCTION	49
3.3.2 DÉFINITION DU CORPUS	49
3.3.3 CARACTÉRISTIQUES DU CORPUS ÉTUDIÉ	51
3.3.3.1 Types de parties	52
3.3.3.2 Observation des phénomènes linguistiques	55
3.4 CONCLUSION	58
<u>IV. MÉTHODOLOGIE</u>	61
4.1 INTRODUCTION	61
4.2 CRÉATION DES DICTIONNAIRES SPÉCIALISÉS	61
4.3 LA PRÉSÉLECTION DES RÈGLES DE CONTRÔLE.	67
4.3.1 RÈGLES GÉNÉRALES	69

4.3.2 RÈGLES SPÉCIFIQUES	79
4.4 CONCLUSION	89

V. PRÉPARATION DE L'ÉVALUATION **91**

5.1 INTRODUCTION	91
5.2 MÉTHODOLOGIE : CORPUS, EXPERTS ET OUTILS	92
5.2.1 LES CORPUS TESTS	92
5.2.2 LES EXPERTS	94
5.2.3 LES OUTILS	95
5.3 DÉFINITION DES CRITÈRES DE QUALITÉ	95
5.3.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	96
5.3.2 LANGAGE CONTRÔLÉ : EFFICACITÉ DES RÈGLES DE CONTRÔLE	97
5.3.3 LANGAGE CONTRÔLÉ : COMPRÉHENSIBILITÉ	97
5.3.4 LISIBILITÉ	97
5.4 DESCRIPTION DES TESTS	97
5.4.1. COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	98
5.4.1.1 Test 1 : Couverture lexicale – Terminologie	98
5.4.1.2 Test 2 : Couverture lexicale – Vocabulaire	99
5.4.2 LANGAGE CONTRÔLÉ	100
5.4.2.1 Test 3 : efficacité des règles de contrôle	100
5.4.3 LANGAGE CONTRÔLÉ	101
5.4.3.1 Test 4 : compréhension	101
5.4.4 LISIBILITÉ	102
5.4.4.1 Test 5 : lisibilité	102
5.5 ACROLINX	102
5.5.1 ACROLINX	103
5.5.1.1 Test 6 : compréhension	103
5.6 CONCLUSION	103

VI. EXÉCUTION DE L'ÉVALUATION **105**

6.1 INTRODUCTION	105
6.2 RÉSULTATS	105
6.2.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	105
6.2.1.1 Test 1 : Couverture lexicale - Terminologie	105
6.2.1.2 Test 2 : Couverture lexicale - Vocabulaire	106
6.2.2 LANGAGE CONTRÔLÉ	107
6.2.2.1 Test 3 : efficacité des règles de contrôle	107
6.2.3 LANGAGE CONTRÔLÉ	108
6.2.3.1 Test 4 : compréhension	108
6.2.4 LISIBILITÉ	110
6.2.4.1 Test 5 : lisibilité	110
6.2.5 ACROLINX	112
6.2.5.1 Test 6 : Compréhension	112
6.3 CONCLUSION	114

<u>VII. POST-ÉDITION : LE CAS ITALIEN</u>	117
7.1 INTRODUCTION	117
7.2 LA POST-ÉDITION (PE)	117
7.2.1 INTRODUCTION	117
7.2.2 QU'EST-CE QUE LA POST-ÉDITION ?	118
7.2.3 TYPES DE POST-ÉDITION	119
7.2.3.1 Post-édition rapide (PER)	120
7.2.3.2 Post-édition conventionnelle (PEC)	120
7.2.3.3 Vers une post-édition personnelle	120
7.3 CRITÈRES D'ÉVALUATION	122
7.3.1 CRITÈRES DE QUALITÉ	122
7.3.1.1 Intelligibilité	123
7.3.1.2 Degré d'amélioration	123
7.4 ÉVALUATION DES ERREURS	123
7.4.1 TABLEAU DES ERREURS	124
7.4.2 CLASSEMENT DES ERREURS	128
7.4.2.1 Application des mesures d'intelligibilité	129
7.4.2.2 Résultats du classement des erreurs pour le critère d'intelligibilité	130
7.4.2.3 Application des mesures de degré d'amélioration	132
7.4.2.4 Résultats du classement des erreurs pour le critère de degré d'amélioration	133
7.5 ÉVALUATION SUR LE CORPUS (J)	138
7.5.1 TABLEAU DES ERREURS POUR CORPUS (J)	138
7.5.2 CLASSEMENTS DES ERREURS SELON LES CRITÈRES DE QUALITÉ POUR LE CORPUS (J) ET RÉSULTATS	141
7.5.2.1 Application et résultats des mesures d'intelligibilité pour le corpus (J)	141
7.5.2.2 Application et résultats des mesures de degré d'amélioration pour le corpus (J)	142
7.6 DISCUSSION	145
7.7 CONCLUSION	147
<u>VIII. SYSTRAN VS. LUCY LT: UNE ÉVALUATION COMPARATIVE DES SORTIES DES SYSTÈMES</u>	149
8.1 INTRODUCTION	149
8.2. MÉTHODOLOGIE : CORPUS, EXPERTS ET OUTILS	149
8.2.1 LES CORPUS	149
8.2.2 LES EXPERTS	151
8.2.3 LES OUTILS	152
8.3 SPÉCIALISATION DES DICTIONNAIRES DE LUCY LT	152
8.4 DÉFINITION DES CRITÈRES DE QUALITÉ	154
8.4.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	155
8.4.2 ÉVALUATION COMPARATIVE : « SENTENCE-BASED RANKING »	155
8.4.3 WER	156
8.5 DESCRIPTION DES TESTS	156
8.5.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	156
8.5.1.1 Test 1 : Couverture lexicale –Vocabulaire	156
8.5.1.2 Test 2 : Couverture lexicale –Terminologie	157

8.5.2. ÉVALUATION COMPARATIVE : « SENTENCE-BASED RANKING »	157
8.5.2.1 Test 3 : ranking des systèmes	157
8.5.3 WER	158
8.5.3.1 Test 4 : WER	158
8.6 RÉSULTATS	158
8.6.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	158
8.6.1.1 Test 1 : Couverture lexicale – Vocabulaire	158
8.6.1.2 Test 2 : Couverture lexicale –Terminologie	159
8.6.2 SENTENCE-BASED RANKING	160
8.6.2.1 Test 3 : ranking des systèmes	160
8.6.3 WER	162
8.6.3.1 Test 4 : WER	162
8.7. CONCLUSION	163
<u>IX. CONCLUSION</u>	<u>165</u>
<u>X. BIBLIOGRAPHIE</u>	<u>169</u>
INDEX DES FIGURES	175
INDEX DES GRAPHIQUES	177
INDEX DES TABLEAUX	179

I. Introduction

La globalisation du flux de production des documents et de leur traitement multilingue a poussé de nombreuses entreprises à se tourner vers les possibilités de traduction automatique afin de relayer l'information de manière rapide et en plusieurs langues. Ainsi, le projet d'entreprise qui se trouve à la base de ce mémoire avait pour but d'évaluer les possibilités de traduction automatique multilingue pour des sites internet consacrés à des lieux touristiques. Nous avons entrepris de travailler sur ce projet à deux, en collaboration, afin de parvenir à un résultat applicable à deux paires de langues, à savoir le français-anglais et le français-italien.

Le thème de notre travail concerne donc la phase de pré-édition dans le processus de traduction automatisée. Notre objectif consiste à évaluer l'efficacité du langage contrôlé sur un texte de type touristique. À cette fin, nous avons travaillé en commun sur les textes sources qui nous ont été proposés par l'administrateur du projet afin de nous rendre compte de leur traductibilité par la machine.

Notre première étape a été de spécialiser les dictionnaires du système afin de rendre le texte cible le plus fidèle possible au texte source. Ensuite, nous avons sélectionné des règles de contrôles inspirées des principes généraux du langage contrôlé et les avons appliquées à notre texte, afin de vérifier l'impact de cette méthode sur la sortie du système. A ce propos, nous avons également tenu à

tester un outil automatique de vérification pour le LC et pour le travail de pré-édition.

Nous avons procédé à l'évaluation des méthodes d'amélioration de la traduction en commun, dans le but de rendre compte des ressemblances et des différences entre paires de langues et de vérifier si le travail de pré-édition avait un impact au-delà des langues traitées. Puis, une fois l'évaluation des méthodes de pré-édition et de spécialisation terminée, nous avons décidé de suivre chacune deux objectifs différents : d'une part, la post-édition de la sortie du système et d'autre part, l'évaluation des performances d'un autre système de traduction.

Nous allons donc brièvement exposer ci-après la structure de notre travail.

Dans le chapitre II, nous présentons le cadre général de notre projet, en introduisant les concepts de traduction automatique et de langage contrôlé et en soulignant les problématiques auxquelles nous devons faire face lors de leur utilisation.

Dans le chapitre III, nous exposons les outils et le corpus que nous avons choisis d'utiliser pour notre travail. Nous décrivons donc les caractéristiques des systèmes et du type de texte sélectionnés.

Dans le chapitre IV, nous introduisons la méthodologie pour la spécialisation des dictionnaires utilisateurs de Systran et la présélection des règles de contrôle que nous avons jugées les plus pertinentes pour l'amélioration de notre type de texte.

La préparation de l'évaluation et la description des différents tests que nous entendons mener afin de rendre compte de l'efficacité de l'application du langage

contrôlé et des dictionnaires personnalisés ainsi que du vérificateur de LC sont détaillés dans le chapitre V.

Nous rapportons dans le chapitre VI les résultats des tests que nous avons réalisés dans le chapitre V, et qui font état de notre travail commun sur le texte source et de son impact sur les sorties.

Enfin, les chapitres VII et VIII ont été élaborés et rédigés de manière individuelle : le premier portant sur la post-édition de la sortie du système pour la paire de langues français-italien et le second sur la comparaison des sorties des systèmes Systran et Lucy LT pour la paire de langues français-anglais. La conclusion apportera les réponses aux questions que nous nous sommes posées tout au long de cette étude.

II. Traduction Automatique et Langages Contrôlés

2.1 Introduction

Ce chapitre présente de manière générale le cadre de notre travail, qui consiste à mesurer l'efficacité du langage contrôlé sur des textes de type touristique en vue d'une traduction automatique multilingue. Nous allons donc y introduire successivement les concepts de traduction automatique et de langage(s) contrôlé(s). Ainsi, la première partie (section 2.2) dresse un bref portrait de la traduction automatique, de son fonctionnement et de ses caractéristiques. La seconde partie (section 2.3) présente le langage contrôlé, son histoire, ses développements et son utilité.

2.2 La Traduction Automatique (TA)

2.2.1 Introduction

Construire un outil qui traduirait de manière entièrement automatisée a été pendant des années un rêve partagé par beaucoup. Comme Somers (2003 : 4) le mentionne, l'idée même de TA précède l'invention de l'ordinateur. En effet, certains philosophes comme Leibniz ou Descartes ont imaginé, au XVII^e siècle déjà, des langues universelles sous la forme de codes numériques. Avec l'apparition des ordinateurs, l'invention d'un programme informatique gérant la traduction de textes d'un langage à un autre n'était donc qu'une question de

temps. Mais qu'est-ce qui est à l'origine de l'invention d'un tel programme ? Hutchins (1986) énumère quatre raisons principales. La première est d'ordre pratique : les professionnels sont souvent amenés à lire des documents et communiquer dans des langues qu'ils ne connaissent parfois pas très bien. En outre, il n'y a pas assez de traducteurs pour le nombre de publications et autres documents à traduire. La TA serait donc une aide précieuse. La deuxième est d'ordre idéaliste : pour promouvoir la paix et la coopération internationale et transmettre notre savoir aux pays en développement, la TA aurait fait tomber les barrières de la langue. La troisième raison est d'ordre stratégique : les armées en conflit se doivent de savoir ce que leur(s) ennemi(s) sait. La quatrième raison est d'ordre scientifique : la recherche sur le langage et la pensée, l'avancée technologique et les limitations de l'ordinateur en font un sujet passionnant pour les informaticiens, les linguistes et les traducteurs.

La traduction automatique a d'abord captivé le public pendant la première moitié du XX^e siècle, puis, n'ayant pas rempli ses promesses, elle a été écartée du devant de la scène. Seuls quelques laboratoires de recherches ont continué à la faire progresser, comme nous le verrons plus en détail dans la sous-section 2.2.3. De nos jours toutefois, la TA bénéficie d'un regain d'intérêt de la part du public : les systèmes de TA sont de plus en plus utilisés, que ce soit par des entreprises ou des particuliers. C'est en cela que notre travail peut être utile : comme nous le verrons au chapitre V, en évaluant les capacités d'un système de TA dans une situation donnée et en essayant de trouver des moyens d'améliorer le résultat de

la traduction automatique, nous essaierons de confirmer l'hypothèse selon laquelle les systèmes de TA peuvent se révéler utiles, voire avantageux.

2.2.2 La Traduction Automatique

La traduction automatique suggère, comme son nom l'indique, l'automatisation du processus de traduction d'un texte en langue source en un texte équivalent en langue cible, à l'aide d'un ou plusieurs programmes informatiques (Arnold, 1994 : 1). Ces programmes sont capables de traduire une grande variété de textes d'une langue à l'autre, mais comme Hutchins & Somers (1997 : 1) le mentionnent très justement, il n'existe pas à l'heure actuelle de « machine à traduire » capable de produire en quelques clics une traduction parfaite. Les programmes de traduction développés jusqu'à présent fournissent des traductions dites « brutes » de textes relativement spécialisés, qui doivent ensuite être révisés par des traducteurs afin d'atteindre une qualité optimale.

Il existe à l'heure actuelle deux grandes catégories de systèmes de TA : linguistiques et statistiques. La première catégorie se fonde sur une compréhension du texte à l'aide de connaissances linguistiques, tandis que, pour la seconde, la compréhension n'est pas requise. En effet, comme nous le verrons plus loin, le système statistique repose sur un corpus de traductions existantes dans lequel le programme va rechercher la traduction la plus probable d'une phrase donnée. Les systèmes linguistiques, quant à eux, se décomposent en deux approches distinctes communément appelées minimaliste et maximaliste. L'approche minimaliste, ou directe, donne une traduction qu'on peut qualifier de

« mot à mot » dans le sens où le système utilise un dictionnaire bilingue pour remplacer le mot source par son équivalent en langue cible (Arnold, 1994 : 63). La désambiguïsation se fait uniquement au niveau du mot, sans véritable compréhension de la phrase. Une des conséquences fâcheuses de ce type d'approche est que le système n'a pas les connaissances suffisantes pour produire des phrases toujours correctes en langue cible. En effet, le système ne peut effectuer qu'une analyse très élémentaire de la langue source, et comme il ne possède pas de véritable connaissance de la langue cible, il n'existe aucune garantie que le texte traduit soit conforme aux usages grammaticaux de la langue cible (Arnold, 1994 : 68). L'approche maximaliste, ou indirecte va plus loin : le système donne un sens aux phrases en effectuant une analyse syntaxique complète à l'aide d'une grammaire propre à chaque langue.

A l'intérieur des systèmes maximalistes indirects, on trouve l'approche par transfert et l'approche par interlangue. Le système par transfert fonctionne en trois étapes : la première analyse les textes pour obtenir des représentations intermédiaires sans tenir compte de l'autre langue ; la seconde étape se charge ensuite de convertir la représentation source dans une représentation cible ; la troisième étape produit le texte en langue cible. La traduction se fonde sur une connaissance approfondie de la langue source et de la langue cible ainsi que de la corrélation des phrases analysées dans les deux langues (Arnold, 1994 : 71). Quant au système par interlangue, sa particularité consiste à convertir les textes en représentations du sens, communes à plus d'une langue. Le processus se fait ici en deux étapes : du texte source à « l'interlangue » et de « l'interlangue » au

texte cible (Hutchins & Somers 1997 : 4). La différence majeure entre ces deux approches concerne la spécificité de la représentation linguistique : alors que le système par transfert se restreint à une représentation spécifique à chaque langue, l'approche interlangue tend à faire disparaître la grammaire comparative au profit du sens, indépendamment des langues. La figure ci-dessous (figure 1) reprend les types de systèmes de TA précédemment cités (Hutchins & Somers, 1997 : 6) :

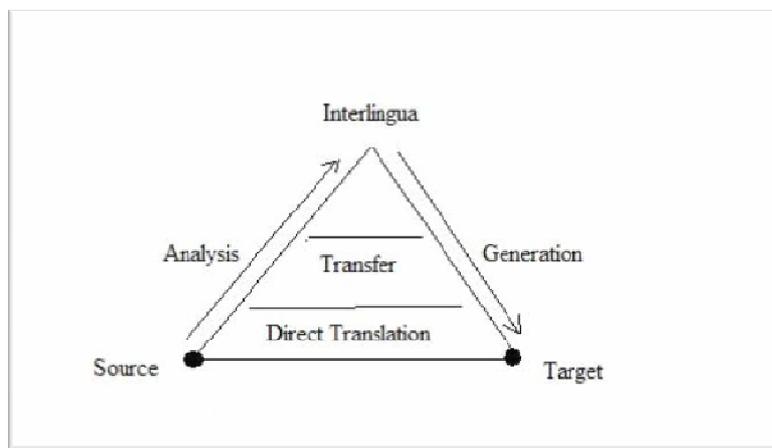


Figure 1. Les systèmes de TA

Depuis quelques années, la recherche s'oriente vers une autre approche de la traduction automatique : les méthodes basées sur les corpus (Hutchins, 2006 : 380). La démarche n'est néanmoins pas nouvelle. En effet, il s'agit d'une des premières techniques utilisées pour la TA dans les années 1950, mais elle a été abandonnée, à partir des années 1960, pour les systèmes basés sur les règles (sous-section 2.2.3). Les méthodes basées sur les corpus englobent la traduction statistique et la traduction guidée par l'exemple. Les systèmes statistiques se

fondent donc sur des probabilités extraites de corpus de traduction existantes. Afin de déterminer la plus haute probabilité qu'une phrase X soit la traduction de la phrase Y, le système statistique utilise deux modèles : le modèle du langage, qui évalue la grammaticalité de la phrase X dans un ou plusieurs corpus monolingues, et le modèle de traduction, qui mesure la fidélité de la traduction sur la base de corpus bilingues. L'approche guidée par l'exemple utilise une base de données constituée d'exemples, c'est-à-dire de paires de fragments de textes source et cible (Somers, 2003 : 136) pour produire une traduction. L'approche par l'exemple considère donc la phrase comme unité de traduction. Le processus consiste donc à rechercher dans une base de données les exemples de référence et de modifier les séquences qui divergent (Carl, 2003 : 4). Toutefois l'approche guidée par l'exemple présente deux problèmes : tout d'abord, un fragment ou une séquence peut se traduire de façons différentes, ce qui peut mener à problème d'ambiguïté ou de synthèse (phrase grammaticalement juste mais non idiomatique) ; ensuite, la base de données doit être suffisamment grande (des millions d'exemples) pour que l'approche fonctionne (Somers, 2003 : 138).

Nous l'avons vu, la TA comprend des techniques différentes. Il n'y a pour l'heure aucun système parfait, mais les différents modèles de TA, linguistiques, statistiques ou même hybrides contribuent chacun à leur manière à l'amélioration de la traduction automatisée. Nous verrons plus en détail en évoquant l'histoire de la TA (sous-section 2.2.3) que les recherches continuent dans le monde entier : le retour de la TA sur le devant de la scène n'y est peut être pas étranger.

2.2.3 Un bref historique

Comme nous l'avons mentionné précédemment, la traduction automatique n'est pas une idée récente, puisque sa création remonte à la première moitié du XX^e siècle. En 1949, Warren Weaver convainc de l'utilité et de la nécessité de la traduction automatique dans un mémorandum destiné à la Rockefeller Foundation (Hutchins & Somers, 1997 : 5). Selon lui, le langage est un code à déchiffrer, ce qui fait de la traduction une application idéale pour les ordinateurs (Arnold, 1994 : 13). La recherche dans ce domaine se développe rapidement en Europe et aux États-Unis, à tel point qu'on assigne à Yehoshua Bar-Hillel un poste de chercheur à plein-temps au MIT. Bar-Hillel organise la première conférence sur la TA, où des idées comme le langage contrôlé sont évoquées pour la première fois. Au départ, il s'agissait de prouver que la traduction entièrement automatisée était possible. En 1954, un groupe de chercheurs de l'université de Georgetown à Washington DC fait la première démonstration publique de TA : un échantillon soigneusement choisi de 49 phrases russes sont traduites en anglais à l'aide de 250 mots et 6 règles grammaticales (Hutchins & Somers, 1997 : 6). Bien que n'ayant aucune valeur scientifique à proprement parler, cette démonstration confirme l'engouement pour la TA.

A la fin des années cinquante cependant, Bar-Hillel soulève les premiers doutes concernant la TA : il déclare dans un rapport datant de 1959 que « la traduction entièrement automatisée de haute qualité était impossible » (Guerra, 2000 : 26).

En 1964, l'ALPAC (Automatic Language Processing Advisory Committee) voit le jour afin d'examiner les résultats des recherches. Deux ans plus tard, le rapport ALPAC met fin aux subsides d'état et en conséquence à la recherche dans ce domaine aux Etats-Unis (Guerra, 2000 : 26). Quelques compagnies privées subsistent malgré tout : Systran sera utilisé par *l'US Air Force* et la Nasa dans les années septante (Bouillon, 2009). L'Europe et le Canada, quant à eux, continuent les recherches en se concentrant sur la construction de systèmes de règles grammaticales pour l'analyse syntaxique des textes (Guerra, 2000 : 26). En 1976, le groupe TAUM (*Traduction Automatique de l'Université de Montréal*) développe le système METEO pour la traduction de prévisions météorologiques. De fait, les années septante marquent un véritable tournant pour la TA : la Communauté Européenne achète les licences *anglais-français* de Systran, redonnant à cette occasion à la TA ses lettres de noblesses (Guerra, 2000 : 28).

Dans les années 80, marquées par les recherches sur l'intelligence artificielle et la linguistique cognitive, les systèmes par interlangue connaissent un regain d'intérêt de la part des chercheurs. Beaucoup de projets voient le jour dans les pays asiatiques, comme le Japon, la Corée, Taïwan ou encore la Malaisie. La fin de la décennie voit arriver les autres méthodes de TA, les « méthodes basées sur des corpus » (Hutchins, 2006 : 380), dont nous avons brièvement parlé dans la sous-section 2.2.2. La première, l'approche statistique, a été au centre des recherches dans les années 90, après le succès de CANDIDE, un modèle statistique développé par IBM, dont le corpus était composé de textes anglais et français tirés des débats parlementaires canadiens. La seconde, l'approche guidée

par l'exemple, est apparue au Japon dès 1981 et, à l'heure actuelle, les recherches continuent activement dans ce pays et aux Etats-Unis. Bien que les dernières décennies se soient concentrées sur les approches basées sur les corpus, les recherches dans le domaine des systèmes linguistiques n'ont pas cessé pour autant : CATALYST pour Caterpillar ou DIPLOMAT pour les opérations militaires, tous deux développés au Carnegie-Mellon University à Pittsburgh, le système ULTRA développé au New Mexico State University ou encore le projet *Universal Networking Language* (UNL) de l'Institute of Advanced Studies of the United Nations University le prouvent (Hutchins, 2006 : 381). Depuis 2007, le Center of Computer and Language Information Engineering (CCLIE) de la Chinese Academy of Sciences et la School of Computer Sci.& Tech du Beijing Institute of Technology (BIT) travaillent sur un projet national de production massive d'information multilingue à l'aide de la TA. (Huang Heyan, 2010).

2.2.4 Avantages et inconvénients de la TA

La réputation de la TA est pour le moins ambivalente. Il est regrettable que la majorité des personnes réduisent la TA à un texte aberrant produit par *Réverso* ou *Google Traduction*. Ils fondent leur opinion sur la partie risible, si ce n'est la partie visible, de l'iceberg de la TA. Cependant, selon Hutchins et Somers (1997 : 1), les progrès de la TA donnent lieu à une avancée significative dans le domaine de la traduction. Les systèmes de TA sont incapables de traduire des textes littéraires mais la majorité des textes à traduire ne sont pas de nature littéraire ou culturelle : ce sont des textes techniques ou scientifiques, des

transactions commerciales ou des brochures publicitaires. Ainsi, l'outil, en évitant au traducteur la tâche pénible de traduire ce type de documents longs et répétitif, se révèle d'une grande utilité.

En outre, la demande pour ce genre de traduction augmente chaque année, ce qui nécessiterait un nombre important de traducteurs et entraînerait ainsi un coût important. L'attrait croissant pour la traduction automatique s'explique tout naturellement par le calcul des gains : gain de temps et d'argent. En effet, il a été calculé qu'un traducteur produisait entre 4 et 6 pages par jour, soit l'équivalent de 2000 mots (Arnold, 1994 : 5). Dans un monde où tout va de plus en plus vite et où les délais sont de plus en plus courts, la TA présente des avantages certains. Le mémoire de Donna Furlani (2009 : 78) en a apporté la preuve : la comparaison des temps de traductions qu'elle a effectué entre la traduction humaine et un système de HAMT (*Human Assisted Machine Translation*) montre que la HAMT possède l'approche la plus rapide dans 70% des traductions. Certes, la sortie du système brute n'est pas toujours de grande qualité, mais elle peut être considérée comme suffisamment bonne pour être post éditée. Si le domaine traité est limité, la sortie du système peut même être fiable la plupart du temps. Le système METEO, par exemple, traduit quotidiennement 45'000 mots environ, et moins de 4 % de la sortie du système a besoin d'être post-éditée (Arnold, 1994 : 11). Dans ces conditions, il n'est pas étonnant que certaines entreprises s'intéressent à la TA et à ses possibilités.

Bien sûr, la TA doit faire face à plusieurs problèmes. Arnold (2003 : 140-141) dénombre quatre problèmes majeurs : (a) la forme est déterminée par le contenu ;

(b) le contenu est déterminé par la forme ; (c) les langues divergent dans leur manière d'exprimer le contenu et (d) il est difficile soit d'exprimer les concepts présents avec la précision nécessaire, soit de trouver les données dont on a besoin pour une approximation statistique.

Pour Hutchins (1992), l'idée qui se cache derrière la TA est qu'une grande partie des langages naturels ainsi qu'une grande partie du processus de traduction peuvent être traités par les programmes informatiques. Les systèmes de TA parviennent dans la plupart des cas à l'aide de l'analyse lexicale, morphologique et syntaxique à produire une sortie du système suffisamment bonne pour être post-éditée. Les difficultés majeures rencontrées par les systèmes de TA concernent l'ambiguïté et le style. En effet, les phrases ayant une construction ambiguë ne sont pas correctement analysées par le système, qui risque de produire un équivalent erroné. Pourquoi les systèmes de TA sont-ils incapables de traduire des phrases ambiguës ? L'explication est simple : il leur manque ce que Hutchins (1992 : 12-13) appelle la « connaissance du monde » (*real-world knowledge*). C'est une certaine connaissance du comportement humain qui permet au traducteur de choisir la bonne traduction pour une phrase donnée.

Hutchins donne un exemple :

- (1) The soldiers shot at the women and some of them fell.
- (2) The soldiers shot at the women and some of them missed.

Seule la connaissance du monde peut permettre au traducteur humain d'identifier les antécédents de « them ». Ainsi :

- (1') Les soldats ont tiré sur les femmes et quelques unes sont tombées.

(2') Les soldats ont tiré sur les femmes et quelques uns ont raté.

Un ordinateur se révèle incapable de comprendre les subtilités de la langue, et par conséquent de saisir tous les niveaux de lecture possible d'une phrase ou d'un texte aussi bien qu'un être humain. C'est pourquoi un système de traduction automatique ne traduira jamais une œuvre littéraire aussi bien qu'un traducteur humain.

En d'autres termes, pour obtenir une traduction acceptable, il faut fournir un texte source que le système peut comprendre sans problème. En effet, si l'on connaît les capacités et les limitations des systèmes de TA, on peut les exploiter au mieux. C'est d'ailleurs le but de notre travail : évaluer l'impact des méthodes d'amélioration de la traduction (dans notre cas, le langage contrôlé) sur un corpus de type touristique soumis à un système de TA. Ainsi, nous allons confronter des versions contrôlées et non contrôlées d'un même texte afin de déterminer s'il est possible ou non d'améliorer la sortie du système de manière significative en intervenant directement au niveau de l'entrée du système à l'aide du langage contrôlé. Dans ce but, nous allons introduire dans la section suivante le concept de langage contrôlé.

2.3 Le Langage Contrôlé

2.3.1 Introduction

De manière générale, la langue est un moyen de communication entre être humains, qui génère beaucoup d'incompréhension. En effet, les ambiguïtés qui la caractérisent sont à la base des difficultés que chacun de nous peut rencontrer. Cet obstacle s'avère d'autant plus important lorsqu'il s'agit d'ordinateurs qui n'ont évidemment pas la même capacité analytique qu'un être humain. Le langage contrôlé (LC) se présente comme la solution trouvée pour réduire les ambiguïtés et augmenter la compréhensibilité.

2.3.2 Le LC

En quelques mots, le langage contrôlé peut être tout simplement décrit comme un langage simplifié, sans pour autant être un langage entièrement artificiel (Ryan, 2009) ou hybride tels que les langages universels comme l'Esperanto (Arnold, 1994 : 147). De manière plus détaillée, il s'agit d'un langage qui est régi par de nombreuses restrictions au niveau du lexique, de la syntaxe et de l'agencement du texte (Ryan, 2009). Ces contraintes se font à l'aide de règles de vocabulaire, de grammaire et de style qui s'ajoutent à celles de la langue standard, sans pour autant les contredire, et qui permettent une meilleure compréhension par la machine (Ryan, 2009). En reprenant un exemple du corpus utilisé pour notre travail (section 3.2), le français ordinaire permet d'écrire de la manière suivante :

L'Espace Paradiski c'est du ski grand format créant ainsi un domaine exceptionnel, avec plus de 420 km de pistes de ski, 2 glaciers skiabiles et 2 sommets culminants à plus de 3000 m d'altitude.

En langage contrôlé, cette phrase serait raccourcie, divisée en plusieurs segments, pour permettre une meilleure compréhension par la machine. Toujours en suivant les règles de LC, on ajouterait des verbes admis par un vocabulaire restreint et on éliminerait les verbes au participe présent. Cela nous amène au texte suivant :

L'Espace Paradiski propose du ski grand format sur un domaine skiable exceptionnel. La station de ski compte plus de 420 km de pistes de ski, 2 glaciers skiabiles et 2 sommets culminants à plus de 3000 m d'altitude.

Cette simplification a pour objectif de réduire toutes les ambiguïtés qui pourraient interférer dans la compréhension par l'être humain et surtout, dans un second temps, par l'ordinateur. En effet, le LC a d'abord été pensé comme une « aide » pour l'être humain, puis il a été utilisé et développé pour les systèmes de traduction, au vu de l'augmentation de volume de la documentation et l'apparition des pratiques d'assurance qualité (Ryan, 2009).

2.3.3 Historique

L'idée de langage contrôlé remonte aux années Trente, lorsqu'un groupe de linguistes et d'intellectuels britanniques décident de concevoir une sorte d'anglais minimal qui serait disponible et accessible au plus grand nombre (Arnold : 1994). Sous la direction de Charles K. Odgen, ce mouvement établit un langage appelé *Basic English*. L'idée principale du *BE* reposait sur une économie lexicale par rapport au vocabulaire général pour toutes sortes d'écrits. Il s'agissait de passer d'une moyenne de 75'000 mots à une centaine. Les mots admis étaient disponibles sous la forme d'une liste, constituée de 850 mots et de

quelques règles concernant l'utilisation des « verbes opérateurs »¹ et de noms/adjectifs remplaçant les verbes dérivés (tableau 1).

Arnold (1994 : 156) en donne un exemple :

(1) The disc controller design was perfected over numerous revisions.

Le Basic English écrirait *was made perfect* :

Make => verbe opérateur

Perfect => un des adjectives dérivés autorisés

Tableau 1. Exemple de Basic English

Le vocabulaire devait être par la suite agrandi pour intégrer les termes techniques et scientifiques propres à chaque domaine (Arnold, 1994 : 156). Le groupe mené par Odgen avait conçu ce langage simplifié comme "un langage international et une base à l'apprentissage de l'anglais standard" (Huisjen, 1998 : 5). A l'époque, l'intérêt porté à ce nouveau langage a été plutôt éphémère et n'a engendré aucune pratique concrète. Ce n'est que plus tard que l'idée refera son apparition, lorsque des grandes entreprises industrielles en feront usage pour la traduction des manuels techniques à un niveau mondial. En écrivant de manière concise, claire et précise, elles pensaient pouvoir être comprises par n'importe quel employé. Ainsi, dans les années 60, est apparu le premier langage contrôlé concrètement utilisé : le *Caterpillar Fundamental English*. Créé par Caterpillar Inc., ce vocabulaire était composé de 900 mots spécialisés et utilisés pour la

¹ Les verbes opérateurs sont les auxiliaires ou les semi-auxiliaires qui permettent de modifier le verbe plein.

documentation internationale de produits. Il a par la suite inspiré bon nombre d'autres LC pour les différentes entreprises et dans des domaines variés tels que la communication maritime, la télécommunication ou l'aéronautique.

Le langage contrôlé le plus connu est peut-être le *AECMA Simplified English* (SE), aujourd'hui appelé le *Simplified Technical English* (O'Brien, 2006 : 17). Il est utilisé pour la documentation concernant l'aéronautique et plus précisément pour les procédures de maintenance des appareils de vols. Le guide SE a été créé dans les années 80 par l'European Association of Aerospace Industries (AECMA) à la demande de l'Association of European Airlines et est devenu un standard mondial de la documentation aéronautique. Ce guide comprend un vocabulaire d'environ 3'100 mots et 57 règles d'écritures divisées en neuf familles (mots, noms, verbes, phrases, procédures, description, mise en garde et précaution). Un guide français (GIFAS Français Rationalisé) a été créé à la suite du SE. L'objectif était de rendre les textes français compréhensibles pour les francophones et facilement traduisibles en SE. Le guide FR contient 50 règles d'écritures divisées en six familles : vocabulaire, verbes, phrases, procédures/modes opératoires, description/fonctionnement, ponctuation et comptage des mots (Barthe, 1998 : 87-102).² Les recherches effectuées à l'AEROSPATIALE sur les deux guides ont montré que leurs règles d'écritures générales pouvaient être réutilisées pour d'autres documents que ceux touchant à l'aéronautique. La réutilisation de règles permet ainsi de gagner du temps et de l'énergie (et des coûts) plutôt que de concevoir de nouvelles règles (Lalaude *et*

² L'article présente toutes les adaptations qui ont été faites pour réaliser un guide FR équivalent au SE mais cohérent au niveau linguistique.

al., 1998 : 104). L'implémentation des règles les plus efficaces est de fait un avantage indéniable « aussi longtemps que leur efficacité peut être confirmée empiriquement » (O'Brien & Roturier, 2007 : 345) Nous avons donc pensé que nous pouvions nous inspirer des règles du SE/GIFAS pour créer notre propre liste de règles de LC, comme nous le verrons dans le chapitre IV.

De nos jours, de nombreuses pistes d'études ont été également explorées sur la possibilité de prendre le LC comme une aide à l'apprentissage de la langue anglaise, à travers le processus de traduction automatique (Shei, 2002 : 89-90). Plus récemment, en 2001, le site Internet Wikipedia (anglais) a également lancé un projet de textes simplifiés pour répondre aux besoins des personnes de langue maternelle étrangère, des enfants et des personnes souffrant de handicaps intellectuels. A l'inverse du *English Wikipedia* standard, le *Simple English Wikipedia* (SimpleEW) utilise un vocabulaire restreint ainsi qu'une grammaire simplifiée pour traiter principalement des articles existants, bien que les articles standards et les articles simplifiés puissent être rédigés indépendamment l'un de l'autre (Woodsend & Lapata, 2011 : 927). Dans ces cas et dans ceux que nous avons vu précédemment, le langage contrôlé est utilisé de manière à faciliter la compréhension du texte pour les êtres humains et non pas dans un but de traduction automatique. Toutefois, dans les années 90, d'autres langages contrôlés sont apparus dans les entreprises pour la traduction automatique de sous-domaines et des sous-langages spécifiques, créant leur propre système de traduction au lieu d'acheter des produits commerciaux (Hutchins, 1999 : 6). Par exemple, la société Cap Volmac en Hollande et la société LANT en Belgique ont

construit des systèmes de traduction spécialisés pour divers clients, en utilisant leurs logiciels pour le contrôle de LC (Hutchins, 1999 : 6). Parmi les langages contrôlés qui sont pensés pour la machine, on peut également citer : IBM's Easy English, Attempto Controlled English, Alcatel's COGRAM, GM's CASL, Sun Mircosystem Controlled English, Avaya's Controlled English, Océ's Controlled English (O'Brien, 2003 :105-106), Controlled Language Optimized for Uniform Translation – Clout (Ryan, 2009) ou encore Carnegie Mellon University' KANTOO system. Nous pouvons donc constater que les LC se sont insérés avec succès depuis cinquante ans dans l'industrie et les technologies plus récentes.

2.3.4 Structure d'un Langage contrôlé

Nous avons vu que le langage contrôlé restreint le langage au niveau du vocabulaire, de la grammaire et du style. De manière plus générale, il s'agit là d'un véritable guide de rédaction que l'auteur et le traducteur se doivent de consulter pour trouver toutes les indications utiles à l'élaboration ou à la traduction d'un texte. Généralement, le LC comprend un dictionnaire dans lequel sont répertoriés tous les mots et termes admis, suivant le principe « un mot, un sens » (Arnold, 1994 : 148). Ainsi par exemple, nous avons reproduit ci-dessous la présentation du vocabulaire du GIFAS (figure 2.):

Mot/Locution (catégorie grammaticale)	Définition/ PROPOSITION(S)	EXEMPLE APPROUVÉ	<i>Non approuvé</i>
ENVOYER (v.)	Faire parvenir, expédier quelque chose.		
<i>effectuer</i> (v.)	FAIRE, ou verbe d'action.	<u>DÉMONTER</u> LE CARTER D'ADMISSION SUR UNE TABLE DE TRAVAIL. <u>FAIRE</u> UN CONTRÔLE DIMENSIONNEL DES PIÈCES QUI ONT DES TRACES D'USURE.	<i>Le démontage du carter d'admission peut être <u>effectué</u> sur une table de travail. <u>Effectuer</u> un contrôle dimensionnel des éléments présentant des traces d'usure.</i>

Figure 2. Exemple de mots autorisés (en majuscule) et non autorisés (en minuscule).

Les règles sont principalement des recommandations sur les constructions de phrases qui doivent être évitées (Arnold, 1994 : 148). Sharon O'Brien (2003) a analysé huit langages contrôlés pour la langue anglaise et en a tiré une classification des règles qui se recoupent dans chaque guide de LC. Les règles se divisent en trois catégories :

- a) Les règles lexicales. Elles concernent l'utilisation du vocabulaire, consistant en une réduction du nombre de mots pouvant être utilisés (Huisjen, 1998 : 2) (figure 3).

<u>Règles concernant les procédures</u>	<u>Règles concernant les rubriques</u> <u>Description/Fonctionnement</u>
<u>Vocabulaire</u>	
<p>RÈGLE 1-1-1 : Utiliser :</p> <ul style="list-style-type: none"> - les mots approuvés du Glossaire, à l'exclusion de tout autre mot, - les mots qui sont des termes techniques, - les mots qui désignent des procédés industriels de fabrication. 	<p>RÈGLE 1-1-2 : Utiliser :</p> <ul style="list-style-type: none"> - les mots approuvés du Glossaire, ou à défaut, tout autre mot nécessaire, - les mots qui sont des termes techniques, - les mots qui désignent des procédés industriels de fabrication.

Figure 3. Règles lexicales du GIFAS.

b) Les règles grammaticales. Elles concernent la syntaxe et la construction des phrases (tableau 2).

Exemple du SE :
SE Rule 3.2 : "Use the approved forms of the verb to make only:
* The infinitive
* The imperative
* The simple present tense
* The simple past tense
* The future tense"

Tableau 2. Exemple de règle grammaticale du SE.

- c) Les règles de style. Elles concernent aussi bien la structure du texte que sa mise en page (figure 4).

<u>Règles concernant les procédures</u>	<u>Règles concernant les rubriques</u> <u>Description/Fonctionnement</u>
<u>Description/Fonctionnement</u>	
Non applicable	RÈGLE 5-3-1 : Ne traiter qu'un seul sujet par paragraphe.

- RÈGLE 6-2-1 :** Vous pouvez utiliser les signes de ponctuation suivants :
- le point .
 - la virgule ,
 - les deux points :
 - le tiret –
 - le trait d'union -
 - les parenthèses ()

Figure 4. Exemples de Règles de style du GIFAS

Comme nous l'avons dit, les guides de rédaction de LC sont pensés pour une utilisation humaine et/ou pour une utilisation par la machine. Certaines règles grammaticales et stylistiques conçues pour l'être humain peuvent se révéler non applicables à la machine, c'est pourquoi il convient toujours d'en vérifier l'impact sur la machine (Roturier, 2004 : 5). D'autre part, les règles peuvent se contredire entre elles. Par exemple, Roturier a pu le vérifier dans les tests qu'il a mené pour Symantec : ainsi la règle de rédaction du Symantec Guideline permettant des constructions de phrases enchâssées se contredit avec le principe du LC « une phrase-un sujet ». Afin d'éviter cela, il faut donc que les programmeurs et les traducteurs se consultent pour établir une sélection satisfaisante (tests Symantec in Roturier, 2004 : 11).

2.3.5 Avantages et désavantages du Langage Contrôlé

L'intérêt majeur de l'utilisation d'un langage contrôlé réside bien évidemment dans l'impact que celui-ci a sur la lisibilité et la compréhensibilité du texte traduit, réduisant ainsi le temps et le coût d'une post-édition (Roturier, 2004 : 2). Sur le plan de la qualité de la traduction automatique, les diverses études de ces vingt dernières années ont en effet démontré que la qualité de la sortie du système peut être sensiblement améliorée par la simplification du vocabulaire général et de la structure des phrases et du texte (Arnold, 1994 : 27 ; Roturier, 2004 : 2). De plus, le développement du LC répond parfaitement aux besoins d'une industrie qui doit gérer des milliers de documents et produire une

information de plus en plus rapide. Ryan (2009) explique dans son article sur la documentation technique industrielle que les langages contrôlés sont devenus un outil essentiel au travail de traducteur, lequel doit faire face à une nombreuse documentation spécialisée et variée et répondant à des exigences de conformité. Le LC a facilité la simplification et la standardisation des documents en établissant des normes textuelles avec pour objectif : la clarté, la bi-univocité terminologique, l'homogénéité, etc. De plus, la qualité du travail rédactionnel contrôlé, effectué sur le texte source, se répercute immédiatement sur « la qualité ergonomique du texte cible » (Ryan, 2009). Ainsi, le but ultime des industries – l'économie – peut être atteint par le gain effectif en termes de coût, d'effort et de temps pour la traduction comme pour la post-édition, obtenant un produit rapidement publiable (Roturier, 2004 : 2).

Toutefois, comme nous l'avons énoncé plus haut, le langage contrôlé consiste en une série de règles qui peuvent faire surgir des difficultés au moment de l'application effective. De fait, si les règles du LC se veulent facilement insérables dans les règles de la langue naturelle, elles restent parfois difficiles voire impossibles à appliquer. Selon Ryan (2009), une communication neutre et factuelle, comme le sous-entend l'emploi des règles de grammaire élémentaire du LC, se heurte aux règles constitutives de la hiérarchisation des informations ; l'utilisation réduite des subordonnées afin de garder la longueur des phrases dans les limites autorisées en est exemple (Ryan, 2009). L'auteur se met en effet à la place de l'utilisateur final, le lecteur, et relève les difficultés suivantes : la déshumanisation du discours, l'inconfort de lecture dû à la répétition lexicale ou

un manque de lien naturel entre les informations (non utilisation des pronoms par exemple). Tout ceci pouvant conduire à un désintérêt du lecteur. Au niveau de la traduction, Ryan (2009) comme Roturier (2004) pointent du doigt la difficulté de se conformer totalement au LC. En effet, les règles de LC dépendent largement des performances du système de traduction automatique utilisé; ainsi, il existe des règles généralisables – notamment concernant la construction grammaticale ou la mise en page textuelle – (Lalaude *et al.*, 1998 : 108) mais la création de règles spécifiques, répondant aux exigences du texte source, s'impose la plupart du temps (Roturier, 2004 : 4). La soumission aux structures formelles du LC peut provoquer chez le lecteur un sentiment de gêne, surtout dues aux concessions faites à l'usage normal de la langue (Ryan, 2009). Les deux auteurs en donne un exemple en reprenant les discussions et pressions des rédacteurs en faveur d'un assouplissement ou d'un compromis. Qui plus est, pour ces derniers, la pratique de LC est considérée comme ennuyeuse. En effet, il est clair que le rythme de travail et donc de productivité se trouve diminué par le devoir de se conformer à un guide d'écriture. D'un point de vue stylistique, le LC peut parfois requérir un travail important de réécriture et amener à un risque d'altération du sens par souci de conformité. De même, les différents desseins du LC – lisibilité et traductibilité par exemple - sont parfois incompatibles (Ryan, 2009).

En parallèle à la création des langages contrôlés, des outils spécifiques ont été développés spécifiquement pour contrôler automatiquement le texte. Ces programmes vérifient si les mots sont autorisés et si les règles d'écriture sont suivies (Huisjen, 1998 : 7). Ces outils aident les êtres humains dans la tâche de

contrôle. Certains, en plus de vérifier, peuvent offrir une aide à la correction. Pour ce qui est de la vérification des règles, il faut souligner que le degré de vérification dépend également du niveau d'analyse atteignable par l'outil (Huisjen, 1998). L'outil Acrolinx, que nous allons utiliser dans le cadre de notre étude (sous-section 3.2.2), aide à la vérification du langage contrôlé au niveau du lexique, de la syntaxe et du style. Par l'implémentation des règles dans le système, il fournit une analyse détaillée des différentes violations commises dans les phrases et qui pourraient empêcher une bonne sortie du système de la machine. Pour corriger les erreurs, le système affiche une boîte de dialogue où les corrections peuvent être effectuées manuellement. Toutefois, l'outil ne fait que rappeler la règle violée et ne propose pas de solutions.

2.4 Conclusion

Après avoir exposé le cadre général de notre travail, nous allons dans le chapitre suivant présenter les outils de traduction et de vérification que nous avons brièvement mentionné et que nous entendons utiliser sur notre type de texte, introduisant également les caractéristiques textuelles du corpus.

III. Les outils et le corpus

3.1 Introduction

Le présent chapitre décrit les outils utilisés ainsi que le texte que nous avons choisi pour évaluer l'efficacité du LC pour la traduction automatique. Tout d'abord, nous allons brièvement introduire les deux outils de traduction, l'outil de vérification et l'outil de corpus que nous avons employés dans le cadre de notre travail (section 3.2). Enfin, dans un second temps, nous allons nous pencher sur le type de texte que nous avons choisi de soumettre aux diverses applications, décrire ses caractéristiques et présenter les difficultés potentielles de ce corpus en ce qui concerne la traduction automatique (section 3.3).

3.2 Présentation des outils utilisés

3.2.1 Systèmes de traduction

3.2.1.1 Systran

La création du système de TA Systran a débuté à la fin des années 60 aux États-Unis lorsque son développeur, Peter Toma, quitte le projet de MT de l'université de Georgetown afin de lancer un système commercialement viable. En 1973, le premier système opérationnel est utilisé par l'US Air Force pour répondre aux enjeux politiques et techniques sur lesquels s'affrontent les Américains et les Russes. La première paire de langues est donc pour la

traduction du russe à l'anglais. Deux ans plus tard, la Commission des Communautés Européennes achète une version pour la paire de langue EN-FR qui sera développée et étendue à d'autres paires (FR-EN, EN-IT). Aujourd'hui Systran compte 22 paires de langues, continuellement développées.

Systran peut être essentiellement vu comme un système direct évolué (Hutchins, 1986), dit parfois de transfert (Senellart *et al.*, 2001), qui se caractérise par trois composantes essentielles :

- *son système de base dit « modulaire »* (Whitelock & Kilby, 1995 : 40-41). En effet, Systran possède un programme commun qui est totalement indépendant des langues qu'il traite. Il est composé de programmes utilitaires et de contrôle (Hutchins, 1986) qui sont responsables de gérer les dictionnaires (mise à jour, création des traductions et vérification) et de contrôler la traduction, etc. (Loffler-Laurian, 1996 : 53).
- *les programmes de traduction*. Contrairement au système de base, ces derniers dépendent de la langue traitée et sont donc sujets aux variations selon la paire de langues (Whitelock & Kilby, 1995 : 41). Systran est souvent classifié comme système de transfert car le système traite séparément l'analyse de la langue source et la synthèse de la langue cible (Whitelock & Kilby, 1995 : 41) mais le processus de traduction principal est mené par les dictionnaires LS-LC

(Hutchins, 1986). Les différentes étapes du processus sont présentées schématiquement dans la figure 5 et peuvent se résumer ainsi:

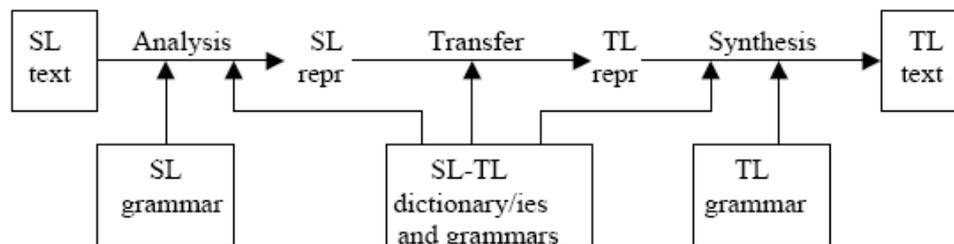


Figure 5. Processus de traduction d'un système direct évolué – Systran (Hutchins : 1986)

Systran effectue un processus linéaire où différents modules opèrent l'un après l'autre. Tout d'abord, dans la phase d'analyse, chaque mot composant une phrase de la langue source (LS/SL) est analysé avec toutes les informations grammaticales (catégories, nombre et relation) et les équivalents de traduction associés aux mots existant dans le dictionnaire LS (Whitelock & Kilby, 1995 : 42). Puis, la phrase entière est analysée par chaque module permettant ainsi la gestion des différents phénomènes linguistiques tels que « [la] résolution des homographes grammaticaux, [l'] analyse des limites d'énoncés [et l'] établissement des divers niveaux de dépendance syntaxique » (Loffler-Laurian, 1996 : 54). Ensuite, la phase de transfert s'effectue selon les spécificités des paires de langues. De manière générale, le programme identifie les structures grammaticales et/ou les termes lexicaux (mot/expressions) de LS puis attribue le sens en langue cible (LC/TL) (Whitelock & Kilby, 1995 : 60). Enfin, la phase de

synthèse se fait en deux temps : premièrement, chaque élément de la phrase est traduit selon les informations récoltées par les divers modules ; deuxièmement, l'ordre des mots de la LC est revue et corrigée selon les règles syntaxiques de cette dernière (Whitelock & Kilby, 1995 : 63).

- *les dictionnaires.* Pour chaque langue source, deux types de dictionnaires sont utilisés. Le premier dictionnaire traite les entrées à un mot appelées « stems » ; son objectif premier est de fournir les informations lexicales pour l'analyse linguistique. Le second dictionnaire traite les entrées consistant en expressions à plusieurs mots appelées « idiomes » ; l'ensemble des mots constitue un sens propre que le système doit reconnaître, comme pour l'expression anglaise «on the one hand » ou pour les mots polysémiques par exemple « to turn » et « to turn off ». En fait, dans les phases d'analyse linguistique et de transfert, ce dictionnaire complète les informations générales attribuées à chaque mot par des informations spécifiques à l'expression entrée (Gerber & Yang, 1997 : 212).

Pendant plus de 30 ans, Systran a été développé progressivement, intégrant les dernières technologies de MT, avec comme souci constant le maintien des ressources linguistiques, leur réutilisation, la maîtrise du processus de traduction et l'ouverture du système à un marché de plus en plus exigeant (Senellart *et al.*, 2001). Le système a été pensé pour les besoins d'une information rapide

puisqu'il fournit des traductions « directement » utilisables pour une grande variété des textes. Pour Systran, la couverture large du vocabulaire, avec des caractéristiques sémantiques et syntaxiques, sans limitation dans le type ou domaine de document est une haute priorité (Gerber & Yang, 1997 : 211). Suivant cette perspective, Systran a développé la technologie du codage intuitif dans le Dictionary Manager, facilitant l'enrichissement des dictionnaires et permettant ainsi une amélioration de la qualité de la traduction automatique. Cette technologie permet la conversion simple et automatique des entrées du dictionnaire utilisateur en un dictionnaire de traduction automatique fonctionnel, sans aucune intervention manuelle (Senellart *et al.*, 2003). L'utilisateur a donc en main un moyen de personnaliser le système. Il peut, comme dans l'exemple ci-dessous, ajouter des équivalents techniques pour un mot général ou définir le sens spécifique d'un mot susceptible d'être traduit de plusieurs façons ou encore spécifier des règles contextuelles, etc. (Senellart *et al.*, 2003).

Exemple :

EN : to save => sauver ; (context : money) = économiser

Le codage intuitif permet une gestion facile de l'information spécialisée par l'utilisateur et une personnalisation du système. Toutefois, la richesse et précision des dictionnaires par le codage implique également un processus de construction de dictionnaire manuel relativement lent (Gerber & Yang, 1997 : 212).

Cette personnalisation facilitée des dictionnaires sera testée dans la suite de notre travail, puisque nous créerons notre dictionnaire utilisateur pour les paires de langue français-anglais et français-italien (section 4.2). Par la suite, nous pourrons observer l'impact de l'apport des dictionnaires spécialisés sur la qualité de la traduction automatique (chapitres V et VI).

3.2.1.2 Lucy LT

Le système de TA Lucy LT est entièrement fondé à partir de METAL (*Mechanical Translation and Analysis of Language*), un système développé grâce à la collaboration de Siemens et de l'Université du Texas entre 1978 et 1988. En 1989, METAL est commercialisé pour la paire de langues EN-ALL (Schneider, 1991 : 41). A partir de cette date, le développement du système est transféré à Sietec et GMS (*Gesellschaft für multilinguale System GmbH*) jusqu'à la faillite de l'acheteur de GMS. Deux systèmes dérivés de METAL apparaissent sur le marché : T1 de Langenscheidt et Compendium, de SailLabs. Compendium est racheté en 2009 et renommé Lucy LT (Bouillon, 2009).

Il s'agit d'un système indirect de transfert, ce qui signifie, comme nous l'avons expliqué dans la sous-section 2.2.2, que la représentation se fait sous forme d'arbre syntaxique annoté avec les informations grammaticales et sémantiques. La traduction se fait donc d'un « arbre » à l'autre, et non pas d'une phrase à l'autre comme dans les systèmes directs. Les arbres étant par défaut identiques en langue source et en langue cible, les différences doivent être signalées à travers l'écriture de règles de transfert. Lucy opère avec deux lexiques monolingues et

un lexique de transfert pour chaque paire de langue. Les lexiques monolingues contiennent des informations morphologiques, syntaxiques et sémantiques nécessaires à l'analyse et à la génération de la langue. Le lexique de transfert fait le lien entre la langue source et la langue cible (Schneider, 1991 : 42).

Le lexique de transfert peut contenir un ou plusieurs équivalents en langue cible pour chaque terme en langue source, selon le nombre de traductions possibles. Puis, pour que le système puisse produire une sortie du système correcte, il est possible d'associer des tests et des actions à ces équivalents. Les tests permettent de définir les conditions dans lesquelles une entrée sera utilisée. Il existe deux types de tests : sur les attributs du terme, et sur le contexte. Les actions, quant à elles, influent sur l'utilisation du terme dans la langue cible, comme son nombre, sa position, sa forme verbale, ses compléments, etc. (Guide *Lexshop*). On compte trois types d'opérations : la première permet de modifier la valeur d'un attribut, la seconde de changer la traduction par défaut selon le contexte et la troisième d'intervenir sur la structure syntaxique (Bouillon, 2009). A l'aide des tests et des actions, l'utilisateur a la possibilité d'affiner le système de manière très rigoureuse afin que la sortie du système soit la plus juste possible.

Lucy LT offre donc à l'utilisateur la capacité de modifier le lexique de manière plus approfondie que ne le ferait Systran. Cela représente à la fois un avantage et un inconvénient, car l'utilisateur de Lucy LT doit de préférence avoir des connaissances linguistiques et une intuitivité informatique qui ne sont pas obligatoirement exigées de l'utilisateur de base de Systran. En outre, les deux outils utilisent des approches différentes pour traduire les textes, c'est pourquoi

nous avons pensé qu'il serait intéressant de comparer les deux systèmes à travers une application pratique, c'est-à-dire en utilisant une partie de notre corpus afin de voir quel système obtient les meilleurs résultats. La comparaison entre Lucy LT et Systran fera l'objet du chapitre VIII.

3.2.2 Outil de vérification

3.2.2.1 Acrolinx IQ

Acrolinx IQ a été développé au Centre de Recherches Allemand pour l'Intelligence Artificielle, et commercialisé en 2002 sous le nom d'Acrocheck. Il a été rebaptisé Acrolinx IQ en 2008 (Lehmann, 2007). Il s'agit d'un vérificateur de style, de grammaire et d'orthographe qui assure la cohérence entre les textes d'un même domaine ayant plusieurs auteurs. Il est utilisé par de nombreuses grandes compagnies, comme Siemens ou encore IBM. L'outil offre une architecture de type client-serveur ; les fonctionnalités d'Acrolinx sont disponibles sous la forme de plug-ins compatibles avec la plupart des systèmes auteurs (www.acrolinx.com). Nous pouvons voir dans la figure 6 ci-dessous le fonctionnement de l'architecture d'Acrolinx :

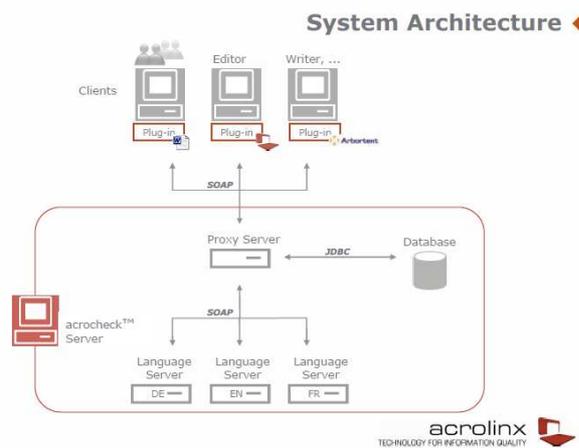


Figure 6. L'architecture d'Acrolinx IQ

Dans notre cas, Acrolinx IQ vient sous la forme d'un plug-in installé sur Word qui permet d'analyser le texte au niveau de la grammaire, de l'orthographe et du style en se basant sur des règles de langage contrôlé. Les phrases qui ne correspondent pas aux règles sont signalées dans un rapport et peuvent être corrigées par l'utilisateur. La figure 7 ci-dessous nous montre les règles les plus fréquemment enfreintes :

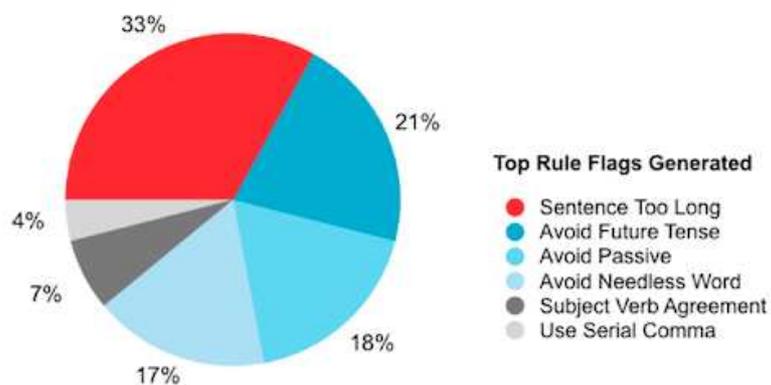


Figure 7. Les erreurs les plus fréquentes relevées par Acrolinx.

Nous n'avons qu'un accès restreint au serveur, et de fait nous ne pouvons ni modifier les règles ni en implémenter de nouvelles. Nous nous appuyerons donc sur les règles existantes pour évaluer l'efficacité et l'utilité des règles de contrôle d'Acrolinx IQ pour la traduction de notre corpus.

3.2.3 Outil pour le corpus

3.2.3.1 Wordsmith Tools

Créé en 1996 par Mike Scott, à l'université de Liverpool, Wordsmith Tool a été pensé pour les étudiants et les chercheurs afin qu'ils puissent construire et maintenir un corpus sur leur PC, facilement accessible à l'aide d'un bon logiciel (Scott, 2001 : 48). Mike Scott (2001) décrit son outil comme une suite intégrée de programmes permettant d'étudier le « comportement » des mots dans le texte. Les trois logiciels principaux sont :

- le concordancier *Concord*. Il permet la localisation de toutes les références d'un mot ou d'une phrase dans le corpus en contexte.
- la *Wordlist*. Elle permet de créer des listes de mots et de les classer par ordre alphabétique et selon leur fréquence dans le corpus. Cet outil est utile pour identifier les caractéristiques d'un texte ou d'un genre (Bowker & Pearson, 2002)
- Le *Keyword*. Il utilise les listes de mots et permet leur comparaison notamment au niveau de la fréquence.

Nous avons uniquement utilisé l’outil de création de liste de mots pour repérer facilement les termes spécialisés du texte source afin de les extraire puis si nécessaire, insérer leur(s) équivalent(s) dans le dictionnaire utilisateur de Systran. De plus, cet outil s’est révélé d’une grande aide pour l’évaluation sur la personnalisation du dictionnaire, comme nous le verrons dans les chapitres V et VI.

3.3 Le Corpus

3.3.1 Introduction

La première grande étape de notre travail consiste à choisir un corpus sur lequel nous allons tester l’impact du langage contrôlé et les dictionnaires utilisateurs. Le corpus en lui-même sera également évalué pour savoir s’il se prête ou non à un processus de traduction automatisé. A la base de ce travail, un projet de traduction pour des sites web touristiques devait nous fournir le contexte à analyser. Nous avons donc reçu plusieurs fichiers de textes extraits de plusieurs sites ayant comme trait commun les stations de ski. L’analyse devait porter sur la faisabilité d’une traduction automatique de textes touristiques, en utilisant le système Systran, les dictionnaires personnalisés par nos soins et le langage contrôlé (en s’aidant également d’Acrolinx pour en vérifier son utilité).

3.3.2 Définition du corpus

Pour établir notre corpus, nous avons dû sélectionner des textes parmi ceux que l’administrateur du projet nous avait envoyés et qui étaient divisés

selon les thèmes suivants : bien-être, hébergement, présentation générale des stations, présentation des remontées mécaniques. Notre choix s'est porté sur ce dernier thème puisqu'il contient une série de textes susceptibles d'être améliorés par le LC. Le document, en format Excel, contient 215 cellules ; chaque cellule étant un texte indépendant qui correspond à une description de stations de ski. Les cellules sont toutes composées de manière plus ou moins identique. Au vu de l'importance du document, nous avons décidé de ne prendre que la moitié des cellules, totalisant ainsi un peu plus de 10'200 mots. Le contenu de notre corpus se référant à un domaine de connaissance spécifique, il peut donc être regardé comme un sous-langage. On peut définir ce dernier par un certain degré de fermeture, dans lequel une série de mots et des constructions de grammaire finies sont utilisés (Pearson, 1998 : 28); par exemple, il est possible de développer une simple liste de règles de phrases-types pour ce langage (McEnery & Wilson, 2001 : 166). Nous tenterons plus loin d'éliminer complètement les dernières traces de créativité par l'application du LC (notamment au niveau lexical par le principe de « un mot, un sens »). A ce stade de notre étude, en partant de l'exemple de McEnery & Wilson (2001), nous avons décidé de décrire ces segments-types, de dégager leur appartenance à un type de discours et en conclusion, d'arriver à en extraire une liste de règles utilisables, qui serait à la base de notre sélection pour le langage contrôlé.

3.3.3 Caractéristiques du corpus étudié

En vue de faire la description des phrases-types qui composent notre corpus, nous avons tout d'abord dû identifier le type de discours auquel ces phrases appartiennent. Nous savons que le discours est « le produit de multiples pratiques discursives » (Desmet, 2007 : 8) et possède des critères linguistiques observables. Les différents domaines sont ainsi sources de types discursifs différents. Selon Lalaude *et al.* (1998), une approche linguistique combinant des dimensions macro-linguistiques (agencement du texte) et micro-linguistiques (structures des phrases) permettrait d'effectuer une typologie du corpus et d'extraire des règles d'écriture spécifiques nécessaire pour le travail de LC. Nous avons donc adapté leur méthodologie en établissant une correspondance entre le type de textes et les critères linguistiques (structures syntaxiques récurrentes, lexicale, etc.) qui s'y trouvent et qui contribuent à leur donner ce caractère particulier immédiatement identifiable, (Druetta³ : 2).

A première vue, la structure d'un document est facilement identifiable par les parties introduites par un titre, accompagnées parfois de signes formels tels que des tableaux ou des listes et qui présentent des caractéristiques linguistiques spécifiques. Appelées « textual modules » par Lalaude *et al.* (1998), ces types de parties ont des spécificités qui nous permettent de les détecter puis d'observer et de décrire comment elles se composent et comment elles fonctionnent. Toutefois, ces dernières n'étaient pas immédiatement détectables dans notre cas. En fait, le

³ Voir : http://www.francaisunivers.unito.it/documents/types_druetta.pdf

texte que nous avons reçu en format Excel n'avait plus ni structure ni formatage puisqu'une cellule représentait plusieurs paragraphes. Après avoir extrait le texte en format RTF, notre première démarche a été de retrouver la structure du texte. Ainsi, les *textual modules* nous sont apparus de manière plus claire.

3.3.3.1 Types de parties

L'identification des parties formelles et, par la même occasion, des objectifs de communication d'un texte découlent d'une approche macro-linguistique. Les signes formels du texte de notre corpus (tels que les paragraphes, les titres, les sous-titres, les listes ou les tables) nous permettent de constater la présence de trois grandes parties, chacune précédées d'un titre et dont deux sont composées de listes (tableau 3). Par ailleurs, nous avons constaté que chaque partie présente une même structure avec une à plusieurs composantes.

1	Ouverture du domaine : - Du 12/12/2010 au 03/04/2011 (soumis aux aléas climatiques et à la décision unilatérale de la station).
2	La situation géographique privilégiée de la station vous fera profiter d'un soleil méditerranéen et d'un excellent enneigement. Le domaine skiable, d'Isola 2000 est partagé en deux vallées, avec d'un côté le versant de la Lombarde et de l'autre côté le versant du Méné. Situé à 2300 m, le snowpark d'Isola 2000 vous fait découvrir de nouvelles sensations de glisse. Cet espace ludique et technique, vous propose des modules pour tous les niveaux : slopestyle, big air, handrails, zone débutants...

3	<p>Le forfait comprend :</p> <ul style="list-style-type: none"> - L'accès au domaine skiable d'Isola 2000. <p>A noter :</p> <ul style="list-style-type: none"> - Gratuité pour les enfants de - de 5 ans et les adultes à partir de 71 ans. - Prévoir un document officiel justifiant l'âge des participants. - Pour les séjours du samedi 18/12 au dimanche 26/12/2010 (séjours de 8 nuits), votre forfait remontées mécaniques est valable à partir du lundi 20/12/2010 au matin. - Pour les séjours du dimanche 02/01 au samedi 08/01/2011 ou du dimanche 09/01 au samedi 15/01/2011 (séjours de 6 nuits), votre forfait remontées mécaniques est valable le jour même de votre arrivée (à retirer dès l'ouverture des pistes). - L'information concernant le nombre de pistes et de remontées mécaniques est donnée à titre indicatif.
---	--

Tableau 3. Exemple de cellules du corpus avec ses trois parties.

Comme nous le voyons dans le tableau 3, la partie (1) est une indication temporelle concernant l'ouverture des pistes skiabiles. La partie (2) est une information générale, accompagnée d'une description géographique et informative du domaine skiable. Enfin, la troisième partie est une recommandation sur les conditions d'utilisation, accompagnée d'informations spécifiques et/ou d'une note. De ces premières observations, nous pouvons voir comment les parties sont rédigées et comment elles fonctionnent.

Ainsi, notre premier constat s'est porté sur l'importance de respecter la mise en forme du texte afin d'être compris par l'être humain comme par le système. Par exemple, la mauvaise mise en forme lorsque nous avons reçu le texte en format Excel a découlé sur une mauvaise ponctuation, à cause de laquelle le système a généré des sorties de systèmes erronées. Bien avant la phase du premier choix des règles de contrôle, cette règle générale d'agencement du texte s'impose

comme la plus importante à ce stade : règle 4.3. du SE (*Use a tabular layout (vertical layout) for complex texts*) pour tout ce qui concerne l'agencement des informations.

En ce qui concerne les parties prises une à une, nous constatons que les premières parties (1) traitant des informations générales sont pratiquement toujours identiques, se répétant dans la forme comme dans le contenu sauf pour ce qui est des dates. Par contre, les parties descriptives (2) (qui pourraient être également une partie indépendante) sont très variées de par leur langage littéraire et résolument publicitaire ; elles demanderont, à ce stade, le plus de contrôle. Les parties (3) sont faites des listes dans lesquelles diverses informations sont données. Toutefois, elles sont composées de segments à la structure courte et simple, présentant un même contenu mais se distinguant par des formulations diverses. Parfois, nous pouvons noter que le contenu et la forme de certains segments sont respectés mais leur emplacement dans le texte global est différent. Ces premières observations se recourent avec la typologie établie par Ruggero Druetta (2011) pour le discours utilisé en communication touristique et nous permettent de percevoir les caractéristiques macro-linguistiques observables pour l'organisation d'un texte de type touristique (tableau 4) :

<p>Fonctions :</p> <ul style="list-style-type: none"> ⇒ Fonction informative ⇒ Donner une image d'un objet qu'on ne voit pas ⇒ Ordonner un comportement au destinataire <p>Caractéristiques d'organisation :</p> <ul style="list-style-type: none"> ⇒ Progression à thème constant ⇒ Organisation de l'action en phases successives ⇒ Constance de ton employé
--

Tableau 4. Fonctions et caractéristiques d'organisation du corpus selon Druetta.

L'organisation des parties comme présentées dans notre corpus (tableau 3) est donc typique du discours touristique, qui comprend en effet des séquences-types de forme descriptive ou explicative (partie 2) et injonctive (parties 1 et 3). Ces indications nous amènent donc à penser le texte du corpus en parties et à imaginer déjà les premières règles spécifiques applicables aux diverses séquences. Par ailleurs, nous sommes en mesure également d'identifier les segments identiques ou répétitifs, nous conduisant ainsi à étudier le texte de manière micro-linguistique (Lalaude *et al.*, 1998).

3.3.3.2 Observation des phénomènes linguistiques

L'identification macro-linguistique permet de glisser à un niveau micro-linguistique puisque les composantes du texte ont déjà révélés les « variations lexicales et syntaxiques ». Nous avons ainsi pu grouper les formulations synonymes en identifiant les segments répétitifs (tableau 5) et en mettant en évidence l'inutilité des variations (Lalaude *et al.*, 1998 :106).

A noter :

- L'information concernant le nombre de pistes et de remontées mécaniques est donnée à titre indicatif.
- Prévoir un document officiel justifiant l'âge des participants.
- Prévoir une photo pour les enfants de - de 5 ans et les adultes de + de 72 ans.
- Forfait gratuit pour les enfants de - de 6 ans et pour les séniors de + de 75 ans.
- Pour les séjours du samedi 18/12 au dimanche 26/12/2010 (séjour de 8 nuits), votre forfait remontées mécaniques est valable à partir du lundi 20/12/2010 au matin.
- Pour les séjours du dimanche 02/01 au samedi 08/01/2011 ou du dimanche 09/01 au samedi 15/01/2011 (séjour de 6 nuits), votre forfait remontées mécaniques est valable le jour même de votre arrivée (à retirer dès l'ouverture des pistes).

Tableau 5. Exemple de segments répétitifs du corpus.

Il est intéressant de constater que ces segments appartiennent tous à la partie injonctive du texte (3) dont les caractéristiques-types sont énumérées au tableau 6. Ces phrases sont de style peu créatif de par la constance du ton, le caractère court des phrases et la fréquence des mêmes verbes d'action.

Caractéristiques lexicales :

- ⇒ Fréquences de verbes indiquant l'action à accomplir
- ⇒ Abondance de formes allocutives d'interpellation : « vous »

Caractéristiques grammaticales :

- ⇒ Temps verbal utilisé: impératif, infinitif,
- ⇒ Phrases courtes

Tableau 6. Caractéristiques lexicales et grammaticales pour les parties injonctives du corpus.

Ainsi, nous avons déjà pu établir que cette partie (3) nécessite une règle de contrôle générale indispensable à une bonne utilisation de la traduction

automatisée : utiliser toujours les mêmes formulations pour un même contenu. Ces segments-types sont donc ceux qui, à première vue, ne devraient pas poser de grands problèmes une fois la règle de contrôle appliquée.

En fait, l'approche micro-linguistique, dans notre cas, s'avère intéressante en ce qui concerne les parties descriptives. En effet, l'analyse micro-linguistique nous amène également à observer le fonctionnement linguistique des segments qui composent le texte. Dans notre corpus, chaque description propose des phrases qui se distinguent par leur longueur et leur structure complexe. Le travail de LC permettra de rendre ces structures plus simples et donc plus lisibles pour la traduction automatique. Toutefois, sans rendre une nouvelle fiche de lecture micro-linguistique détaillée, nous avons essayé de proposer une analyse sommaire mais qui donne un bon aperçu du fonctionnement de la description dans ce type de texte⁴. Ainsi nous pouvons observer que les descriptions de notre corpus:

- définissent un objet concret, un lieu (utilisation des verbes suivants : *être, constituer de, compter, bénéficier*)
- définissent les divers objets qui composent ce lieu (utilisation des verbes suivants : *proposer, offrir*)

⁴ Une analyse micro-linguistique approfondie pourrait être faite à l'aide d'outils informatiques pour l'étude des corpus monolingues tels que le logiciel Wordsmith. Dans le cadre de cette étude, le but final se trouvant ailleurs, nous n'avons fourni qu'une analyse sommaire mais suffisante.

- définissent les sujets à qui le lieu et ses composantes s'adressent ou du moins ce qu'ils y recherchent (utilisation de *vous* + *futur/présent* + *découvrir, profiter, promettre, permettre*).

Au niveau des caractéristiques observables (tableau 7), nous constatons que la typologie présentée par Druetta (2011) se recoupe encore une fois avec nos propres observations et ajoute même quelques pistes qui, au fil de notre travail sur le texte, se vérifieront.

<p>Caractéristiques lexicales :</p> <ul style="list-style-type: none"> ⇒ Emploi de vocabulaire spécialisé du domaine dont on parle ⇒ Emploi de vocabulaire analogique ⇒ Abondance de repères spatiaux ou temporels <p>Caractéristiques grammaticales :</p> <ul style="list-style-type: none"> ⇒ Temps verbal utilisé : présent ⇒ Fréquence des présentatifs : « c'est » ⇒ Abondance de groupes nominaux

Tableau 7. Caractéristiques observables pour les parties descriptives et explicatives du corpus.

A partir de là, nous pouvons dire que les structures de toutes les descriptions de notre corpus diffèrent très peu. Lors de notre évaluation du LC, il s'agira de constater si ces structures (descriptives) peuvent être conservées tout en étant simplifiées sans, si possible, sacrifier le ton publicitaire et vendeur de ces parties. Il faudra également voir si l'uniformisation, d'une façon ou d'une autre, de la structure des segments « similaires » est réalisable.

3.4 Conclusion

Les approches méthodologiques que nous avons entreprises à deux niveaux – micro-linguistique et macro-linguistique - nous ont amené à travailler sur le

texte de manière globale avant même d'avoir véritablement commencé le processus de pré-édition. En effet, nous avons pu distinguer les parties de textes et quel était leur fonctionnement, ce qui nous a permis de pouvoir établir deux premières règles générales et d'entrevoir les « travaux » nécessaires au bon déroulement de cette évaluation. Ainsi, en traitant le texte de manière factuelle (une fois manuellement et une fois technologiquement), nous pourrions établir des règles plus spécifiques. Pour le moment, il s'agit de commencer à travailler avec les différents outils afin de rendre notre texte le plus facile à lire pour la machine.

IV. Méthodologie

4.1 Introduction

Dans le chapitre précédant nous avons abordé une partie de notre marche à suivre en présentant le texte que nous avons choisi et les outils dont nous sommes servis afin d'atteindre l'objectif annoncé dans le chapitre I. Dans le présent chapitre, nous allons plus avant dans la description de la méthodologie choisie en présentant notre travail de pré-édition dans le détail.

La première section se penche sur la personnalisation du système de TA Systran ; c'est-à-dire sur la création des dictionnaires spécifiques au domaine traité. La seconde partie, quant à elle, concerne la sélection des règles de contrôle que nous avons jugées utiles à l'amélioration de la traduction de notre corpus par Systran. Nous jugerons de leur efficacité et de leur importance dans les chapitres V et VI.

4.2 Création des dictionnaires spécialisés

Le dictionnaire est la composante essentielle d'un système de TA. Plus que pour tout autre élément, la taille et la qualité du dictionnaire influe directement sur la qualité de la sortie du système, c'est pourquoi la création de dictionnaires spécialisés est une étape décisive du processus de traduction dans le cadre de la TA. Toutefois, bien que le dictionnaire d'un système de TA soit à l'heure actuelle d'une qualité équivalente à celle d'un dictionnaire « papier », il est loin d'être exhaustif et ne couvre pas tous les domaines potentiellement traités

par la TA (Arnold, 1991 : 91). De fait, le programme ne peut pas toujours traduire un texte donné de manière adéquate, tout simplement parce que les termes demandés ne se trouvent pas dans le lexique de base (Guerra, 2000: 100). C'est pourquoi les utilisateurs des systèmes de TA ont la possibilité de consulter, de modifier, d'ajouter et même d'effacer des termes. C'est la participation la plus importante de l'utilisateur au sein même du système de TA. En effet, l'ajout d'une terminologie propre au domaine concerné permet d'améliorer sans conteste la qualité des traductions des documents dits techniques (Guerra, 2000 : 91).

Comme nous l'avons expliqué dans la section 3.3, notre corpus est un texte de type touristique qui possède un domaine et un sous-domaine : le ski et les domaines skiables. Il s'agit donc d'un texte spécialisé comprenant un vocabulaire spécifique propre au domaine. Après une première traduction du texte par Systran, nous avons pu constater que la plupart des termes techniques inhérents au ski ne figuraient pas dans le(s) dictionnaire(s) existant(s). La première sortie du système contenait donc des mots inconnus du système, par conséquent non traduits ainsi que des mots mal traduits.

Systran, par le truchement du *Systran Dictionary Manager*, permet de créer trois types de dictionnaires : les dictionnaires utilisateurs, les dictionnaires de normalisation et les mémoires de traduction. Chaque dictionnaire est associé à un domaine, soit interne au programme, soit établi par l'utilisateur, dans notre cas, le ski. Deux autres paramètres permettent à l'utilisateur de personnaliser le dictionnaire selon ses besoins spécifiques. Tout d'abord, l'option « priorité » allant de 1 (la plus haute, l'entrée prend le pas sur n'importe quelle autre règle

interne au système ou expression longue et ne garde pas les homographes se trouvant dans le dictionnaire principal) à 9 (la plus basse, l'entrée n'est pas prise en compte). Selon le niveau de priorité choisi par l'utilisateur, les homographes présents dans le dictionnaire principal sont ou ne sont pas conservés et l'entrée a ou n'a pas la prévalence sur les plus longues expressions présentes dans le système. Enfin, l'option « codage expert », par rapport au codage intuitif dont nous avons parlé plus en détail dans la sous-section 3.2.1.1 permet à l'utilisateur d'affiner son entrée, par exemple en attribuant manuellement une catégorie grammaticale à l'entrée en choisissant dans le menu déroulant, et de là faire les modifications nécessaires (Systran UserGuide).

Nous avons donc décidé de créer trois dictionnaires utilisateurs et une mémoire de traduction, distincts par langue. Nous avons établi la priorité de nos dictionnaires à 3, afin que les homographes du dictionnaire principal ne soient pas pris en compte et que les entrées n'aient pas la priorité sur les règles grammaticales et les expressions plus longues du programme. Nous ne voulons pas dénaturer la sortie du système en choisissant une trop haute priorité. Autant que possible, nous avons utilisé l'option de codage expert, afin de contrôler au mieux nos entrées. Le premier dictionnaire que nous avons créé est de type DNT (*Do Not Translate*). Il contient tous les noms propres de notre corpus, afin que le système n'en tienne pas compte lors de la traduction, évitant de générer des erreurs comme le tableau 8 le démontre ci-dessous.

FR (...) et de l'autre côté le versant du Méné .
EN (...) and the other side the slope of Carried Out .
IT (...) e dell'altro lato la pendenza del condotto .
(Cellule 177)

Tableau 8. Exemple de sortie de système brute – nom propre

Le dictionnaire DNT français-anglais totalise 168 noms propres. Le dictionnaire français-italien en compte 164 (annexe 1).

Le second dictionnaire regroupe les termes techniques appartenant au domaine du ski, comme le *skwal* ou le *ski joëring*, mais aussi les « homographes », c'est-à-dire les termes figurant dans le lexique de Systran mais dont l'équivalence n'est pas correcte concernant notre domaine, comme « station » (*station/stazione*), « poudreuse » (*duster/spolveratore*) ou « forfait » (*fixed price/multa*). Nous avons également dû faire des modifications dans le texte source afin de spécialiser des termes trop généraux, comme « domaine » (*field/settore*), qui est devenu « domaine skiable » (*ski area/regione sciistica*) ou « station » qui est devenu « station de ski ». Nous avons entré 52 termes pour le dictionnaire français-anglais et 51 pour le français-italien (annexe 2).

Le troisième dictionnaire regroupe les mots inconnus du système ne faisant pas partie de la terminologie spécifique au ski ainsi que les expressions idiomatiques, comme « conditions climatiques » (*vagaries of the weather/condizioni climatiche* et non pas *climatic risks/rischi climatici*). Ce dictionnaire regroupe également les termes mal analysés par Systran, tel que le nom « dénivelé », qui en anglais est devenu l'adjectif *uneven* tandis qu'en italien il est devenu une

grossière périphrase : *mutato il livello*. Nous avons entré 51 termes pour l'anglais et 82 pour l'italien (annexe 3). A ce propos, il faut noter que le dictionnaire italien a été divisé en deux : un pour les mots mal traduits (54) et un autre pour les mots inconnus du système (28). Cette division nous a semblé importante au niveau de la post-édition que nous traiterons exclusivement pour l'italien dans le chapitre VII.

La mémoire de traduction (annexe 4) nous a été utile pour la partie répétitive de notre corpus : nous avons rentré 16 segments pour l'anglais et 14 pour l'italien.

Après l'implémentation des trois dictionnaires et de la mémoire de traduction dans le *Project Manager*, la sortie du système s'est révélée autrement plus compréhensible. Les noms propres ambigus ont été traduits par le système, et les termes techniques sont rendus correctement : *station* devient *resort* et *fixed price, ski pass*. Le terme « forfait » devient *skipass* en italien aussi. Il va sans dire que la création d'une terminologie spécifique au domaine du ski se révèle être une démarche importante pour améliorer la qualité de la sortie du système. En effet, l'utilisateur se doit de faire quelques modifications du dictionnaire de base pour rendre le système réellement exploitable. Bien sûr, la sortie du système n'est pas encore une traduction satisfaisante du texte source, mais, nous l'avons constaté, les dictionnaires spécialisés permettent une nette amélioration de celui-ci (tableau 9). Les dictionnaires sont donc une première étape vers une sortie satisfaisante du système, la seconde étant l'application des règles de contrôle que nous avons sélectionnées ce que nous pourrions vérifier et quantifier dans notre évaluation (chapitre V et VI).

Version brute :

FR : (...)Le domaine skiable d'Isola 2000 se divise en 2 vallées : d'un côté le versant de la Lombarde et de l'autre côté le versant du Méné.

(...)Le forfait comprend : - L'accès au domaine skiable d'Isola 2000.

(Cellule 177)

IT:(...) Il settore sciistico, di isolò 2000 è condiviso in due valli, con da un lato la pendenza del Lombarde e dell'altro lato la pendenza del condotto.

(...)La multa comprende: L'accesso al settore sciistico di isolò 2000.

EN:(...) The skiable field from Isolated 2000 is shared in two valleys, with side the slope on Lombarde and the other side the slope of Carried Out.

(...) The fixed price includes/understands : - the access to the skiable field of Insulated 2000.

Remarque : Nous pouvons voir que Systran a traduit « Isola 2000 » de deux façons différentes pour l'anglais.

Version avec dictionnaire :

IT:(...) La zona sciistica, d'Isola 2000 è condivisa in due valli con da un lato la pendenza di Lombarde e dell'altro lato la pendenza di Méné.

(...) Il skipass comprende: L'accesso alla zona sciistica d'Isola 2000.

EN:(...) The ski area of Isola 2000 is shared in two valleys, with side the slope on Lombarde and the other side the slope of Méné.

(...)The ski pass includes: - the access to the ski area of Isola 2000.

Tableau 9. Sorties de systèmes avec et sans dictionnaires utilisateurs.

4.3 La présélection des règles de contrôle.

Notre travail consiste à évaluer l'impact du langage contrôlé sur un corpus donné ; en sélectionnant des règles de LC qui nous paraissaient utiles, et en les appliquant à notre texte, la sortie du système qui en résulte devrait être grandement améliorée. Afin d'établir cet ensemble de règles de contrôle, nous nous sommes basé sur plusieurs sources : le guide GIFAS (1999), une série de règles créées par Uwe Muegge dans le but de faciliter la traduction automatique, l'article « Mtranslability » d'Arendse Bernth et de Claudia Gdaniec (2001) et l'article de Lalaude *et al* (1998).

Lalaude *et al.* (1998) définissent deux types de règles d'écritures : générales et spécifiques. Les règles dites générales peuvent s'appliquer à de nombreux documents, tandis que les règles spécifiques tiennent compte des particularités d'un document donné. Les règles générales se fondent sur l'observation des erreurs ou des ambiguïtés les plus répandues, et qui peuvent créer des difficultés de traduction. Les règles spécifiques sont définies après une analyse précise d'un texte ou d'un corpus particulier, comme nous l'avons fait à la section 3.3 (Lalaude *et al.*, 1998 :103-113).

Dans notre cas, certaines règles générales se sont imposées de manière assez évidente (sous-section 3.3.3.1) : l'orthographe, la ponctuation ou encore l'ajout de l'article sont finalement des principes relevant plus du bon sens que du langage contrôlé. Néanmoins, il est important de les inclure dans notre liste car nous avons observé qu'elles n'étaient pas toutes suivies avec constance. Dans un deuxième temps, nous nous sommes inspirées du Français Rationalisé en

reprenant quelques règles générales du guide GIFAS, notamment concernant le temps des verbes (règle n°2. 2. 1) et la longueur des phrases (règle n°3.4.3). Nous avons également consulté le site web d'Uwe Muegge (2006) où nous avons trouvé quelques règles intéressantes qui rejoignaient le guide GIFAS, comme le fait d'exprimer une seule idée par phrase, de faire attention à la longueur des mots, à la grammaticalité de la phrase ou encore d'éviter le passif.

En somme donc, nous avons défini les caractéristiques de notre corpus (section 3.3) en nous appuyant sur l'article de Lalaude *et al.* (1998) : un texte de type touristique, appartenant au domaine du ski et au sous-domaine des stations de ski. Ensuite, nous avons identifié les différentes parties qui composent notre corpus (appelées *textual modules*) afin de déterminer les règles spécifiques. Nous l'avons déjà dit, notre corpus est composé de cellules qui comportent pour la majorité d'entre elles trois parties distinctes, que nous avons appelé (1) indication (2) information et (3) recommandations (sous-section 3.3.3.1). En partant de ce principe, nous avons défini la mise en page des cellules, sur la base de la règle 4.3 du guide GIFAS. Les autres règles spécifiques que nous avons choisies nous ont été fournies par une analyse attentive de l'entrée du système et des difficultés de traduction présentes dans la sortie du système.

Voici les règles que nous avons sélectionnées :

4.3.1 Règles générales

i. Rédiger des phrases avec des mots correctement orthographiés.

Version brute :

FR : (...) L'accès illimité à tout le domaine de la **fôret** blanche.

IT : (...) L'accesso illimitato a tutto il settore del **trapano** bianco

EN : (...) The access unlimited to all the field of the **drill** white.

Version contrôlée sans dictionnaires :

FR : (...) L'accès illimité à tout le domaine skiable de la **Forêt Blanche**.

IT : (...) L'accesso illimitato a tutto il settore skiable della **foresta bianca**.

EN : (...) The access unlimited to all the skiable field of the **White Forest**.

(Cellule 324)

Tableau 10 . Entrées et sorties de systèmes avant et après application de la règle i.

Nous avons constaté que pour les fautes les plus communes Systran traduisait correctement les mots mal orthographiés. Dans d'autres cas, il a signalé le mot comme inconnu et a donc repris le mot français dans la sortie du système. Ce n'est que par l'adjonction du dictionnaire et/ou correction du mot que la sortie du système devient correcte. Toutefois, nous avons remarqué que la sortie du système italienne ne respectait pas toujours les majuscules (dans certains cas, le

ystème ne reconnaissait même plus le mot). En revanche, si l'erreur d'orthographe ou de typographie crée une ambiguïté, notamment au niveau de la catégorie grammaticale, le système délivre une traduction grammaticalement fausse.

Version brute :

FR : (...) À deux pas d'Annecy, la Clusaz cultive l'excellence en entrant dans le très sélect "Top of the French alps", qui **garantie** des prestations de qualité aux clientèles internationales les plus exigeantes.

(Cellule 2021)

IT : (...) A due pas d'Annecy, Clusaz coltiva l'eccellenza entrando in molto sélect "top of the French alps,, **che garanzia** delle prestazioni di qualità alle clientele internazionali più esigenti.

EN : (...) With two steps of Annecy, Clusaz cultivates excellence while entering very the sélect "Signal off the French alps", which **guaranteed** high quality services with the international customers most demanding.

Version contrôlée sans dictionnaires :

FR : (...) À deux pas d'Annecy, la Clusaz cultive l'excellence en entrant dans le très sélect "Top of the French Alps". La Clusaz **garantit** des prestations de qualité aux clientèles internationales les plus exigeantes.

IT : (...) A due pas d'Annecy, Clusaz coltiva l'eccellenza entrando in molto sélect "top of the French Alps,, Clusaz **garantisce** prestazioni di qualità alle clientele internazionali più esigenti.

EN : (...) With two steps of Annecy, Clusaz cultivates excellence while entering very the sélect "Signal off the French alps". Clusaz **guarantees** high quality services with the international customers most demanding.

Tableau 11. Entrées et sorties de systèmes avant et après application de la règle i – ex. 2

ii. Traiter d'un seul sujet par phrase.

Version brute :

FR : (...) Un **domaine** adapté aux débutants avec des pistes accessibles, ainsi qu'aux **skieurs chevronnés** qui prendront plaisir à découvrir ou redécouvrir les pistes olympiques ainsi que **le vertigineux couloir** de Casserousse qui permet de descendre sur plus de 850m de dénivelé.

(Cellule 359)

IT : (...) Un **settore** adeguato ai principianti con tracciati accessibili, come pure ai **skieurs esperti** che prenderanno piacere da scoprire o riscoprire i tracciati olimpici e **il corridoio vertigineux** di Casserousse che permette di scendere su più 850m di mutato il livello.

EN : (...) **A field** adapted to the beginners with accessible tracks, as with the **senior skiers** who will take pleasure to discover or **redécouvrir** the Olympic tracks as well as the **vertiginous corridor** of Casserousse which makes it possible to go down on more 850m of uneven.

Version contrôlée sans dictionnaires :

FR : (...) **Le domaine skiable** est adapté aux débutants avec des pistes accessibles ainsi qu'aux skieurs chevronnés. **Les skieurs chevronnés** prennent plaisir à découvrir ou redécouvrir les pistes olympiques ainsi que le vertigineux couloir de Casserousse. **Le couloir** de Casserousse permet de descendre sur plus de 850 m de dénivelé.

IT : (...) **Il settore skiable** è adattato ai principianti con tracciati accessibili e ai skieurs esperti. **I skieurs esperti** prendono piacere da scoprire o riscoprire i tracciati olimpici e il corridoio vertigineux di Casserousse. **Il corridoio** di Casserousse permette di scendere su oltre 850 m di mutato il livello.

EN: (...) **The skiable field** is adapted to the beginners with accessible tracks. **The senior skiers** take pleasure to discover or **redécouvrir** the Olympic tracks as well as the vertiginous corridor of Casserousse. **The corridor of Casserousse** makes it possible to go down on more 850m of uneven.

Tableau 12. Entrées et sorties de système avant et après application de la règle ii.

iii. Utiliser la même phrase pour exprimer le même contenu.

Version brute :

FR : (...) - Forfait **offert** pour les enfants de - de 5 ans et pour **les séniors de + de 75 ans**.

IT: (...) - Multa **offerta** per i bambini di - di 5 anni e per **i séniors di +de 75 anni**.

EN: (...) - Fixed price offered for the children of – 5 years and for the seniors of **+de 75 years**.

(Cellule 437)

FR : (...) - Forfait **gratuit** pour les enfants de - de 5 ans et pour **les séniors de + de 75 ans**.

IT: (...) - Multa **gratuita** per i bambini di - di 5 anni e per **i séniors di + di 75 anni**.

EN: (...) - **Free** Fixed price for the children of – 5 years and for the seniors of + 75 years.

(Cellule 491)

FR : (...) - **Gratuit** pour les enfants de - de 5 ans et **les adultes à partir de 75 ans (prévoir un document officiel justifiant l'âge des participants)**.

IT: (...) - **gratuito** per i bambini di - di 5 anni e **gli adulti da partire da 75 anni (prevedere un documento ufficiale che giustifica l'età dei partecipanti)**.

EN: (...) - **Free** for the children of – 5 years and **adults as from 75 years (to envisage an official document justifying the age of the participants.)**

(Cellule 847)

Version contrôlée sans dictionnaires :

FR : (...) Forfait **gratuit** pour les enfants de moins de 5 ans et pour les **adultes de plus** de 75 ans.

IT : (...) Multa **gratuita** per i bambini di meno di 5 anni e **gli adulti di oltre** 75 anni.

EN : (...) **Free** fixed price for the children of less than 5 years and **for the adults of** more than 75 years.

(Cellule 437, 491, 847)

Tableau 13. Entrées et sorties de système avant et après application de la règle iii.

iv. Utiliser le même terme à chaque fois.

Version brute :

FR : (...) Cet espace ludique et technique, vous propose des modules pour tous les niveaux : slopestyle, big air, handrails, **zone débutants...**

IT : (...) Questo spazio ludico e tecnico, vi propone moduli per tutti i livelli: slopestyle, big aria, handrails, **zona principianti...**

EN : (...) This ludic and technical space, proposes to you modules for all the levels: **initial** slopestyle, big air, handrails, **zone...**

(Cellule 177)

FR : (...) Ce réseau optimisé de canons à neige permet l'enneigement de 10 pistes tous niveaux ainsi que des **espaces débutants.**

IT : (...) Questa rete ottimizzata di cannoni da neve permette l'altezza del manto nevoso di 10 tracciati qualsiasi livello e

degli **spazi iniziali**.

EN: (...) This optimized network of snow-blowers allows the snowing up of 10 tracks all levels as well as **initial spaces**.

(Cellule 3594)

Version contrôlée sans dictionnaires :

FR: (...) Cet espace ludique et technique, vous propose des modules pour tous les niveaux : slopestyle, big air, handrails, **zone pour débutants...**

IT: (...) Questo spazio ludico e tecnica propone moduli per tutti i livelli: slopestyle, big aria, handrails, **zona per principianti...**

EN:(...) This ludic and technical space proposes modules for all the levels: slopestyle, big air, handrails, **zone for beginners...**

(Cellule 177)

FR: (...) Ce réseau optimisé de canons à neige permet l'enneigement de 10 pistes tous niveaux ainsi que des **zones pour débutants**.

IT: (...) Questa rete ottimizzata di cannoni da neve permette l'altezza del manto nevoso di 10 piste qualsiasi livello e delle **zone per principianti**.

EN: (...) This optimized network of snow-blowers allows the snow cover of 10 slopes all levels as well as **zones for beginners**.

(Cellule 3594)

Tableau 14. Entrées et sorties de système avant et après application de la règle iv.

v. **Répéter le nom pour éviter le pronom.**

Version brute :

FR : (...) **Le Domaine** de la Forêt Blanche, union des domaines de Vars et Risoul, est l'un des plus beaux espaces skiabiles des Alpes du Sud. **Il** est niché au milieu de splendides forêts de mélèzes, avec des sommets environnants (Massif des Ecrins, de l'Ubaye et du Queyras) et des pistes à près de 3000 m. Un enneigement de qualité est assuré toute la saison grâce à sa haute altitude et un réseau de neige de culture sur l'ensemble des axes principaux. **Il** offre une belle luminosité avec son climat déjà méditerranéen.

(Cellule 847)

IT : (...) **Il settore** della foresta bianca, unione dei settori di Vars e Risoul, è uno dei più begli spazi skiabiles delle Alpi del Sud. **È** nidificato in mezzo a foreste splendide di mélèzes, con vertici circostanti (Massif degli Ecrins, dello Ubaye e del Queyras) e tracciati quasi a 3000 m. un'altezza del manto nevoso di qualità è garantita tutta la stagione grazie alla sua alta altitudine ed una rete di neve di coltura su tutti gli assi principali. **Offre** una bella luminosità con il suo clima già mediterraneo.

EN : (...) **The Field** of the White Forest, union of the fields of VAr and Risoul, is one of most beautiful spaces skiabiles of the Alps of the South. **It** is nested in the middle of splendid forests of larches, with surrounding tops (Solid mass of the Jewel cases, Ubaye and Queyras) and tracks with nearly 3000 Mr. a snowing up of quality is assured all the season thanks to its high-altitude and a network snow of culture on the whole of the main axes. **It** offers a beautiful luminosity with its already Mediterranean climate.

Version contrôlée sans dictionnaires :

FR : (...) **Le domaine skiable** de la Forêt Blanche réunit les stations de ski de Vars et Risoul. La Forêt Blanche est l'un des

plus beaux espaces skiabiles des Alpes du Sud. **Le domaine skiable** est niché au milieu de splendides forêts de mélèzes, entouré de sommets (Massif des Ecrins, de l'Ubaye et du Queyras). Les pistes se situent à près de 3000 m. L'altitude et un réseau de neige de culture assure un enneigement de qualité sur l'ensemble des axes principaux. **Le domaine skiable** offre une belle luminosité avec son climat déjà méditerranéen.

IT: (...) **Il settore skiable** della foresta bianca riunisce le stazioni di sci di Vars e Risoul. La foresta bianca è uno dei più begli spazi skiabiles delle Alpi del Sud. **Il settore skiable** è nidificato in mezzo a foreste splendide di mélèzes, circondato di vertici (massiccio degli scrigni, dello Ubaye e del Queyras). I tracciati si situano quasi a 3000 m. l'altitudine ed una rete di neve di cultura garantisce un'altezza del manto nevoso di qualità su tutti gli assi principali. **Il settore skiable** offre una bella luminosità con il suo clima già mediterraneo.

EN: (...) **The ski area** of the Forêt Blanche joins together the resorts of Vars and Risoul. The Forêt Blanche is one of most beautiful skiable spaces of the Alps of the South. **The ski area** is nested in the middle of splendid forests of larches, surrounded by tops (Massif of Ecrins, Ubaye and Queyras). The ski slopes are located at nearly 3000 Mr. altitude and a network of snow of culture ensures a snow cover of quality on the whole of the main axes. **The ski area** offers a beautiful luminosity with its already Mediterranean climate.

Tableau 15. Entrées et sorties de système avant et après application de la règle v.

Nous avons pu constater que les pronoms ne posaient pas de réels problèmes, ni en italien ni en anglais. Il s'agit simplement, en proposant cette règle, de simplifier et de désambiguïser le texte source au maximum. En effet, le deuxième pronom peut se référer au domaine ou à l'enneigement et il convient de clarifier la situation pour un meilleur confort de lecture.

vi. Un terme ne peut appartenir qu'à une seule catégorie syntaxique.

Version brute :

FR : (...) la variété des pistes de ski autorise la pratique de toutes les formes de ski : **ski débutant**, ski facile (...)

IT : (...) Valfréjus offre un settore skiabile de plus di 70 km² e la varietà dei tracciati di sci autorizza la pratica di tutte le forme di sci: **sci iniziale**, sci facile(...)

EN : (...) the variety of the ski slopes authorizes the practice of all shapes of ski: **ski beginning**, easy ski (...)

(Cellule 288)

FR : (...) Un domaine adapté **aux débutants** avec des pistes accessibles, ainsi qu'aux skieurs chevronnés qui prendront plaisir à découvrir ou redécouvrir les pistes olympiques (...)

IT : (...) Un settore adeguato **ai principianti** con tracciati accessibili, come pure ai skieurs esperti che prenderanno piacere da scoprire o riscoprire i tracciati olimpici (...)

EN : (...) A field adapted **to the beginners** with accessible tracks, as with the senior skiers who will take pleasure to discover or redécouvrir the Olympic tracks (...)

(Cellule 359)

FR : (...) **Glisseurs débutants**, plus sportifs ou passionnés de sensations fortes, la station d'Ax vous propose 3 domaines (...).

IT : (...) **Glisseurs iniziali**, più sportivi o appassionati di sensazioni forti, la stazione di Ax vi propone 3 settori tutt'al più vicino alla neve ed ai grandi spazi.

EN : (...) **Initial slide blocks**, more sporting or impassioned strong feelings, the station of Ax proposes to you 3 fields with more close to snow and big spaces.

(Cellule 1874)

Version contrôlée sans dictionnaires :

FR : (...) La variété des pistes de ski autorise la pratique de tous les styles de ski : **ski pour débutants**, ski facile, (...)

IT: (...) La varietà dei tracciati di sci autorizza la pratica di tutti gli stili di sci: **sci per principianti**, sci facile, sci di scuola, sci di gobbe, (...)

EN: (...) The variety of the ski slopes authorizes the practice of all the styles of ski : **ski for the beginners**, easy ski, (...)

(Cellule 288)

FR: (...) Le domaine skiable est adapté **aux débutants** avec des pistes accessibles ainsi qu'aux skieurs chevronnés.

IT: (...) Il settore **skiable** è adattato **ai principianti** con tracciati accessibili e ai skieurs esperti.

EN: (...) The skiable field is adapted **to the beginners** with accessible tracks as with the senior skiers.

(Cellule 359)

FR: (...) La station d'Ax propose 3 domaines skiabiles au plus près de la neige et des grands espaces pour **les débutants**, les skieurs plus sportifs et les amateurs de sensations fortes.

IT: (...) La stazione di Ax propone 3 settori skiabiles tutt'al più vicino alla neve ed ai grandi spazi per **i principianti**, gli sciatori più sportivi e gli appassionati di sensazioni forti.

EN: (...) The station of Ax proposes 3 skiabiles fields with more close to snow and big spaces for **the more sporting beginners**, skiers and the enthusiasts for big thrills.

(Cellule 1874)

Tableau 16. Entrées et sorties de système avant et après application de la règle vi.

4.3.2 Règles spécifiques

vii. Découper phrases longues en phrases courtes: 20 mots pour les recommandations et 30 pour les descriptions (GIFAS).

Version brute :

FR : L'Espace Paradiski c'est du ski grand format créant ainsi un domaine exceptionnel, avec plus de 420 km de pistes de ski, 2 glaciers skiabiles et 2 sommets culminants à plus de 3000 m d'altitude. **34 mots**

(Cellule 700)

IT: Lo spazio Paradiski è dello sci grande formato che crea così un settore eccezionale, con oltre 420 km di tracciati di sci, 2 ghiacciai skiabiles e 2 vertici culminants a oltre 3000 m d'altitudine.

EN: Space Paradiski it is ski large size thus creating an exceptional field, with more than 420 m of ski slopes, 2 glaciers skiabiles and 2 tops culminating with more than 3000 m of altitude.

Version contrôlée sans dictionnaires :

FR : L'Espace Paradiski propose du ski grand format sur un domaine skiable exceptionnel. **12 mots**

La station de ski compte plus de 420 km de pistes de ski, 2 glaciers skiabiles et 2 sommets culminants à plus de 3000 m d'altitude. **27 mots**

IT: Lo spazio Paradiski propone dello sci grande formato su un settore skiable eccezionale.

La stazione di sci conta oltre 420 km di tracciati di sci, 2

ghiacciai skiables e 2 vertici culminanti a oltre 3000 m d'altitudine.

EN: Paradiski space proposes ski large size on an exceptional skiable field.

The station counts more than 420 km of ski slopes, 2 glaciers skiables and 2 tops culminating with more than 3000 m of altitude.

Tableau 17. Entrées et sorties de système avant et après application de la règle vii – ex. 1.

Il faut préciser que la limite de 30 mots pour les descriptions a été parfois difficile à respecter en italien comme en anglais. En effet, certains termes comme « station » devait impérativement être suivi du terme « de ski » afin de désambiguïser le texte au maximum : les termes *station* et *stazione* désignent communément la gare. Ainsi, la limite a été souvent dépassée de deux ou trois mots. De même pour les recommandations, dont les phrases ne peuvent pas dépasser 20 mots. Certaines phrases, notamment celles qui indiquaient des périodes de séjours, n'ont pu être réduites à la limite imposée.

Version brute :

FR : Pour les séjours du dimanche 02/01 au samedi 08/01/2011 ou du dimanche 09/01 au samedi 15/01/2011 (séjour de 6 nuits), votre forfait remontées mécaniques est valable le jour même de votre arrivée (à retirer dès l'ouverture des pistes). **38 mots**

(Cellule 142)

IT: Per i soggiorni della domenica 2 /01 alla sabato 8/01/2011 o della domenica 9 /01 alla sabato 15/01/2011 (soggiorno di 6 notti), la vostra multa risalite meccanici è valida il giorno stesso del vostro arrivo (a ritirare fin dall'apertura tracciati).

EN: For the stays of Sunday the 2nd/01 at the Saturday 1/8/2011 or

Sunday the 9th/01 at the Saturday 1/15/2011 (6 nights stay), your ski pass ski lift is valid the very same day your arrival (to be withdrawn as of the opening of the slopes).

Version contrôlée sans dictionnaires :

FR : - Pour les séjours du dimanche 02/01/2011 au samedi 08/01/2011 (6 nuits) le forfait pour les remontées mécaniques est valable le jour même de votre arrivée. Le forfait est à retirer dès l'ouverture des pistes. **25 + 9 mots**

- Pour les séjours du dimanche 09/01/2011 au samedi 15/01/2011 (6 nuits), votre forfait pour remontées mécaniques est valable le jour même de votre arrivée. Le forfait est à retirer dès l'ouverture des pistes.

IT: - Per i soggiorni dal del domenica 2/01/2011 al sabato 8/01/2011 (6 notti), la vostra multa risalite meccanici è valida il giorno stesso del vostro arrivo. La multa è da ritirare al momento dell'apertura dei tracciati.

- Per i soggiorni dal del domenica 9/01/2011 al sabato 15/01/2011 (6 notti), la vostra multa risalite meccanici è valida il giorno stesso del vostro arrivo. La multa è da ritirare al momento dell'apertura dei tracciati.

EN: - For the stays from Sunday 1/2/2011 to Saturday 1/8/2011 (6 nights) your fixed price for the ski lifts is valid the very same day of the arrival. The fixed price is to be withdrawn at the time of the opening of the slopes.

- For the stays from Sunday 1/9/2011 to Saturday 1/15/2011 (6 nights) your fixed price for the ski lifts is valid the very same day of the arrival. The fixed price is to be withdrawn at the time of the opening of the slopes.

(Cellule 142)

Tableau 18. Entrées et sorties de système avant et après application de la règle vii – ex2.

viii. Mise en page verticale à l'aide de puces : une majuscule au début, un point final lorsque la phrase est complète. Introduction de chaque point à l'aide d'une puce, lettre ou nombre (cf. recommandations AECMA SE règle n°4.3).

Version brute :

FR : (...) Le forfait Grande Galaxie Visalp comprend : - L'accès illimité au domaine des Grandes Rousses. - Grande Galaxie : 1 jour dans chacun des trois domaines suivants : Serre-Chevalier, Puy-Saint-Vincent et la Voie Lactée italienne, ainsi que 2 jours aux 2 Alpes. - Ski en nocturne le mardi et le jeudi. - L'accès aux patinoires de l'Alpe d'Huez, Vaujany ou d'Auris (location de matériel en sus) - L'accès aux 2 piscines chauffées (couverte et de plein air) - L'accès activités du Palais des Sports et des Congrès (location de matériel en sus) - L'accès au Musée d'Huez et de l'Oisans. - 1 concert d'orgue. A noter : - L'échange "Grande Galaxie" peut être interrompu à tout moment. - L'information concernant le nombre de pistes et de remontées mécaniques est donnée à titre indicatif. - Prévoir un document officiel justifiant l'âge des participants. - Prévoir une photo pour les enfants de - de 5 ans et les adultes de + de 72 ans.

(Cellule 514)

IT : (...) La multa grande galassia Visalp comprende: - L'accesso illimitato al settore delle grandi Rousses. - Grande galassia: 1 giorno in ciascuno dei tre settori seguenti: Serre-Chevalier, Puy-Saint-Vincent e la via lattea italiana, come pure 2 giorni alle 2 Alpi. - Sci notturno nel martedì ed il giovedì. - L'accesso alle patinoires dello Alpe di Huez, Vaujany o di Auris (affitto di materiale per di più) - l'accesso alle 2 piscine riscaldate (glassa e d'aria piena) - l'accesso attività del palazzo degli sport e dei congressi (affitto di materiale per di più) - l'accesso al museo di Huez e del Oisans. - 1 concerto d'organo. Notare: - Lo scambio "grande galassia,, può interrompersi in qualsiasi momento. - L'informazione riguardante il numero di tracciati e di risalite meccaniche è fornita a titolo indicativo. - Prevedere un documento ufficiale che giustifica l'età dei partecipanti. - Prevedere una fotografia per i bambini di meno di

5 anni e gli adulti di + di 72 anni.

EN: The fixed price Grande Visalp Galaxy includes/understands: - The access unlimited to the field of Large Russet-red. - Large Galaxy: 1 day in each of the three following fields: Serre-Chevalier, Puy-Saint-Vincent and the Italian Milky Way, like 2 days in the 2 Alps. - Ski into night Tuesday and Thursday. - The access to the skating rinks of Alpe d' Huez, Vaujany or of Auris (hiring of material in addition) - the access to the 2 heated swimming pools (covered and with outdoor) - the access activities of the Congress and Sport hall (hiring of material in addition) - the access to the Museum of Hoot and Oisans. - 1 concert of organ. To note: - The exchange "Large Galaxy" can be stopped constantly. - Information concerning the number of tracks and ski lifts is given as an indication. - To envisage an official document justifying the age of the participants. - To envisage a photograph for the children of - 5 years and adults of + 72 years.

Version contrôlée sans dictionnaires :

FR : (...) Le forfait Grande Galaxie Visalp comprend :

- L'accès illimité au domaine skiable des Grandes Rousses
- 1 jour dans chacun des trois domaines skiabiles suivants : Serre-Chevalier, Puy-Saint-Vincent et la Voie Lactée italienne
- 2 jours aux 2 Alpes
- Ski en nocturne le mardi et le jeudi
- L'accès aux patinoires de l'Alpe d'Huez, de Vaujany ou d'Auris (location de matériel en sus)
- L'accès aux 2 piscines chauffées (couverte et de plein air)
- L'accès aux activités du Palais des Sports et des Congrès (location de matériel en sus)
- L'accès au Musée d'Huez et de l'Oisans
- 1 concert d'orgue.

Remarques :

- L'échange "Grande Galaxie" peut être interrompu à tout moment
- Prévoir un document officiel justifiant l'âge des participants.
- Prévoir une photo pour les enfants de moins de 5 ans et les adultes de plus de 72 ans.

IT: (...) La multa grande galassia Visalp comprende:

- L'accesso illimitato al settore skiable delle grandi Rousses
- 1 giorno in ogni tre settore skiabiles seguenti: Serre-Chevalier, Puy-Saint-Vincent e la via lattea italiana
- 2 giorni alle 2 Alpi
- Sci notturno nel martedì ed il giovedì
- L'accesso alle patinoires dello Alpe di Huez, di Vaujany o di Auris (affitto di materiale per di più)

- L'accesso alle 2 piscine riscaldate (glassa e d'aria piena)
- L'accesso alle attività del palazzo degli sport e dei congressi (affitto di materiale per di più)
- L'accesso al museo di Huez e del Oisans
- 1 concerto d'organo.

Osservazioni:

- Lo scambio "grande galassia,, può interrompersi in qualsiasi momento.
- Prevedere un documento ufficiale che giustifica l'età dei partecipanti.
- Prevedere una fotografia per i bambini di meno di 5 anni e gli adulti di oltre 72 anni.

EN: The fixed price Grande Visalp Galaxy includes/understands:

- The access unlimited to the skiable field of Large Russet-red.
- 1 day in each of the three following fields: Serre-Chevalier, Puy-Saint-Vincent and the Italian Milky Way
- 2 days in the 2 Alps
- Ski into night Tuesday and Thursday
- The access to the skating rinks of Alpe d' Huez, Vaujany or Auris (hiring of material in addition)
- The access to the 2 heated swimming pools (covered and with outdoor)
- The access to the activities of the Congress and Sport hall (hiring of material in addition)
- The access to the Museum of Hoot and Oisans
- 1 concert of organ.

Note:

- The exchange "Large Galaxy" can be stopped constantly.
- To envisage an official document justifying the age of the participants.
- To envisage a photograph for the children of less than 5 years and the adults of more than 72 years.

Tableau 19. Entrées et sorties de système avant et après application de la règle viii.

ix. Eviter l'utilisation des clitiques.

Version brute :

FR : (...) la Toussuire ouvre directement sur 310 km de pistes reliées. Elle **vous** offre en plus une vue panoramique à 360° et un taux d'ensoleillement exceptionnel.

(Cellule 434)

IT: (...) Toussuire apre direttamente su 310 km di tracciati collegati. **Li** offre più in una vista panoramica a 360° ed un tasso di ensoleillement eccezionale.

EN: (...) Toussuire opens directly on 310 km of tracks connected. It offers **to you** in more one panoramic sight with 360° and a rate of exceptional sunning.

Version contrôlée sans dictionnaires :

FR : (...) **La station de ski offre** en plus une vue panoramique à 360° et un taux d'ensoleillement exceptionnel.

IT: (...) **La stazione di sci offre** più in una vista panoramica a 360° ed un tasso di ensoleillement eccezionale.

EN: (...) The ski station offers in more one panoramic sight to 360° and a rate of exceptional sunning.

Tableau 20. Entrées et sorties de système avant et après application de la règle ix.

Le discours publicitaire tend à interpeller le lecteur en s'adressant directement à lui, dans l'exemple ci-dessus par l'emploi du pronom « vous ». Toutefois, la langue italienne a deux équivalents, selon que l'on s'adresse à un groupe – « voi » - ou comme forme de politesse – « Lei ». Dans notre cas, la sortie de système est plutôt inattendue puisqu'elle propose un pronom qui ne correspond pas du tout au « vous » mais au pronom « les ».

x. Rajouter les prépositions.

Version brute :

FR : (...) Surplombé par des sommets Pyrénéens de plus de 3000m. Le domaine de Peyragudes vous accueille pour **des séjours sport ou détente.**

(Cellule 389)

IT : (...) Stato sospeso su da vertici Pyrénéens di più 3000m. Il settore di Peyragudes li accoglie per **soggiorni sport o rilassamento.**

EN : (...) Overhung by Pyrenean tops of more than 3000m. The field of Peyragudes accommodates you for **stays sport or relaxation.**

Version contrôlée sans dictionnaires :

FR : (...) Les sommets Pyrénéens de plus de 3000 m surplombent le domaine skiable de Peyragudes. La station de ski propose **des séjours de sport ou de détente.**

IT : (...) I vertici Pyrénéens de plus di 3000 m sono sospeso su il settore skiable di Peyragudes. La stazione di sci propone soggiorni **di sport o di rilassamento.**

EN : (...) The Pyrénéens tops of more than 3000 m overang the skiable field of Peyragudes. The ski station proposes **stays of sport or relaxation.**

Tableau 21. Entrées et sorties de système avant et après application de la règle x.

Ajouter les prépositions s’insère dans une règle plus générale : éviter l’utilisation des ellipses. Dans notre cas, toutefois, nous avons constaté que les problèmes les plus fréquents concernaient les prépositions, c’est pourquoi nous avons décidé d’omettre la règle générale au profit de la règle plus spécifique.

xi. Eviter l'utilisation du gérondif et participe présent.

Version brute :

FR : (...) Son orientation assure un ensoleillement important et c'est un plaisir de parcourir les pistes balisées **en sachant** que l'on rentrera skis aux pieds à la station.

(Cellule 7075)

IT: (...) Il suo orientamento garantisce un soleggiamento importante ed è un piacere di percorrere i traccati segnalati **in tenuto conto** che si rientreranno sci ai piedi alla stazione.

EN: (...) Its orientation ensures an important sunning and it is a pleasure of traversing the tracks marked **by knowing** that one will return skis to the feet to the station.

Version contrôlée sans dictionnaires :

FR : Son orientation assure un ensoleillement important. Parcourir les pistes de ski balisées est un plaisir quand on sait qu'on rentrera skis aux pieds à la station de ski.

IT: Il suo orientamento garantisce un soleggiamento importante. Percorrere le piste segnalate è un piacere **quando si sa** che si rientreranno sci ai piedi alla stazione sciistica.

EN: Its orientation ensures an important sunning. To traverse the marked out ski slopes is a pleasure **when you know** that one will return skis to the feet to the ski station.

Tableau 22. Entrées et sorties de système avant et après application de la règle xi.

xii. Eviter l'utilisation des verbes réfléchis.

Pour l'italien, le logiciel ne pose pas de problème quant il s'agit de simplement traduire un verbe réfléchi déjà présent dans le dictionnaire général de Systran. Par contre l'introduction d'un verbe réfléchi inconnu dans le dictionnaire pose effectivement problème puisque le système ne sait plus où placer la marque du réfléchi (« se » ou « si »). Donc dans ces cas-là, le mieux est de recourir à un verbe synonyme qui ne serait pas réfléchi, dans la mesure du possible.

Version brute :

FR : (...) Les familles, elles, peuvent profiter d'un cadre reposant où il est aisé de se ressourcer et de se changer la vie.

(Cellule 433)

IT : (...) Le famiglia, possono approfittare di un quadro riposante dove è facile **ressourcer** e di **cambiarsi la vita**.

EN : (...) The families, they, can benefit from a resting framework where it is easy **ressourcer** and to **change the life**.

Version contrôlée avec dictionnaires :

FR : (...) Les familles peuvent profiter d'un cadre reposant pour se ressourcer et se changer la vie.

IT : (...) Le famiglia possono approfittare di un quadro riposante per rimetter in forma **si** e **cambiarsi la vita**.

EN : (...) The families can benefit from a peaceful setting for **to recharge one's batteries** and **change the life**.

Tableau 23. Entrées et sorties de système avant et après application de la règle xii.

4.4 Conclusion

La spécialisation des dictionnaires a permis d'améliorer la qualité des sorties brutes et contrôlées du système, en réduisant les ambiguïtés et les inconnues. Nous le verrons plus en détail dans les chapitres V et VI lorsque nous en ferons l'évaluation. De même, les règles que nous avons sélectionnées, qu'elles soient générales ou spécifiques, ont permis une optimisation significative de la sortie du système générée par Systran, comme nous pouvons le voir dans les exemples ci-dessus. Il ne faut pas oublier qu'un système de traduction automatique, même très performant, n'a pas les capacités d'un traducteur humain ni sa connaissance du monde. Or, le simple fait de traduire exige du traducteur un ensemble de connaissances et une expérience de vie qui lui permet de résoudre les difficultés et les ambiguïtés de la langue. En résumé, la limitation principale d'un système de traduction automatique se situe dans le fait que la traduction ne se résume pas à une combinaison d'algorithmes (Arnold, 2003 : 122). Mais comme les systèmes de traduction automatique fonctionnent essentiellement grâce à une combinaison d'algorithmes, il est important de faire en sorte que la machine puisse lire et comprendre l'entrée du système afin de produire une sortie du système prête à être post-édité.

C'est là que les langages contrôlés et les dictionnaires utilisateurs entrent en jeu : plus l'entrée du système suit les règles d'un LC donné, plus elle sera compréhensible pour le système de TA ; plus l'utilisateur personnalisera les dictionnaires par ses entrées et modifications, plus le système sera exploitable. Comme nous l'avons constaté avec les exemples donnés, la sortie du système

contrôlée donne de meilleurs résultats que la sortie du système brute, nous fournissant ainsi un argument de poids pour notre évaluation.

V. Préparation de l'évaluation

5.1 Introduction

Ce chapitre présente la préparation de l'évaluation (proprement dite) des méthodes d'amélioration de la traduction (le langage contrôlé et les dictionnaires utilisateurs) dont nous avons parlé dans les chapitres précédents. Nous nous sommes inspirées librement des principes du 7-step recipe pour le « EAGLES Working Evaluation Group » (1996) afin de répondre aux questions suivantes :

- 1) l'ajout de dictionnaires utilisateurs est-il nécessaire?
- 2) le LC est-il essentiel à la bonne compréhension du texte ?
- 3) peut-on obtenir une bonne traduction en optant pour l'une ou l'autre de ces méthodes ou faut-il nécessairement les deux ?
- 4) l'outil de vérification Acrolinx, dans sa version de base, offre-t-il une aide efficace à ce stade du processus de la traduction ?

Avant de répondre à ces questions, nous allons exposer la méthodologie utilisée pour ce travail (section 5.2), puis, dans un second temps, introduire les critères de qualité que nous avons choisis afin de vérifier l'impact des différentes méthodes (section 5.3) et enfin les tests effectués (section 5.4). Les résultats des tests feront l'objet du chapitre suivant.

5.2 Méthodologie : corpus, experts et outils

5.2.1 Les corpus tests

Notre travail consiste à mesurer l'efficacité des dictionnaires spécialisés et du LC pour la traduction d'un texte de type touristique à l'aide d'un système de traduction automatique. Comme nous l'avons expliqué dans la sous-section 3.2.2 du chapitre III, notre corpus provient du document sur les remontées mécaniques que nous a envoyé l'administrateur d'Altiservice. Étant donné que le document original était conséquent, nous avons décidé de ne garder que la moitié des cellules, soit un corpus source totalisant un peu plus de 10'200 mots que nous appellerons (A). Depuis (A), nous avons créé le corpus source contrôlé (C), qui totalise 11'736 mots.

A partir du corpus (A), nous avons produit deux sorties pour chaque langue (annexes 5 et 6):

(A_Systran) brute sans dictionnaire,

(B_Systran) brute avec dictionnaires.

A partir du corpus (C), nous avons produit les sorties suivantes (annexes 7 et 8) :

(C_Systran) contrôlé sans dictionnaire,

(D_Systran) contrôlé avec dictionnaires.

Par la suite, nous avons créé un mini-corpus à partir du corpus (A) que nous avons nommé (E) : nous avons pris les cellules 142 à 342, autrement dit 1522 mots en tout. Nous avons soumis ce corpus à l'outil de vérification Acrolinx IQ, présenté dans la première partie du chapitre précédant dans le but de mesurer son efficacité. Nous avons donc deux sorties pour le mini corpus (annexe 9 et 10) :

(E_Systran) brut sans dictionnaire

(F_Systran) contrôlé par Acrolinx sans dictionnaire.

Enfin, nous avons créés deux sous-corpus pour les tests 3 et 4 l'un en fonction des phrases à contrôler et l'autre de manière aléatoire. Le premier sous-corpus, que nous appellerons (G), est composé de 28 phrases provenant du corpus source brut sans dictionnaire et leurs traductions en italien et en anglais, que nous avons appelé α , et de 28 phrases provenant du corpus contrôlé sans dictionnaire, et de leurs traductions en italien et en anglais que nous avons appelé β (pour plus de détails, voir annexe 11). Le second sous-corpus, que nous appellerons (H), est composé de 78 phrases sélectionnées parmi les versions (A_Systran), (B_Systran), (C_Systran), (D_Systran), (E_Systran) et (F_Systran) en fonction des problèmes posés par la langue ; par conséquent, les phrases du test anglais ne sont pas les mêmes que celles du test italien (pour plus de détails, voir annexe 12).

Nous avons fait un récapitulatif des corpus dans le tableau ci-dessous (tableau 24) :

(A)	(A_Systran)
	(B_Systran)
(C)	(C_Systran)
	(D_Systran)
(E)	(E_Systran)
	(F_Systran)
(G)	28 phrases de (A_Systran)
	28 phrases de (C_Systran)
(H)	78 phrases de (A_Systran), (B_Systran), (C_Systran), (D_Systran), (E_Systran) et

	(F_Systran)
--	-------------

Tableau 24. Récapitulatif des corpus

5.2.2 Les experts

A l'exception du test de compréhensibilité (sous-section 5.4.3), nous avons réalisé et noté nous-mêmes les différents tests portant sur l'importance des dictionnaires utilisateurs pour le système de TA et l'impact du langage contrôlé. Nous sommes de langue maternelle française et avons une excellente connaissance de l'anglais et de l'italien, avec une expérience non professionnelle de la traduction automatique.

Pour le test de compréhensibilité (sous-section 5.4.3.1), nous avons soumis le corpus (H) (sous-section 5.2.1) à quatre experts, deux pour l'italien, deux pour l'anglais, afin qu'ils puissent juger de la qualité des versions traduites de notre corpus source.

Pour la langue italienne, nous avons fait appel à deux personnes qui présentent les caractéristiques suivantes :

- Le 1^{er} expert est une professeure d'italien en école secondaire II. D'origine italienne, elle a fait toutes ses classes en Suisse et est parfaitement bilingue (français-italien). Pour information, elle ne possède aucune connaissance spécifique sur le domaine traité dans notre corpus.
- Le 2nd expert est constructeur d'appareils industriels. Vivant en Italie, il a fait toutes ses classes dans son pays et n'a aucune notion de français. Contrairement au premier expert, il est un grand sportif et a donc de bonnes connaissances du domaine traité dans notre corpus.

Pour la langue anglaise, nous avons fait appel à deux personnes qui présentent les caractéristiques suivantes :

- Le 1^{er} expert est diplômé d'histoire de l'art et travaille dans une galerie à Londres. Elle est d'origine américaine et française et a fait ses classes en France voisine et en Suisse. Elle est donc parfaitement bilingue. Elle possède des connaissances basiques de la terminologie du domaine traité.
- Le 2nd expert est traducteur professionnel au Portugal. Anglaise, elle traduit du portugais vers l'anglais et a des notions de français. Elle ne possède pas non plus une connaissance approfondie du domaine traité, mais étant traductrice professionnelle, elle peut apporter une perspective intéressante à notre évaluation.

5.2.3 Les outils

Les outils que nous allons utiliser et commenter dans ce chapitre ont déjà fait l'objet d'une présentation plus approfondie dans le chapitre III. Il s'agit du système de traduction automatique Systran, de l'outil de vérification Acrolinx IQ et de l'outil de corpus Wordsmith.

5.3 Définition des critères de qualité

En l'absence d'un véritable standard de mesures, il est plutôt difficile de déterminer ce qui fait une bonne traduction ou non (Miller 2001 : 1). Les mesures généralement utilisées pour l'évaluation des textes traduits automatiquement, telles que BLEU (*Bilingual Evaluation Understudy* - Papineni et al., 2001), ROUGE (*Recall-*

Oriented Understudy for Gisting Evaluation – Lin, 2004) ou WER (*Word Error Rate*) nécessitent impérativement des textes de références traduits manuellement. Comme nous ne disposons pas de telles références, nous sommes nous sommes limitées à une évaluation humaine, en établissant des critères d'évaluations personnels, bien qu'inspirés d'autres recherches académiques.

Idéalement, la sortie du système devrait être compréhensible à la première lecture. Nous savons que ce n'est pas le cas : les traductions du corpus A le montrent (annexe 5).

Nos hypothèses sont les suivantes :

- La spécialisation des dictionnaires du système de TA doit permettre de produire une sortie du système conforme à la terminologie du texte source.
- Le langage contrôlé améliore la traduction automatique.

Pour tester ces hypothèses, nous avons sélectionné différents critères de qualité qui nous permettront de juger de l'impact de la spécialisation des dictionnaires et du travail de pré-édition.

5.3.1 Couverture lexicale : termes et vocabulaire

Par couverture lexicale, nous entendons que la sortie du système doit contenir précisément les mêmes informations que l'entrée du système. Lorsque le système de TA traite d'un domaine dont la terminologie ne figure pas dans les dictionnaires inhérents au système, il y a des chances pour que l'information ne soit pas retransmise fidèlement. Dans ce cas, la spécialisation des dictionnaires s'avère nécessaire pour s'assurer de l'exactitude des informations contenues dans la sortie du système.

5.3.2 Langage contrôlé : efficacité des règles de contrôle

Nous entendons par « efficacité des règles de contrôle » l'impact que celles-ci ont sur le texte cible pour le système de TA. En effet, les règles de contrôle devraient permettre au système de TA de mieux appréhender le texte source, et donc de fournir une meilleure traduction.

5.3.3 Langage contrôlé : compréhensibilité

L'objectif de ce travail est d'obtenir un texte publiable sur un site internet, qui soit grammaticalement et lexicalement compréhensible par tous. La compréhensibilité concerne la facilité de compréhension à la première lecture (Miller *et al.* : 2001). Il s'agit donc de l'impact du langage contrôlé sur la compréhension du texte cible par des êtres humains.

5.3.4 Lisibilité

Comme Miller *et al.* (2001) l'expliquent, la lisibilité se distingue de la compréhensibilité par la rapidité avec laquelle la sortie du système peut être lue. De fait, plus un texte est facile, plus le temps requis pour le lire diminue.

5.4 Description des tests

Suite à la spécialisation du système et la pré-édition effectuée sur le texte source en français, il est nécessaire à présent d'évaluer les méthodes d'amélioration de la sortie du système à travers la qualité linguistique de celui-ci. Les tests que nous allons

présenter ci-après permettront de voir si les diverses tâches qui ont été entreprises ont apporté des améliorations, à quel point les changements ont été significatifs et si la traduction répond aux attentes de l'utilisateur final. Pour effectuer les tests, nous nous sommes appuyées sur les travaux de recherche d'ISLE MT Evaluation Workshop à l'ISSCO de l'Université de Genève ainsi que sur les articles de Dabbadie *et al.* (2002), Braun-Chen et Yuste-Rodrigo (2001) et Miller *et al.* (2001).

Nous calculerons les pourcentages selon la formule suivante :

$$(\text{nombre de points obtenus}) / (\text{nombre total de points}) \times 100$$

Nous commenterons les résultats obtenus dans le chapitre suivant.

5.4.1. Couverture lexicale : termes et vocabulaire

Les tests 1 et 2 ont été faits à partir de la sortie brute de notre corpus, c'est-à-dire la version (A_Systran).

5.4.1.1 Test 1 : Couverture lexicale – Terminologie

L'ajout des dictionnaires spécialisés est la première tâche que nous avons effectuée une fois le corpus déterminé. Pour reprendre Kittredge (1985), un sous-langage se définit également par l'usage de phrases (ou de termes techniques) fait par une « communauté » de personnes, liées par un savoir commun sur le domaine en question. La terminologie est donc un critère très important pour l'information et la traduction (Braun-Chen & Yuste-Rodrigo. 2001:4). L'importance des termes techniques et du degré de leur exactitude est une caractéristique essentielle de notre travail de pré-édition et sa valeur peut être estimée dans une évaluation subjective de la sortie du système (Braun-Chen & Yuste-Rodrigo. 2001:4). Nous voulions donc savoir jusqu'à quel point la sortie du système de Systran a été modifiée par l'ajout de nos dictionnaires.

Dans un premier temps, nous avons déterminé quels mots en langue source appartenaient au lexique spécialisé, c'est-à-dire les termes appartenant au domaine du ski, en les recherchant manuellement dans le corpus (A). Nous avons ensuite déterminé le pourcentage de termes mal traduits (Braun-Chen & Yuste-Rodrigo 2001 : 4) en les recherchant manuellement dans la sortie brute (A_Systran).

Pour avoir un aperçu des termes techniques que nous devons prendre en compte, nous avons soumis notre corpus (A) à l'outil Wordsmith Tools version 5.0, qui nous a permis de générer une liste des mots les plus fréquemment utilisés dans notre corpus. Afin d'éviter les mots les plus communs tels que les déterminants, les prépositions et autres articles de liaison, nous avons créé une stop list (voir annexe 8). Nous avons donc produit une wordlist et extrait 56 termes qui selon nous peuvent être considérés comme techniques parmi les 1'138 termes de la wordlist. Nous avons ajouté 16 termes à la liste produite car Wordsmith ne permet pas d'extraire les termes composés de plus d'un mot, par exemple *ski joëring*. Pour chaque langue, nous avons utilisé la formule suivante :

$$\frac{(\text{nombre de termes techniques mal traduits})}{(\text{nombre total de termes techniques})} \times 100 = \text{pourcentage de termes techniques mal traduits}$$

D'après les résultats, nous avons été en mesure de déduire le pourcentage de couverture lexicale des termes techniques pour chacune de nos versions (A_Systran).

5.4.1.2 Test 2 : Couverture lexicale – Vocabulaire

La couverture lexicale de nos versions (A_Systran) comprend également des entrées non-spécialisées, qu'il nous fallait comptabiliser afin d'être sûres d'évaluer la totalité de celle-ci. Nous avons donc relevé manuellement toutes les erreurs de vocabulaire des sorties brutes. Nous avons donc divisé le nombre de mots non traduits

ou mal traduits par le nombre total de mots traduits (Miller *et al.*, 2001 : 3) en utilisant la formule suivante :

$$(\text{nombre de termes inconnus ou mal traduits}) / (\text{nombre total de mots traduits}) \times 100 = \text{pourcentage de termes inconnus ou mal traduits.}$$

5.4.2 Langage contrôlé

5.4.2.1 Test 3 : efficacité des règles de contrôle

Pour ce test, nous avons évalué les règles sélectionnées (section 4.2), en adaptant le travail de Roturier (2004). Afin de pouvoir juger de leur efficacité dans le traitement par la machine, il a fallu les prendre séparément une à une. Pour cela, nous avons composé le corpus (G) de la manière suivante :

- 1) Pour chacune de nos 12 règles de contrôle (section 4.2), nous avons sélectionné dans le corpus brut (A) deux ou trois phrases en langue source qui ne sont pas conformes à la règle testée. Nous avons fait traduire ces phrases :

➤ Ces exemples sont classés sous α .

- 2) Nous avons appliqué les règles de LC aux exemples α et les avons traduites :

➤ Ces exemples sont classés sous β .

Nous avons ensuite comparé les phrases traduites classées sous β aux phrases traduites classées sous α (voir annexe 9), selon une échelle de 0 à 2 :

- **(0)** = la règle n'a aucun impact sur la traduction,
- **(1)** = la règle a un léger impact sur la traduction,
- **(2)** = la règle a un fort impact sur la traduction.

Nous avons ensuite calculé le pourcentage d'amélioration des exemples après l'application des règles.

5.4.3 Langage contrôlé

5.4.3.1 Test 4 : compréhensibilité

Pour ce test, nous nous sommes servis du corpus (H), c'est-à-dire d'une sélection de 78 phrases provenant des six versions traduites des corpus dont nous avons parlé dans la sous-section 5.2.1.

Il est important de préciser que nous n'avons pas pris les mêmes phrases d'une langue à l'autre car les problèmes de traduction ne se situaient pas forcément aux mêmes endroits, mais nous avons cependant pris des phrases que nous pensons être les plus représentatives des six versions. Nous avons demandé à deux personnes par langue d'évaluer le corpus (H) selon les consignes suivantes :

Après avoir lu la phrase, les experts devaient répondre à la question : « **La phrase est-elle compréhensible ?** » et attribuer une note en fonction du degré de compréhensibilité :

- **(0)** = non pas du tout, je ne comprends pas de quoi il est question
- **(1)** = oui, suffisamment, je comprends le sens mais j'ai eu recours à un moment de réflexion ou à une seconde lecture
- **(2)** = oui tout à fait, dès la première lecture, tout était parfaitement clair.

Pour les notes **(0)** et **(1)**, les experts devaient expliquer brièvement ce qui avait empêché la parfaite compréhension du texte. Il faut préciser que les experts n'ont pas eu accès à la langue source.

Après avoir comptabilisé le nombre de « 0 », de « 1 » et de « 2 » attribués par chaque experts, puis, à l'aide du pourcentage, nous avons produit une courbe de progression de la compréhensibilité pour chaque langue.

5.4.4 Lisibilité

5.4.4.1 Test 5 : lisibilité

En comptabilisant le nombre de (1) du test 4, nous avons pu mesurer la lisibilité du corpus (H). En effet, les experts ont eu la possibilité de signaler les phrases qui leur ont demandé le plus de temps pour arriver à en saisir le sens. Pour cela, ils devaient leur attribuer la note de 1 :

- (1) = oui, suffisamment, je comprends le sens mais j'ai eu recours à un moment de réflexion ou à une seconde lecture.

Nous pouvons donc essayer de juger de la lisibilité au vu du nombre de note 1 sur l'ensemble des phrases (Miller *et al.* 2001).

5.5 Acrolinx

Dans cette dernière partie, nous avons voulu évaluer l'apport effectif sur nos sorties d'un outil de vérification qui se base sur du LC. Pour ce faire, nous avons pris le corpus (E) (sous-section 5.2.1) pour le soumettre à l'outil et suivre manuellement les indications données pour correspondre au LC. Pour cette section, nous avons décidé de tester uniquement le critère d'évaluation suivant : la compréhensibilité (sous-section 5.3.3).

5.5.1 Acrolinx

5.5.1.1 Test 6 : compréhension

Ce test reprend le corpus (H) et le barème du test 4. Nous avons relevé les phrases correspondant aux corpus (E_Systran) et (F_Systran) (sous-section 5.2.1). Nous avons traité ce test-ci séparément du test 4 car nous voulions voir l'impact d'Acrolinx uniquement.

5.6 Conclusion

Nous avons présenté les critères que nous avons choisis d'utiliser ainsi que les mesures que nous allons appliquer aux différentes versions de notre corpus afin de déterminer l'efficacité des méthodes d'amélioration de la traduction (les dictionnaires utilisateurs et le langage contrôlé). Dans le tableau ci-dessous, nous dressons un récapitulatif des tests que nous allons effectuer :

Critères	Test(s)	Version du corpus utilisée	Experts
<i>Couverture lexicale : termes et vocabulaire</i>	<u>Test 1</u> : Calcul du pourcentage de couverture des termes techniques. <u>Test 2</u> : Calcul du pourcentage de couverture lexicale incorrecte	Corpus (A) + (A_Systran)	Nous-mêmes
<i>LC : efficacité des règles de contrôle</i>	<u>Test 3</u> : Efficacité des règles de contrôle pour le système de TA	Corpus (G)	Nous-mêmes
<i>LC : compréhension</i>	<u>Test 4</u> : Calcul du pourcentage d'amélioration de la traduction grâce au LC	Corpus (H)	Deux experts par langue
<i>Lisibilité</i>	<u>Test 5</u> : Calcul de la facilité et	Corpus (H)	Deux experts par langue

	de la rapidité de lecture		
<i>Compréhensibilité Acrolinx</i>	<u>Test 6 :</u> Calcul du pourcentage d'amélioration de la traduction grâce au LC	Corpus (H)	Deux experts par langue

Tableau 25. Récapitulatif des tests, des critères et des corpus

VI. Exécution de l'évaluation

6.1 Introduction

Ce chapitre présente l'évaluation des méthodes d'améliorations de la traduction, en reprenant dans l'ordre chacun des critères que nous avons évalué. Nous donnerons nos résultats sous forme de tableaux et/ou de graphiques après avoir brièvement expliqué notre méthodologie.

6.2 Résultats

6.2.1 Couverture lexicale : termes et vocabulaire

6.2.1.1 Test 1 : Couverture lexicale - Terminologie

Nous avons établi une liste de 72 termes qui correspondent au 100% de couverture technique du texte. Pour l'anglais, 52 termes mal traduits ont été entrés dans le dictionnaire ; pour l'italien, 50. Nous sommes arrivées à un total d'environ 70 % de couverture lexicale incorrecte des termes techniques pour chacune des langues:

	Anglais	Italien
Nombre de termes techniques mal traduits	52/72	51/72
Pourcentage de termes techniques mal traduits	72.22%	70.83%

Tableau 26. Couverture lexicale technique du corpus

Le système traduit tout de même plus ou moins 30 % des termes techniques de manière correcte, ce qui est tout à fait acceptable. Au vu des ces résultats, il va sans dire que

l'ajout d'un dictionnaire est essentiel pour retrouver le même degré de fidélité technique dans la sortie que dans le texte source.

6.2.1.2 Test 2 : Couverture lexicale - Vocabulaire

Le corpus source est composé de 10'200 mots, dont 1'176 occurrences uniques, qui correspondent au 100% de couverture lexicale du texte. Pour l'anglais, nous avons comptabilisé 271 occurrences de mots inconnus ou mal traduits. Nous avons relevé pour l'italien 297 occurrences de mots inconnus ou mal traduits. Après calculs, nous sommes arrivées à un total de 23,04% de couverture lexicale incorrecte pour l'anglais et de 25.25% pour l'italien :

	Anglais	Italien
Nombre d'occurrences de mots mal traduits ou inconnus	271/1'176	297/1'176
Pourcentage de couverture lexicale incorrecte	23,04%	25,25%

Tableau 27. Couverture lexicale incorrecte du corpus

Les résultats montrent que plus de la moitié du texte est traduite correctement au niveau du vocabulaire, prouvant que Systran est un système plutôt performant pour notre type de texte. Toutefois, nous devons souligner que ces résultats concernent la couverture lexicale et non pas la syntaxe et le style. En effet, afin d'obtenir une évaluation plus complète, nous devons prendre en compte tous les critères exposés dans la section 5.3. La manque de fidélité lexicale est un problème au niveau de la compréhensibilité car il déstabilise le lecteur, qui s'attend à trouver un certain vocabulaire.

6.2.2 Langage contrôlé

6.2.2.1 Test 3 : efficacité des règles de contrôle

Le présent test détermine l'efficacité des règles de contrôle sur les sorties du système, lorsque nous les appliquons au texte source. La notation concernera uniquement les 28 phrases traduites du corpus (G) énumérées sous β en comparaison avec les phrases traduites α . Chaque phrase peut atteindre la note maximum de 2 points, évaluant le degré d'amélioration. Étant donné que nous avons 28 exemples, nous avons calculé notre barème sur le nombre total de points potentiellement atteignables, qui est de 56. Ce nombre équivaut au 100% d'amélioration de la sortie du système par le LC.

	Anglais	Italien
Note obtenue	37/56	40/56
Pourcentage d'amélioration	66,07%	71,42%

Tableau 28. Pourcentage d'amélioration de la sortie du système par le LC

Nous constatons que les résultats pour les deux langues sont très proches et que le taux d'amélioration est important. Il est cependant utile de mentionner que les phrases β pour lesquelles nous avons appliqué les règles iii et iv (sous-section 4.3.1) n'ont pas reçu de points. En effet, utiliser la même phrase pour le même contenu (iii) et le même terme à chaque fois (iv) présuppose que nous ayons choisi une construction de phrase ou un terme dans le corpus (A) que nous avons appliqué au corpus contrôlé (C). Ainsi, le quota d'amélioration pour ces phrases retombe à zéro. Malgré cela, le tableau nous montre clairement que l'ajout des règles sélectionnées aide à l'amélioration de la sortie

du système. Toutefois, nous pensons que celles-ci seraient plus utiles si elles étaient appliquées au moment de la rédaction.

6.2.3 Langage contrôlé

6.2.3.1 Test 4 : compréhension

Chaque partie contient 16 phrases, qui peuvent atteindre la note maximum de 2. Comme nous l'avons fait auparavant, nous avons multiplié le nombre de phrases par le nombre de points maximum potentiellement atteignables. Le total est de 32 points, correspondant à 100% d'amélioration. Nous avons compté le nombre de points que les experts ont attribués à chaque test, puis nous avons reporté les notes dans un tableau, un pour l'italien, et un pour l'anglais :

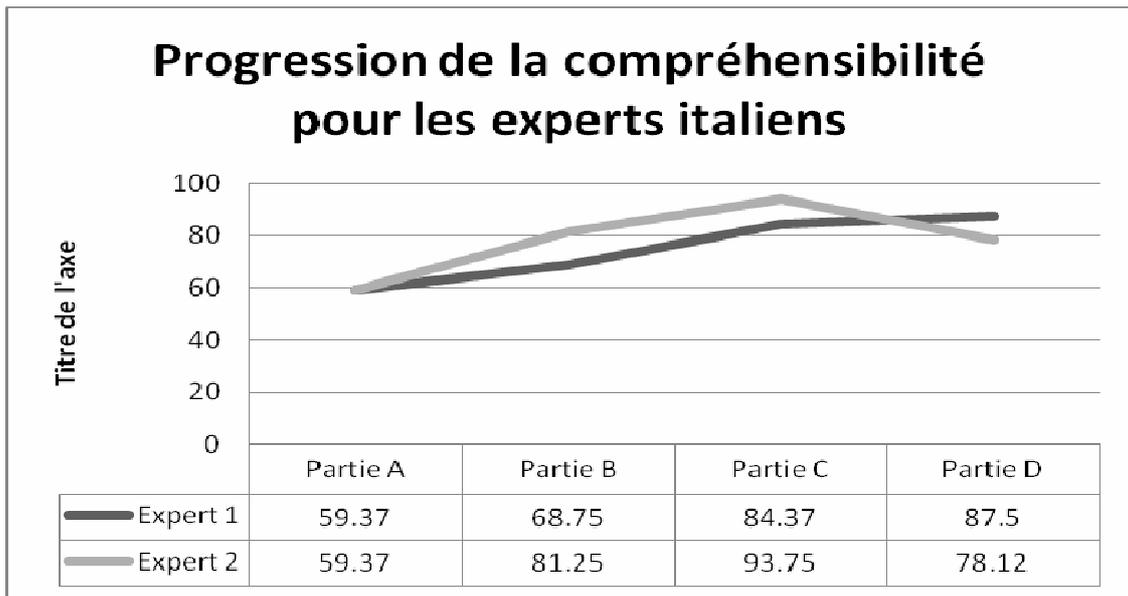
Experts		1	2
Corpus H	Partie A	19/32	19/32
	Partie B	22/32	26/32
	Partie C	27/32	30/32
	Partie D	28/32	25/32

Tableau 29. Résultats pour l'italien

Experts		1	2
Corpus H	Partie A	15/32	15/32
	Partie B	21/32	13/32
	Partie C	23/32	13/32
	Partie D	26/32	24/32

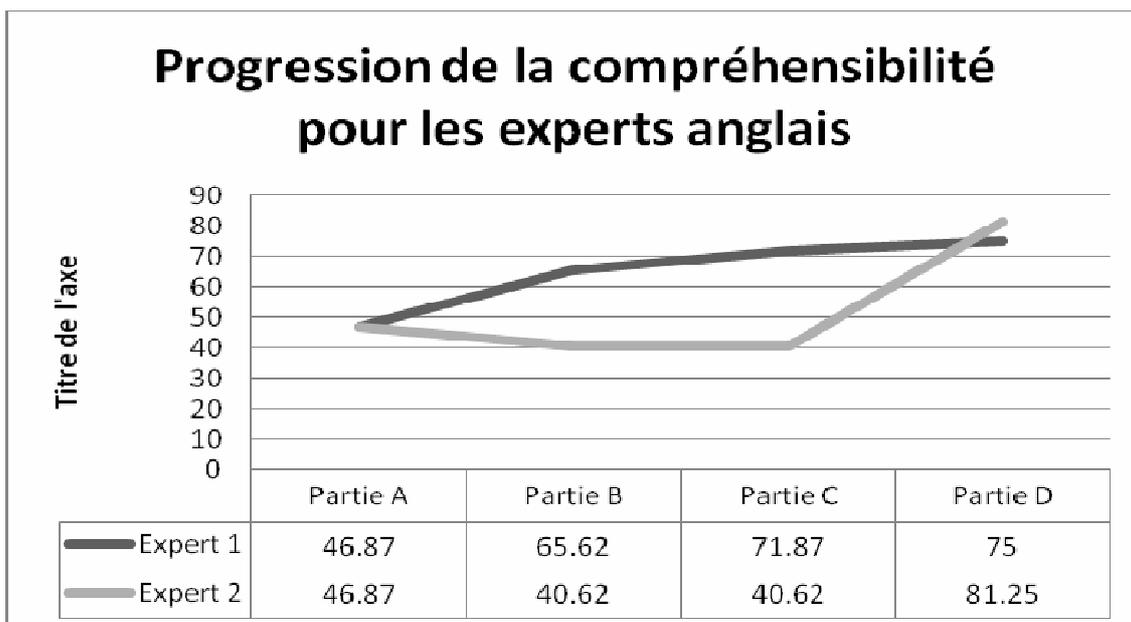
Tableau 30. Résultats pour l'anglais

Comme nous pouvons le voir, les résultats sont cohérents avec nos prévisions : les notes augmentent au fur et à mesure que la sortie du système est modifiée par l'ajout des dictionnaires et du LC. Nous pouvons voir très clairement la progression de la compréhension dans les graphiques 1 et 2 ci-dessous :



Graphique 1. Progression de la compréhensibilité des experts italiens.

Il est intéressant de remarquer que la compréhensibilité de l'expert n°1 italien suit une progression linéaire : les résultats augmentent au fur et à mesure que les parties « s'améliorent » (spécialisation des dictionnaire et ajout du LC). En revanche, il est plus surprenant de constater que pour l'expert n°2 italien, la partie C (reprenant le corpus C_Systran) soit la meilleure.



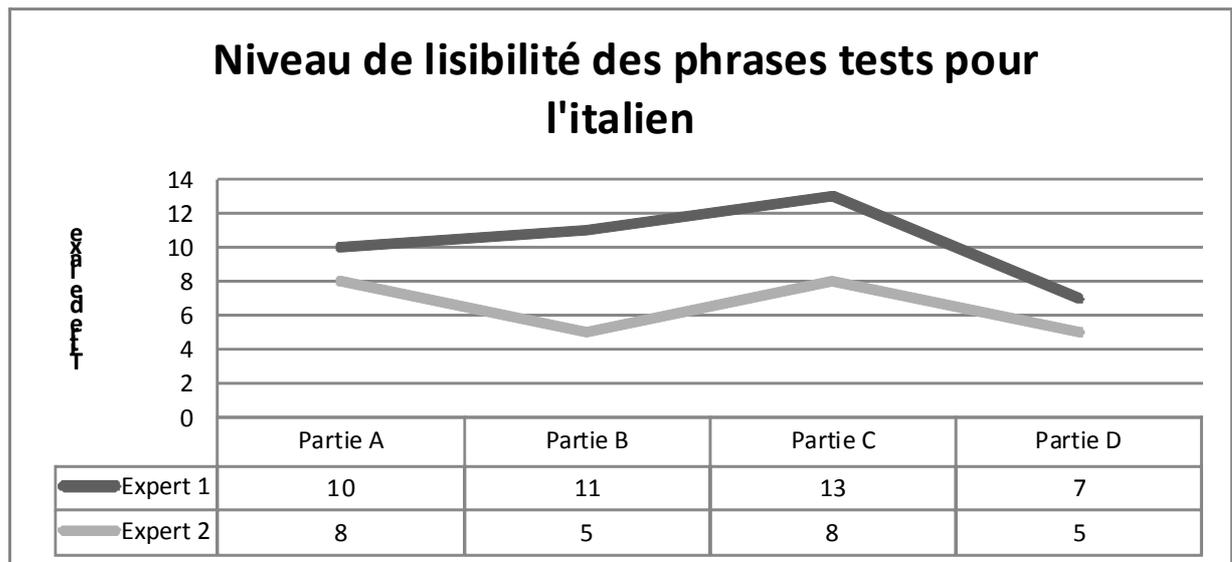
Graphique 2. Progression de la compréhension des experts anglais.

Comme pour l'italien, la courbe de progression de l'expert n°1 anglais est linéaire, ce qui s'avère cohérent avec nos attentes. Étonnamment, l'expert n°2 anglais ne semble pas noter d'amélioration entre les parties B et C, et considère qu'elles sont plus mauvaises que la partie A.

6.2.4 Lisibilité

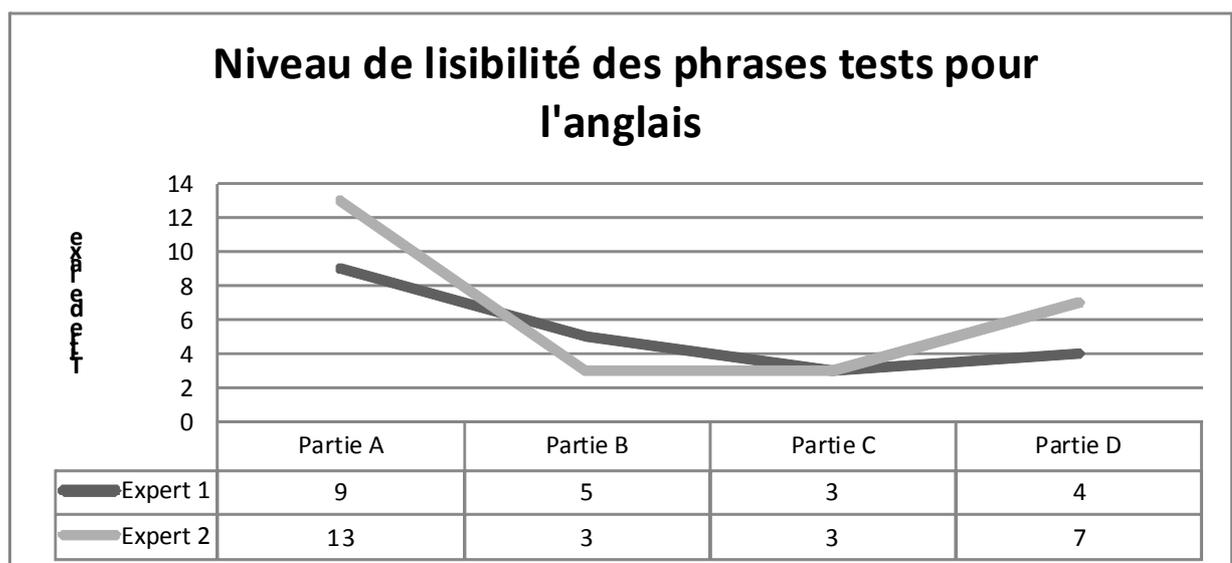
6.2.4.1 Test 5 : lisibilité

Nous avons comptabilisé le nombre de note « 1 » que les experts ont attribué à chaque corpus. Dans les deux langues, nous pouvons voir que le nombre varie selon les versions du corpus : il y a effectivement une progression (dès que l'on ajoute les dictionnaires, le nombre de « 1 » diminue) mais elle n'est pas linéaire. Les graphiques 3 et 4 ci-dessous le prouvent :



Graphique 3. Niveau de lisibilité des phrases tests pour l'italien.

En effet, nous pouvons observer dans le graphique 3 que la lisibilité de la partie C (version contrôlée sans les dictionnaires) n'est pas évidente pour les sujets italophones puisque la courbe remonte légèrement avant de descendre à son niveau le plus bas. Nous pouvons donc constater que l'ajout des dictionnaires semble avoir un impact plus important sur la compréhensibilité des experts italiens que l'application du LC.



Graphique 4. Niveau de lisibilité des phrases tests pour l'anglais.

Chose curieuse, les experts anglais considèrent la partie C comme la plus lisible, au vu de la courbe. Nous pouvons en déduire que le LC semble être mieux perçu par les sujets anglophones que la spécialisation des dictionnaires. Il est intéressant de noter que la partie D, combinant le LC + les dictionnaires, n'obtient pas de très bons résultats de la part de l'expert n°2 anglais.

Il est intéressant de remarquer que la lisibilité des experts italiens suit une courbe opposée à celle des experts anglais dans les deux graphiques ci-dessus. Nous avons constaté que la sensibilité divergeait : les italophones répondent mieux aux parties contenant les dictionnaires alors que les anglophones semblent privilégier le LC. Cette divergence peut s'expliquer par la nature différente des langues évaluées, mais aussi par la subjectivité des experts face aux erreurs.

6.2.5 Acrolinx

6.2.5.1 Test 6 : Compréhensibilité

La sélection de phrases traduites du corpus (H) provient des corpus (E_Systran) et (F_Systran), composés de 1'522 mots. Les parties E et F contiennent donc 7 phrases au lieu des 16 des parties précédentes. La note maximum potentielle étant également de 2, le 100% d'amélioration correspond à 14 points :

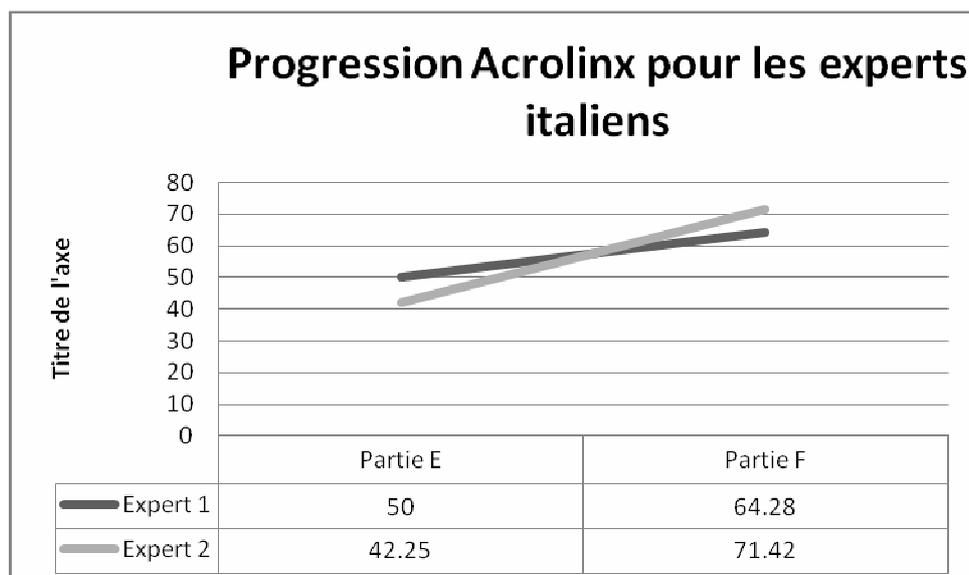
Experts		1	2
Corpus H	Partie E	7/14	6/14
	Partie F	9/14	10/14

Tableau 31. Résultats pour l'italien

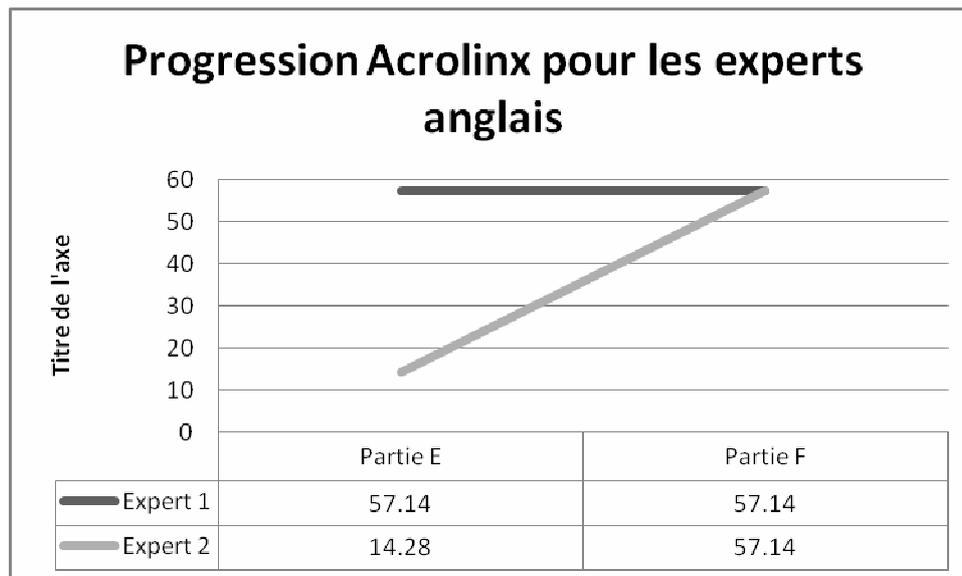
Experts		1	2
Corpus H	Partie E	8/14	2/14
	Partie F	8/14	8/14

Tableau 32. Résultats pour l'anglais

Étonnamment, l'expert n°1 anglais a attribué le même nombre de points aux parties E et F, qui correspondent à un peu plus de 55% d'intelligibilité. L'expert n°2 anglais ainsi que les deux experts italiens ont remarqué une amélioration significative de la sortie du système contrôlée par Acrolinx. Ainsi, dans les commentaires émis par les experts, ce sont les termes mal traduits qui sont pointés du doigt pour expliquer la mauvaise compréhension. Les graphiques 5 et 6 ci-dessous illustrent le phénomène :



Graphique 5. Progression de la compréhension des experts italiens avec Acrolinx.



Graphique 6. Progression de la compréhension des experts anglais avec Acrolinx.

6.3 Conclusion

Au terme de cette évaluation, nous pouvons voir que les méthodes d'amélioration pour la TA ont prouvé leur efficacité. Les différents tests nous ont permis de répondre adéquatement aux questions que nous avons posées dans l'introduction de ce chapitre.

En ce qui concerne l'ajout des dictionnaires utilisateurs, nous avons observé plusieurs choses. Globalement, la sortie du système brute, que ce soit pour l'anglais ou l'italien, n'était pas si mauvaise au départ. Bien que notre corpus traite d'un domaine particulier, le vocabulaire utilisé est suffisamment général pour que le système obtienne de bons résultats. Toutefois, comme nous l'avons dit, il convient de nuancer ces résultats, puisque le texte est relativement répétitif.

Pourtant, notre corpus traite d'un domaine particulier nécessitant une terminologie spécifique. Ces termes techniques n'étant pas disponibles dans le dictionnaire général du système, la sortie du système, bien que relativement compréhensible au niveau de la grammaire et du style, ne remplit pas les exigences de fidélité à la terminologie. Par

ailleurs, lors du dépouillement du questionnaire de compréhensibilité (voir sous-section 6.2.3 et 6.2.4), nous avons constaté que les experts notaient plus favorablement les phrases contenant les dictionnaires, que ce soit dans la version brute ou contrôlée. En conséquence, la création des dictionnaires utilisateurs a démontré de manière convaincante sa pertinence dans l'amélioration de la traduction automatique.

Quant au LC, nous avons pu voir que le remaniement des phrases pour qu'elles correspondent aux règles de contrôle que nous avons choisies s'est révélé nécessaire à la compréhensibilité du texte source pour la machine et du texte cible pour l'être humain. Si on se réfère au tableau 28, en effet, nous pouvons voir que le pourcentage d'amélioration de la sortie du système atteint environ 65% d'amélioration pour l'anglais contre environ 70% pour l'italien. De même, les graphiques 1 et 2 attestent de l'amélioration de la traduction à l'aide du LC : si on compare les résultats obtenus pour la partie A à ceux obtenus pour la partie C, l'augmentation de la courbe de niveau indique que les experts ont donné une meilleure note de compréhensibilité à la version contrôlée. Par conséquent, le taux élevé de pourcentage et les résultats du graphique prouvent de manière indéniable que l'application de règles de contrôle optimise la compréhension du texte par le système de TA, produisant ainsi une sortie du système suffisamment bonne pour être post-éditée.

Comme nous venons de le montrer, l'une et l'autre méthode ont fait leurs preuves. En analysant les résultats du graphique n°1 tant pour l'italien que pour l'anglais, nous pouvons voir que la courbe de progression de la compréhensibilité atteint son apogée pour la partie D, laquelle correspond à la version contrôlée avec dictionnaire de notre corpus.

En ce qui concerne Acrolinx, les tests que nous avons effectués ne se sont pas montrés aussi efficaces que nous l'avions espéré. Les règles présentes dans l'outil sont trop

générales, il faudrait pouvoir personnaliser Acrolinx afin que les règles correspondent au mieux aux types de textes utilisés. Comme la version à laquelle nous avons accès ne nous permettait pas d'ajuster l'outil à notre domaine, les résultats, bien qu'acceptables, ne sont pas suffisamment performants pour que nous puissions utiliser Acrolinx indépendamment des autres méthodes d'amélioration de la traduction que nous avons testé dans ce chapitre.

VII. Post-édition : le cas italien

7.1 Introduction

Dans cette partie, nous nous concentrons sur le travail de post-édition du texte cible pour la langue italienne. Même si nous avons déjà pu constater une amélioration effective de la sortie du système grâce au LC et aux dictionnaires utilisateurs, quelques erreurs subsistent encore. Ces erreurs sont intéressantes dans la mesure où elles nous permettent :

- d'évaluer les performances d'un système de traduction avant un investissement éventuel ;
- de déterminer l'effort de post-édition nécessaire selon un type de texte pour une paire de langue donnée (l'italien).

Nous allons donc expliquer en quoi consiste la post-édition (section 7.2), établir des critères d'évaluation par rapport à la qualité de la sortie du système (section 7.3), proposer une typologie des erreurs ainsi qu'un classement (section 7.4) et enfin vérifier une nouvelle fois l'impact du langage contrôlé (section 7.5).

7.2 La post-édition (PE)

7.2.1 Introduction

La traduction automatique a changé l'environnement du traducteur en proposant au fil des années des systèmes de plus en plus performants, afin d'obtenir un texte acceptable. Toutefois, les sorties du système sont loin d'être exemptes d'erreurs ce qui a conduit à l'arrivée d'un nouvel acteur dans le processus de traduction : le post-éditeur.

7.2.2 Qu'est-ce que la post-édition ?

Contrairement au travail de pré-édition qui modifie l'entrée du système afin de faciliter le processus de traduction automatique, la tâche de post-édition s'effectue une fois que la traduction par la machine a été opérée. Elle a été précédemment définie comme une correction de la sortie du système par des linguistes ou des réviseurs (Allen, 2003 ; Veale and Way, 1997) ou comme le fait de réviser, modifier et/ou corriger un texte pré-traduit par une machine de traduction d'une langue source à une langue cible (Allen, 2003 : 297).

A première vue, il semblerait que le rôle de post-éditeur ne se différencie pas du rôle classique du réviseur. Il faut toutefois mettre en garde contre une superposition de ces deux rôles. Face à une sortie de système brute. Le réviseur corrige comme s'il s'agissait d'une traduction humaine, c'est-à-dire avec un certain nombre de critères standardisés et sensiblement semblables à ceux des professeurs (Loffler-Laurian, 1996 : 82-83). Le post-éditeur revoit une sortie du système brute selon des consignes qui lui ont été transmises. Ces consignes peuvent être codifiées. De plus en plus d'agences de traduction ou de localisation ont en effet créé leur propre guide de post-édition (PE) répertoriant les critères à suivre. Par ailleurs, certains ont vu cette technique de modification comme un processus faisant partie du langage contrôlé puisque les changements effectués sur les sorties du système ressemblent à ceux faits sur les entrées du système, c'est-à-dire selon des règles de contrôle (Allen & Hogan, 2000). Lorsque Loffler-Laurian (1996) parle de la post-édition comme d'une adaptation à un style pour un même type de document, on voit le parallèle avec les règles de cohérence stylistique du LC.

Cependant, il n'existe pas de manuel universel et peu de documents de PE disponibles au grand public comme pour le LC. Il faut toutefois noter que durant les années 2000, de nombreuses entreprises se sont lancées dans l'activité de PE et n'ont donc pas encore rendus publics leurs travaux (Allen, 2003 : 307). Malgré tout, quelques guides de PE sont disponibles⁵. De même, sur le Net, le projet de recherche TAUS propose à la

⁵ Citons General Motors, Pan-American Health Organization (PAHO), European Commission Translation Service (ECTS), Post-Editing Special Interest Group ainsi que les travaux d'Anne-Marie Loffler-Laurian auprès de la Commission Européenne.

vente des guides de PE dont quelques pistes sont ouvertes au grand public. Le guide du projet GALE est également disponible online (2007).

Somme toute, la PE est surtout une activité qui a débuté par quelques règles évoluant au fil des expériences des traducteurs qui se sont lancés dans cette tâche. L'entraînement à la pratique de PE et la compilation de consignes claires sont autant d'éléments nécessaires au bon rendement de cette activité (Guzman, 2007).

7.2.3 Types de post-édition

Le travail de post-édition, ou plutôt l'ampleur de ce travail, a été l'objet de bon nombre de discussions, notamment en ce qui concerne de l'effort à fournir pour obtenir un produit viable (Allen, 2003 : 298). Les facteurs déterminant le niveau de PE auquel est soumis un texte découlent en fait de tout le processus de production des documents. En effet, l'intégration de traduction automatique dans le flux de travail répond souvent à une forte demande de globalisation rapide et efficace. De même, les attentes concernant la qualité du produit peuvent également être influencées par le souhait d'une information rapide et multilingue comme elle est dorénavant proposée sur Internet au travers des portails de traduction gratuit (Allen, 2003 : 300). Un document est ainsi produit pour être compris superficiellement ou alors pour communiquer un message correct au plus grand nombre. Le degré de PE dépend donc :

- des requêtes de l'utilisateur,
- du volume de production,
- des attentes de qualité,
- de la péremption de l'information,
- de la fonction du texte (Allen, 2003 ; Loffler-Laurian, 1996 ; O'Brien, 2010).

La PE doit se situer dans ce contexte de consommation du document traduit et c'est pourquoi le traducteur se retrouve face à deux types d'activités de PE qui dépendent de l'utilisation future de la traduction (Loffler-Laurian, 1996) : la PE rapide et la PE conventionnelle, dont nous allons brièvement parler maintenant.

7.2.3.1 Post-édition rapide (PER)

La post-édition rapide se définit comme un minimum de modifications effectuées sur des documents contenant des informations d'une durée de vie courte et souvent de domaine non-public (Allen, 2003). L'objectif de cette PE est de rendre le texte suffisamment compréhensible et clair sans pour autant être stylistiquement irréprochable. De fait, la structure des phrases et l'ordre des mots sont tout à fait secondaires dans la mesure où ils n'ont pas d'impact négatif sur l'information véhiculée (Loffler-Laurian, 1996 : 93). Au vu du volume toujours plus grand de documents, de plus en plus d'agences visent à ce type de PE.

7.2.3.2 Post-édition conventionnelle (PEC)

Par post-édition conventionnelle, nous entendons que les changements effectués sont les plus nombreux possibles afin de parvenir à un haut niveau de qualité pour des documents ciblés pour la grande distribution (Allen, 2003). La PEC répond en fait aux attentes de l'utilisateur qui veut retrouver dans le texte cible une terminologie adéquate et un style fluide et créatif comme on l'attendrait d'une traduction humaine. Ce type de PE soulève, malgré tout, quelques questions. En effet, dans un contexte de consommation rapide et de distribution à grande échelle, est-il possible de pratiquer une PE totale qui demande presque autant de temps qu'une traduction humaine ? Quelles sont alors les limites de la PE à ne pas dépasser pour être rentable en termes de coût et de temps ?

7.2.3.3 Vers une post-édition personnelle

Dans ce processus de traduction et de post-édition, les divers auteurs que nous avons pu consulter sont tous d'avis que l'expérience et les réactions personnelles aux sorties du système et à leurs corrections ont abouti à établir des critères de PE selon le

document traité. Nous avons donc décidé de nous confronter nous-mêmes à cet effort de post-édition en produisant une traduction acceptable et correcte.

Nous avons donc tout d'abord repris les deux sorties pour la langue italienne produites précédemment (sous-section 5.2.1):

(A_Systran) brute sans dictionnaire

(D_Systran) contrôlé avec dictionnaires.

Pour des raisons pratiques, nous avons décidé de créer un mini-corpus composé de phrases du corpus (A_Systran), appelé (I). Les phrases choisies sont majoritairement extraites des parties descriptives (section 3.1) puisque les autres parties (indication et recommandations) contenaient uniquement des phrases qui se répétaient dans chaque cellule. Néanmoins nous avons sélectionné quelques phrases représentatives de ces dernières. Notre corpus (I) est donc composé de 166 phrases et de 3'727 mots. A partir du corpus (I), nous avons produit une traduction post-éditée appelée (K) (annexe 14). Cette dernière a été faite en suivant les principes de la PE rapide, au fur et à mesure que les erreurs étaient relevées, et par la suite elle a été vérifiée par une personne italophone que nous nommerons au long de ce chapitre l'expert en post-édition.

Enfin, nous avons créé un second mini-corpus composé des phrases du corpus (D_Systran), appelé (J). Ce corpus est composé de 225 phrases et de 3'946 mots. Il va s'en dire que les phrases du corpus (J) sont les équivalents des phrases du corpus (I). A partir du corpus (J), nous avons produit une traduction post-éditée appelée (L) (annexe 15). Le texte post-édité (K) sera par la suite comparé avec la post-édition du texte (L) afin de vérifier l'impact du LC.

Nous avons fait un récapitulatif des différents corpus utilisés dans ce chapitre :

Corpus entier	Mini-corpus	Sorties post-éditées
(A_Systran)	(I)	(K)
(D_Systran)	(J)	(L)

Tableau 33. Récapitulatif des corpus.

Pour le moment, nous allons nous concentrer sur la sortie du système (I), sur les erreurs qui ont été produites par la machine et sur les mesures d'évaluation de ces dernières.

7.3 Critères d'évaluation

Démontrer la qualité de la sortie du système consiste également à évaluer les performances de la MT choisie, dans notre cas Systran. Une bonne sortie du système est le gage d'une post-édition dont l'effort et surtout les coûts sont moindres. Pour notre travail, nous avons entrepris des mesures d'évaluation manuelle (section 5.3). Cette dernière a l'avantage de prendre en compte l'avis subjectif de l'évaluateur (Federmann, 2011) et d'établir une séparation claire des erreurs (Bircher, 2009 : 60)

Dans la suite, nous présentons donc les critères de qualité retenus pour évaluer nos sorties de systèmes et leurs erreurs.

7.3.1 Critères de qualité

Les machines de traduction sont utilisées comme un outil d'aide à la production rapide de textes traduits. Les attentes continuent de se concentrer sur une traduction par la machine post-éditée aussi acceptable qu'une traduction humaine. Or la qualité des sorties de systèmes peut varier selon le système de traduction, les paires de langues, le type de texte et le degré de contrôle de l'entrée de système (O'Brien, 2010).

Pour évaluer la qualité de notre sortie de système (I), nous avons choisi deux mesures pertinentes pour mesurer le degré de qualité de la sortie du système et de la MT. Nous avons repris les critères proposées par Bircher (2009) et Flanagan (1994) : l'intelligibilité et le degré d'amélioration.

7.3.1.1 Intelligibilité

Par le critère d'intelligibilité, nous entendons que la sortie du système est grammaticalement acceptable et suffisamment compréhensible pour véhiculer les informations principales et obtenir un sens équivalent à celui du texte source. Les erreurs peuvent avoir un impact négatif sur l'intelligibilité du texte et le rendre provisoirement inutilisable dans l'attente d'une post-édition. Cela a son importance dans l'évaluation d'une MT, surtout si on veut utiliser et diffuser les textes sans post-édition (Flanagan, 1994 : 69). Comme nous le verrons dans l'évaluation, en général, il s'agit d'erreurs grammaticales graves, de mots manquants ou non traduits, etc. (Bircher, 2009 : 33).

7.3.1.2 Degré d'amélioration

Par le critère de degré d'amélioration, nous entendons que la sortie du système est facilement post-éditable. Toute modification demande un effort et un coût qu'il s'agira d'évaluer au moment de l'investissement dans un système de TA (Flanagan, 1994 : 69). Selon le type d'erreurs que nous trouvons dans notre document, nous pourrons évaluer le type d'amélioration possible pour nos corpus.

Dans la section suivante, nous allons procéder à l'évaluation de la sortie du système (I) pour la langue italienne, selon ces deux critères, après avoir établi un tableau des erreurs de notre corpus.

7.4 Evaluation des erreurs

Pour pouvoir mesurer la qualité de la sortie du système, nous nous sommes concentrées sur les erreurs produites. Nous avons relevé manuellement toutes les erreurs présentes (annexe 14) et réalisé une typologie de ces dernières sous forme de tableau (sous-section 7.4.1). Par la suite, nous avons choisi de traiter les erreurs par catégorie (sous-section 7.4.2) suivant la méthode de classement de Flanagan (1994), ce

qui nous permettra par la suite de mesurer l'intelligibilité et le degré d'amélioration de la sortie (I) (sous-section 7.4.2.1/3).

7.4.1 Tableau des erreurs

La catégorisation des erreurs permet de voir le type de fautes qu'un système de traduction produit selon le type de texte et la paire de langue. Pour obtenir cet ensemble de catégories, il faut observer les fautes commises dans la sortie du système (I), les comptabiliser (tableau 34) ce qui permettrait dans le cadre d'un autre travail de comparer les résultats avec ceux d'un autre système de traduction.

Pour ce qui est du décompte, nous signalons que les catégories concernant le vocabulaire et les noms propres ont été scindées en deux sous-catégories : pour le vocabulaire, nous avons repris la distinction entre mots inconnus et mots mal traduits (section 4.2); pour les noms propres, nous avons séparé les mots inconnus des traductions littérales. Pour les autres catégories de fautes, aucune distinction n'a été faite.

Catégories de fautes		Description	Nombres d'erreurs dans le corpus (I)
1.	Vocabulaire /terminologie	Mots mal traduits	224
		Mots inconnus	112
2.	Noms propres	Mots inconnus	139
		Traduction littérale	45
3.	Préposition	Préposition incorrecte, absente ou non-nécessaire	47
4.	Casse	Extraits de phrases/mots effacés ou non traduits (alors que le mot existe dans le dictionnaire général)	44
5.	Expression	Traduction incorrecte d'une expression à plusieurs mots	38
6.	Catégorie grammaticale	Erreur de catégorie (ex : nom vs verbe)	30
7.	Accord	Relation incorrecte entre sujet-verbe/sujet-adjectif/flexion du nom	24
8.	Ponctuation	Majuscules, orthographe	22
9.	Articles	Déterminant incorrect, absent ou non nécessaire	21
10.	Ordre des mots	Ordre des mots, position des syntagmes ou structure incorrect	15
11.	Pronom personnel	Pronom incorrect, absent ou non nécessaire	14
12.	Périphrase	Remplacer un mot par une expression ou définition plus	12

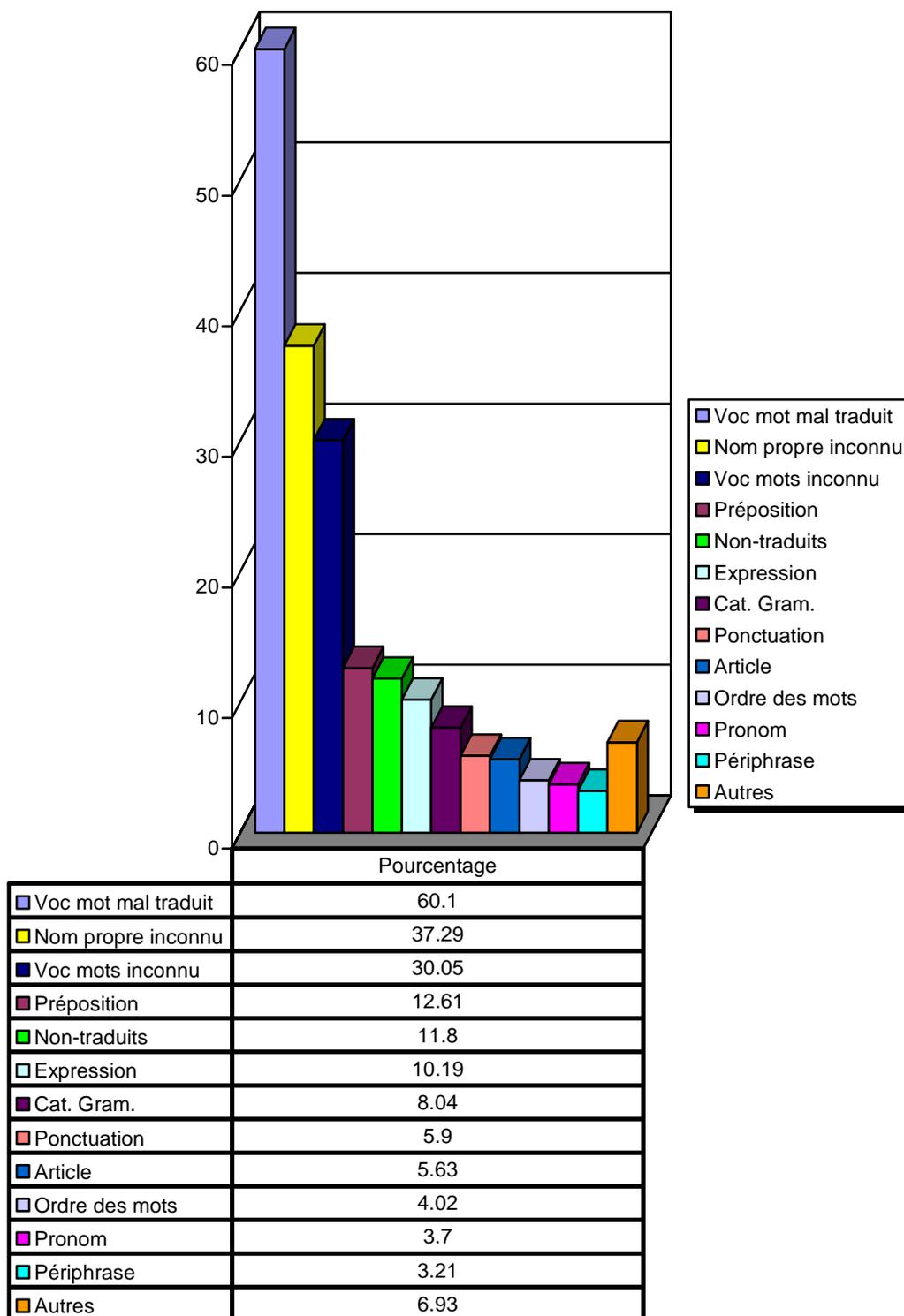
		longue	
13.	Chiffre	retranscription incorrecte des dates et heures	2
Nombre total d'erreurs			789

Tableau 34. Tableau des erreurs de la sortie du système (I).

Les résultats obtenus dans le tableau 34 peuvent être visualisés sous forme de pourcentage dans le graphique 7. Le calcul des pourcentages s'est fait selon le calcul suivant :

$$(\text{nombre d'erreurs}) / (\text{nombre de mots total}) * 1000 = \text{pourcentage d'erreur de la catégorie x}$$

Pour information, sur les 166 phrases du corpus (I), seules 4 phrases étaient exemptes de toute erreur.



Graphique 7. Pourcentage d'erreurs du corpus (I).

Les mots inconnus au système sont évidemment ceux qui posent le plus de problème, représentant près de 67% du texte global. Toutefois, nous pouvons penser que ceux faisant partie de la catégorie « Nom propre » ne présentent pas de frein à la compréhension s'ils restent non traduits et en français. Seuls quelques cas de lieux très connus ont nécessité une traduction : ex. le Mont-Blanc > il Monte Bianco ou la Forêt Blanche > la foresta bianca (reconnu par le système mais sans majuscules !). Au contraire, la mauvaise traduction de la plupart du vocabulaire (mots mal traduits) représente un frein indéniable à la compréhension ce que nous avons pu voir dans les résultats du test 4 (sous-section 6.2.3). Les prépositions ont entraîné également beaucoup de fautes dans les phrases, que ce soit dans leur forme simple ou contractée (préposition + déterminant). De même, les expressions se situent en bonne place, notamment parce qu'elles se forment avec des prépositions (ex : le plus beau du monde > il più bello *del* mondo > il più bello al mondo). Le système a produit également un taux élevé de bruit, supprimant ainsi des phrases entières, ce qui avait été peu constaté sur les sorties anglaises du texte source

Il sera intéressant de comparer ces résultats avec ceux que nous obtiendrons pour la sortie du système (J) avec dictionnaires et langage contrôlé (section 7.5). En attendant, nous allons établir le classement de ces erreurs selon les critères choisis pour connaître leur impact sur l'intelligibilité (sous-section 7.4.2.1/2) et sur le degré possible d'amélioration (sous-section 7.4.2.3/4).

7.4.2 Classement des erreurs

Classer les erreurs fait partie de la seconde phase de notre travail sur la sortie du système. Après avoir relevé toutes les fautes trouvées dans notre corpus (I), nous avons établi un classement des erreurs en fonction de leur gravité afin de juger la sortie du système selon les critères établis, à savoir l'intelligibilité et le degré d'amélioration. Il s'agit d'un processus manuel assez subjectif. Toutefois, les classements d'erreurs sont facilement adaptables puisqu'on peut y ajouter des catégories et en supprimer d'autres au fur et à mesure qu'on soumet des textes divers à des systèmes de TA divers selon les paires de langues (Flanagan, 1994). Le classement peut donc évoluer dans le temps.

Ainsi, nous avons établi une échelle qui qualifie le degré de gravité des erreurs (Green, 1982), selon qu'elles altèrent ou non le sens véhiculé dans le texte source (critère d'intelligibilité) et selon qu'elles soient facilement ou non corrigées par une intervention sur la machine ou de post-édition (critère de degré d'amélioration).

7.4.2.1 Application des mesures d'intelligibilité

Pour notre classement des erreurs jugées selon le critère d'intelligibilité (tableau 35), nous avons pris en compte l'impact que les fautes avaient sur l'intelligibilité de l'information véhiculée et nous avons établi l'échelle suivante :

- La classe 1 regroupe toutes les erreurs qui n'altèrent en rien le sens du texte car l'information principale est correctement véhiculée. Par exemple, une erreur de déterminant n'empêche pas la compréhension du texte global (ex : *un* snowpark > *uno* snowpark). Cette classe est donc considérée sans effet majeur.
- La classe 2 regroupe toutes les erreurs qui altèrent l'intelligibilité en proposant de solutions grammaticalement incorrectes ou des sens proches du mot exact mais l'information principale peut être aisément retrouvée à partir du contexte de la phrase (ex : *senza simile* > *senza eguali*). Cette classe est donc considérée comme acceptable.
- La classe 3 regroupe toutes les erreurs qui altèrent complètement l'intelligibilité du texte, en créant du non-sens et en ne véhiculant pas l'information principale (ex : *mutato il livello* > *dislivello*). Cette classe est donc considérée comme inacceptable.

Classe 1	Classe 2	Classe 3
Ponctuation	Vocabulaire (mots mal traduits)	Nom Propre (Traduction littérale)
Nom propre (mots inconnus)	Expression	Catégorie grammaticale
Article	Chiffre	Casse
Accord	Préposition	Vocabulaire (mots inconnus)
Ordre des mots (13)		Périphrase
		Pronom
		Ordre des mots (2)

Tableau 35. Classement des erreurs pour le critère d'intelligibilité.

En ce qui concerne la catégorie d'erreur « ordre des mots », celle-ci apparaît dans le tableau 35 deux fois car 2 fautes (structure de la phrase) ont été jugées plus graves que les autres (position de l'adjectif).

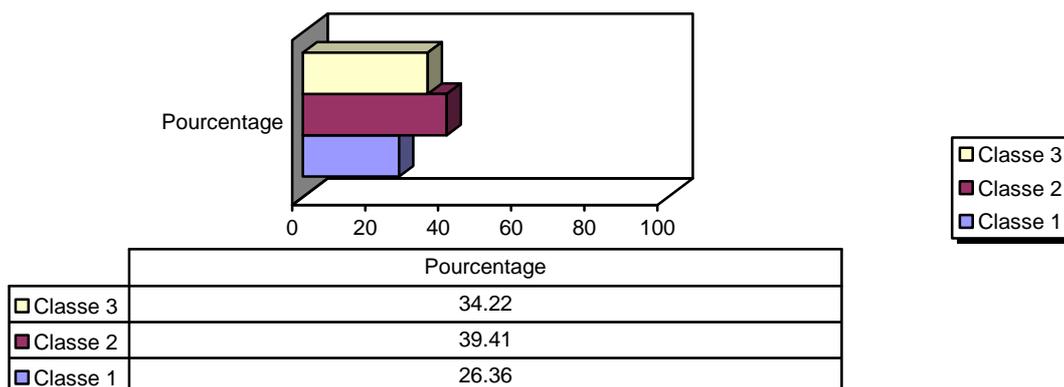
7.4.2.2 Résultats du classement des erreurs pour le critère d'intelligibilité

Les résultats obtenus dans le tableau 35 peuvent être visualisés sous forme de pourcentage dans le graphique 8. Le calcul des pourcentages s'est fait selon le calcul suivant :

$$(\text{addition du nombre d'erreurs par catégorie}) / (\text{nombre d'erreurs total}) *$$

$$100 = \text{pourcentage d'erreur de la classe } x$$

Dans l'idéal, pour obtenir une traduction acceptablement intelligible, sans aucun impact des erreurs, il faudrait que le taux d'erreurs de la classe 1 soit le plus élevé, c'est-à-dire 100%.



Graphique 8. Pourcentages du classement des erreurs selon le critère d'intelligibilité.

A première vue, nous obtenons un résultat décevant puisque la classe 1 obtient le plus petit score. Toutefois, la classe 3 (représentant le plus haut niveau d'altération du sens) n'est pas loin et les résultats de chaque classe se tiennent à peu près à une même moyenne. Le score le plus élevé (classe 2) démontre que le contexte de la phrase ou du texte global est nécessaire à la bonne compréhension du lecteur. Néanmoins, en regardant l'ensemble des résultats, nous pouvons être satisfaites car un peu plus de 65% de l'information parvient au lecteur malgré les fautes, dans cette traduction du texte brut.

Les erreurs qui ont le plus créé de dommages sont les traductions littérales des noms propres (classe 3), lesquels sont assez nombreux dans notre texte (annexe 14) : ex. Les Contamines > *loro contaminati* > *eux contaminés*. De même, les mauvais choix de catégorie lexicale ont freiné la compréhension, notamment pour un terme qui est au centre de l'information véhiculé par notre texte : ex. la glisse > *scivola* (verbe) > sport di scivolamento (nom). Nous avons été également surpris de constater que Systran avait formé quelques périphrases (ex. le dénivelé > *mutato il livello* > il dislivello), rendant la compréhension difficile comme cela avait d'ailleurs été souligné par les experts italiens dans le test de compréhensibilité (section 5.4.3). Tout ceci a pu être évidemment résolu une fois les dictionnaires utilisateurs appliqués. De plus, par rapport au type de texte qui est le nôtre, nous avons remarqué que les erreurs de pronom avaient sensiblement fait perdre le ton vendeur du texte. En effet, tout le texte s'adresse directement au potentiel touriste (à travers le vous) et la mauvaise gestion de ces pronoms personnels

(vous > *li* > *lui*) a effacé cet aspect. Ceci avait également été signalé par les experts italiens qui ne comprenaient pas à qui s'adresse l'auteur du texte.

7.4.2.3 Application des mesures de degré d'amélioration

Pour le classement des erreurs jugées selon le critère de degré d'amélioration (tableau 36), nous avons pris en compte le degré d'amélioration qui pouvait être obtenu pour les différentes fautes et nous avons établi l'échelle suivante :

- La classe 4 regroupe les erreurs facilement corrigeables par l'ajout d'un dictionnaire utilisateur ou par une intervention de pré-édition. Par exemple, les catégories lexicales peuvent être désambiguïsées en n'autorisant qu'une catégorie (technique (adjectif) > *tecnica* (nom) > *tecnico* (adjectif)).
- La classe 5 regroupe les erreurs qui ne peuvent être résolues qu'au moment de la post-édition sans demander un grand effort. Par exemple, les prépositions ont vraiment été mal gérées par le système (ex : d'un côté > *dell'altro lato* > *dall'altro lato*).
- La classe 6 regroupe les erreurs qui sont difficiles à résoudre et demandent un effort de post-édition notable, voire de traduction complète. Le système a produit quelques phrases dont les tournures françaises ne sont pas adaptées à la langue italienne (ex : au départ de la station de ski au col > *alla partenza della stazione sciistica al collare* > *dal principio al valico della stazione*). De plus, certaines phrases du texte source ont été tout bonnement ignorées par le système qui les a effacées dans la sortie.

Classe 4	Classe 5	Classe 6
Vocabulaire (tous)	Accord	Casse
Nom propre (tous)	Article	Ordre des mots (2)
Périphrase	Préposition	
Catégorie grammaticale	Ordre des mots (13)	
Chiffre	Ponctuation	
Expression	Pronom	

Tableau 36. Classement des erreurs selon le critère de degré d'amélioration.

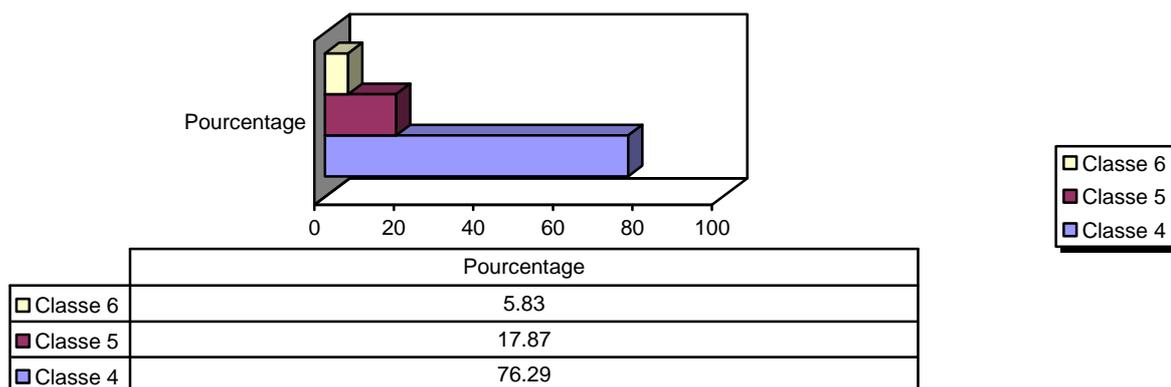
En ce qui concerne la catégorie d'erreur « ordre des mots », celle-ci apparaît dans le tableau 36 deux fois car 2 fautes (structure de la phrase) ont été jugées plus graves que les autres (position de l'adjectif).

7.4.2.4 Résultats du classement des erreurs pour le critère de degré d'amélioration

Les résultats obtenus dans le tableau 36 peuvent être visualisées sous forme de pourcentage dans le graphique 9. Le calcul des pourcentages s'est fait selon le calcul suivant :

$$\text{(addition du nombre d'erreurs par catégorie) / (nombre d'erreurs total) * 100 = \text{pourcentage d'erreur de la classe } x$$

Dans l'idéal, pour obtenir une traduction hautement améliorable en pré-édition, il faudrait que le taux d'erreurs de la classe 4 soit le plus élevé, c'est-à-dire 100%.



Graphique 9. Pourcentages du classement des erreurs selon le critère de degré d'amélioration.

Cette fois-ci nous sommes particulièrement satisfaites des résultats, puisque les erreurs de la classe 4 sont les plus nombreuses. La sortie du système (I) se prêterait donc bien aux corrections de pré-édition, laissant une marge de post-édition raisonnable. De plus, elle présente des erreurs qui sont facilement post-éditables et peu de problèmes nécessitant une reconstruction complète de la phrase. Nous pouvons nous demander si cela est dû à l'appartenance commune de la paire de langues (français-italien) au groupe linguistique latin, favorisant ainsi une traduction acceptable quoique imparfaite. Toutefois, les locutions culturelles ou expressions ont montré les diversités des deux langues sans en altérer complètement le sens.

Nous avons précédemment signalé que certains types d'erreurs comme l'« ordre des mots » auraient pu, au vu de leur gravité, apparaître dans deux classes. Au moment de réaliser le classement, nous nous sommes souvent confrontées à une réflexion sur le positionnement d'une catégorie dans une classe. Ainsi, nous aurions pu également faire une distinction de gravité pour le vocabulaire. De fait, si certains mots peuvent être gérés au niveau de la spécialisation des dictionnaires et appartiennent donc à la classe 4, d'autres, au contraire, peuvent ou doivent être traité au niveau de la post-édition (classe 5). En effet, certains termes français se sont trouvés avoir deux voire plusieurs équivalents en italien selon leur contexte. Par exemple, le terme « enneigement » a été reconnu par Systran comme équivalent de « altezza del manto nevoso » (contexte : météorologie ou lors d'utilisation de canon à neige). De manière générale, on parlera plus simplement de « manto nevoso », mais la nuance est subtile. Nous nous sommes

donc demandé si nous devons traiter l'erreur au moment de la spécialisation du dictionnaire ou de la post-édition. Après discussion avec notre expert en post-édition

I. <u>INTRODUCTION</u>	9
II. <u>TRADUCTION AUTOMATIQUE ET LANGAGES CONTRÔLÉS</u>	13
2.1 INTRODUCTION	13
2.2 LA TRADUCTION AUTOMATIQUE (TA)	13
2.2.1 INTRODUCTION	13
2.2.2 LA TRADUCTION AUTOMATIQUE	15
2.2.3 UN BREF HISTORIQUE.....	19
2.2.4 AVANTAGES ET INCONVÉNIENTS DE LA TA	21
2.3 LE LANGAGE CONTRÔLÉ	24
2.3.1 INTRODUCTION	24
2.3.2 LE LC.....	25
2.3.3 HISTORIQUE	26
2.3.4 STRUCTURE D'UN LANGAGE CONTRÔLÉ	30
2.3.5 AVANTAGES ET DÉSAVANTAGES DU LANGAGE CONTRÔLÉ.....	34
2.4 CONCLUSION	37
III. <u>LES OUTILS ET LE CORPUS</u>	39
3.1 INTRODUCTION	39
3.2 PRÉSENTATION DES OUTILS UTILISÉS	39
3.2.1 SYSTÈMES DE TRADUCTION	39
3.2.2 OUTIL DE VÉRIFICATION	46
3.2.3 OUTIL POUR LE CORPUS	48
3.3 LE CORPUS	49
3.3.1 INTRODUCTION	49
3.3.2 DÉFINITION DU CORPUS	49
3.3.3 CARACTÉRISTIQUES DU CORPUS ÉTUDIÉ	51
3.4 CONCLUSION	58
IV. <u>MÉTHODOLOGIE</u>	61
4.1 INTRODUCTION	61
4.2 CRÉATION DES DICTIONNAIRES SPÉCIALISÉS	61
4.3 LA PRÉSÉLECTION DES RÈGLES DE CONTRÔLE	67
4.3.1 RÈGLES GÉNÉRALES	69
4.3.2 RÈGLES SPÉCIFIQUES	79
4.4 CONCLUSION	89
V. <u>PRÉPARATION DE L'ÉVALUATION</u>	91

5.1 INTRODUCTION.....	91
5.2 MÉTHODOLOGIE : CORPUS, EXPERTS ET OUTILS.....	92
5.2.1 LES CORPUS TESTS	92
5.2.2 LES EXPERTS	94
5.2.3 LES OUTILS	95
5.3 DÉFINITION DES CRITÈRES DE QUALITÉ.....	95
5.3.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	96
5.3.2 LANGAGE CONTRÔLÉ : EFFICACITÉ DES RÈGLES DE CONTRÔLE.....	97
5.3.3 LANGAGE CONTRÔLÉ : COMPRÉHENSIBILITÉ	97
5.3.4 LISIBILITÉ	97
5.4 DESCRIPTION DES TESTS.....	97
5.4.1. COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	98
5.4.2 LANGAGE CONTRÔLÉ.....	100
5.4.3 LANGAGE CONTRÔLÉ.....	101
5.4.4 LISIBILITÉ	102
5.5 ACROLINX.....	102
5.5.1 ACROLINX.....	103
5.6 CONCLUSION	103
VI. <u>EXÉCUTION DE L'ÉVALUATION</u>	105
6.1 INTRODUCTION.....	105
6.2 RÉSULTATS	105
6.2.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	105
6.2.2 LANGAGE CONTRÔLÉ : EFFICACITÉ DES RÈGLES DE CONTRÔLE.....	107
6.2.3 LANGAGE CONTRÔLÉ : COMPRÉHENSIBILITÉ	108
6.2.4 LISIBILITÉ	110
6.2.5 COMPRÉHENSIBILITÉ ACROLINX	112
6.3 CONCLUSION	114
VII. <u>POST-ÉDITION : LE CAS ITALIEN</u>	117
7.1 INTRODUCTION.....	117
7.2 LA POST-ÉDITION (PE).....	117
7.2.1 INTRODUCTION	117
7.2.2 QU'EST-CE QUE LA POST-ÉDITION ?	118
7.2.3 TYPES DE POST-ÉDITION	119
7.3 CRITÈRES D'ÉVALUATION	122
7.3.1 CRITÈRES DE QUALITÉ	122
7.4 ÉVALUATION DES ERREURS	123
7.4.1 TABLEAU DES ERREURS	124
7.4.2 CLASSEMENT DES ERREURS.....	128
7.5 ÉVALUATION SUR LE CORPUS (J)	138
7.5.1 TABLEAU DES ERREURS POUR CORPUS (J).....	138
7.5.2 CLASSEMENTS DES ERREURS SELON LES CRITÈRES DE QUALITÉ POUR LE CORPUS (J) ET RÉSULTATS	141
7.6 DISCUSSION.....	145

7.7 CONCLUSION	147
-----------------------------	------------

VIII. SYSTRAN VS. LUCY LT: UNE ÉVALUATION COMPARATIVE DES SORTIES DES SYSTÈMES..... 149

8.1 INTRODUCTION.....	149
8.2. MÉTHODOLOGIE : CORPUS, EXPERTS ET OUTILS.....	149
8.2.1 LES CORPUS	149
8.2.2 LES EXPERTS	151
8.2.3 LES OUTILS	152
8.3 SPÉCIALISATION DES DICTIONNAIRES DE LUCY LT	152
8.4 DÉFINITION DES CRITÈRES DE QUALITÉ.....	154
8.4.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	155
8.4.2 ÉVALUATION COMPARATIVE : « SENTENCE-BASED RANKING »	155
8.4.3 WER	156
8.5 DESCRIPTION DES TESTS.....	156
8.5.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	156
8.5.2. ÉVALUATION COMPARATIVE : « SENTENCE-BASED RANKING »	157
8.5.3 WER	158
8.6 RÉSULTATS	158
8.6.1 COUVERTURE LEXICALE : TERMES ET VOCABULAIRE	158
8.6.2 SENTENCE-BASED RANKING	160
8.6.3 WER	162
8.7. CONCLUSION	163

IX. CONCLUSION..... 165

X. BIBLIOGRAPHIE..... 169

INDEX DES FIGURES.....	175
INDEX DES GRAPHIQUES	177
INDEX DES TABLEAUX.....	179

et pour des raisons pratiques (le terme général apparaissant plus souvent dans le texte), nous avons décidé de créer une entrée dans le dictionnaire de Systran et d'intervenir au moment de la post-édition lorsque le contexte le demandait (annexe 11, corpus (J)). En conclusion, ce terme aurait donc pu appartenir à la classe 4 comme à la classe 5. Or, comme précédemment dit, ce genre de décision est définitivement subjectif.

Somme toute, ces résultats nous ont permis de voir que les erreurs du corpus (I) présentent un niveau de gravité acceptable (35%) et surtout qu'elles sont facilement corrigibles. Un travail sur la spécialisation des dictionnaires pourrait résoudre la majorité des problèmes puisque le vocabulaire et les noms propres inconnus présentent le pourcentage le plus haut. La post-édition intervient sur des erreurs dites mineures

pour l'intelligibilité du texte et ne demandant pas ou peu de traduction humaine. Toutefois, au vu du nombre d'erreurs (789), l'effort de post-édition pour le corpus (I) reste conséquent.

7.5 Evaluation sur le corpus (J)

Ayant produit un corpus avec dictionnaires utilisateurs et langage contrôlé (J), nous avons décidé de le soumettre au processus d'évaluation des erreurs jusqu'ici appliqué au corpus (I). Par ce biais, nous avons le moyen de démontrer l'impact sur le texte cible du travail commun qui a été fait sur le texte source (chapitre III et IV). Idéalement, nous devrions atteindre l'objectif suivant : une diminution significative du nombre d'erreurs (sous-section 7.5.1), de la gravité de ces dernières (sous-section 7.5.2) et par conséquent, de l'effort de post-édition (section 7.6).

7.5.1 Tableau des erreurs pour corpus (J)

Comme nous l'avons fait précédemment, nous avons relevé toutes les fautes présentes dans le corpus (J) (annexe 15) puis nous avons établi une typologie des erreurs trouvées (tableau 37) en le comparant à la traduction post-éditée (L). En suivant notre méthodologie (sous-section 7.4.1), nous avons réalisé un graphique (graphique 10) reproduisant le pourcentage d'erreurs selon le calcul :

$$(\text{nombre d'erreurs}) / (\text{nombre de mots total}) * 1000 = \text{pourcentage d'erreur de la catégorie x}$$

Il est important de souligner que le nombre de phrases total pour le corpus (J) est plus élevé que celui du corpus (I), car nous avons dû notamment nous conformer à la règle de contrôle vii. : découper les phrases longues en phrases courtes (sous-section 4.2.2). Notre corpus (J) est donc composé de 225 phrases et de 3'946 mots.

Catégories de fautes	Nombres d'erreurs dans
----------------------	------------------------

		le corpus (J)
1.	Préposition	45
2.	Casse	33
3.	Ordre des mots	28
4.	Article	28
5.	Accord	22
6.	Expression	18
7.	Vocabulaire	16
8.	Catégorie grammaticale	7
9.	Pronom	6
10.	Ponctuation	3
11.	Chiffre	2
12.	Nom propre (trad.littérale)	1
Nombre d'erreurs total		209

Tableau 37. Tableau des erreurs de la sortie du système (J).

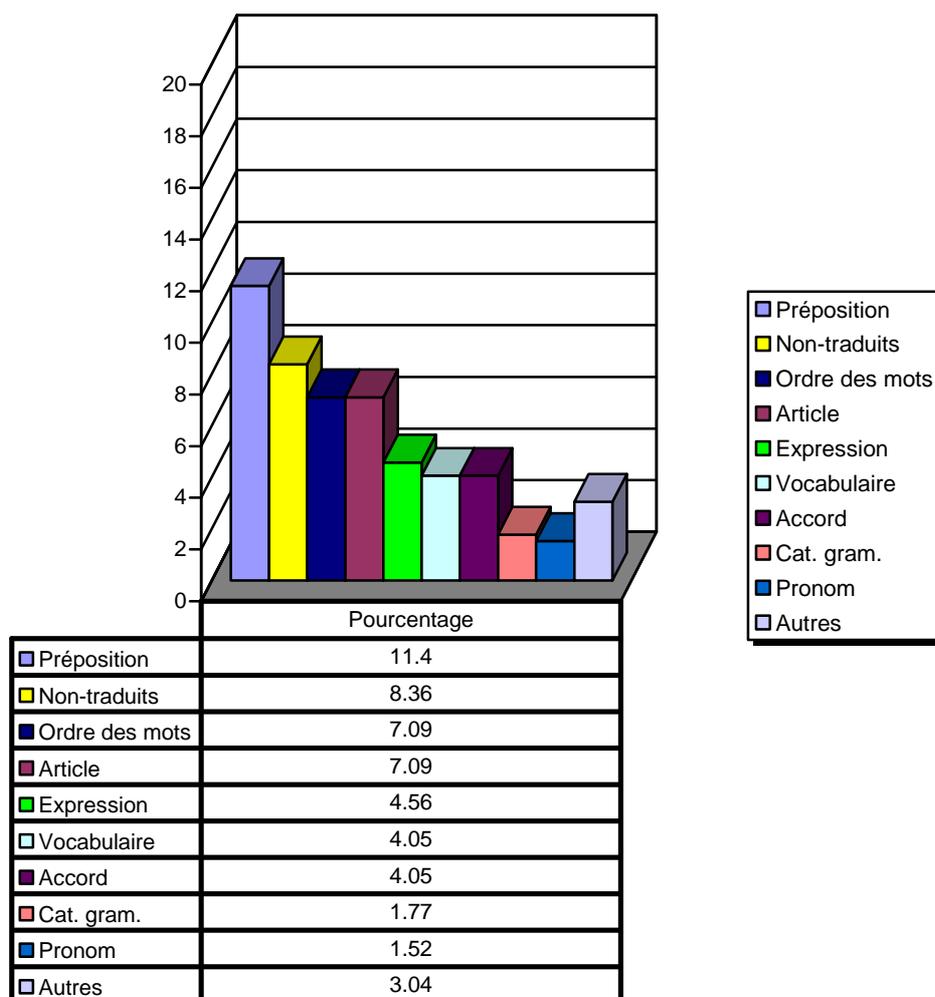
La première information que nous pouvons tirer de notre document recensant les erreurs du corpus (J) (annexe 15), est l'obtention de 101 phrases complètement exemptes d'erreurs, à savoir 44% du texte comparé à seulement 2,5% pour le corpus brut (I). Par ailleurs, dans le tableau 37, nous pouvons immédiatement constater que le nombre de catégories diminue par rapport au tableau 34 (sous-section 7.4.1). De plus, au vu des chiffres du graphique 10, les fautes les plus représentées ne dépassent pas les 12% du texte global. Cette diminution est dû entre autre à la réduction de corrections touchant à la terminologie (vocabulaire général/spécifique et noms propres) qui faisait défaut dans le corpus (I) (sous-section 7.4.1). Par contre le nombre de fautes relatives

aux prépositions est toujours très élevé, montrant que Systran ne parvient pas du tout à gérer de manière correcte le passage du français à l'italien. L'exemple le plus surprenant que nous ayons trouvé dans notre corpus (J) est le suivant : « Apertura degli impianti sciistici : *Dal del* 11/12/2010 al 25/04/11 ». Le système a proposé les deux équivalents italiens de la préposition « de », n'ayant évidemment pas la connaissance du monde nécessaire pour faire une distinction :

dal > du 11/12/2010 au 25...

del > le 11/12/2010

En ce qui concerne les phrases où le problème touchait des dates, nous avons pu régler le problème par l'application de la mémoire de traduction créée dans Systran pour les segments répétitifs (sous-section 4.2). Sinon la post-édition a dû être utilisée.



Graphique 10. Pourcentages d'erreurs du corpus (J)

7.5.2 Classements des erreurs selon les critères de qualité pour le corpus (J) et résultats

La seconde phase de notre évaluation sur le corpus (J) nous permet d'obtenir des résultats quant aux performances de la MT après l'application du LC et au niveau de la qualité. Nous avons donc repris nos critères de qualité, intelligibilité et degré d'amélioration, et avons appliqué les mêmes échelles (section 7.4.2.).

7.5.2.1 Application et résultats des mesures d'intelligibilité pour le corpus (J)

Nous avons repris les mesures d'intelligibilité (7.4.2.1) et avons classé les erreurs trouvées dans le corpus (J) selon le degré d'altération de l'intelligibilité du texte. La classe 1 représentant les fautes qui ont peu ou pas d'impact et la classe 3, celles qui altèrent complètement le sens (tableau 38).

Classe 1	Classe 2	Classe 3
Ponctuation	Expression	Catégorie grammaticale
Article	Préposition	Pronom personnel
Ordre des mots	Chiffre	Casse
Accord		Nom propre (trad. littérale)
Vocabulaire		

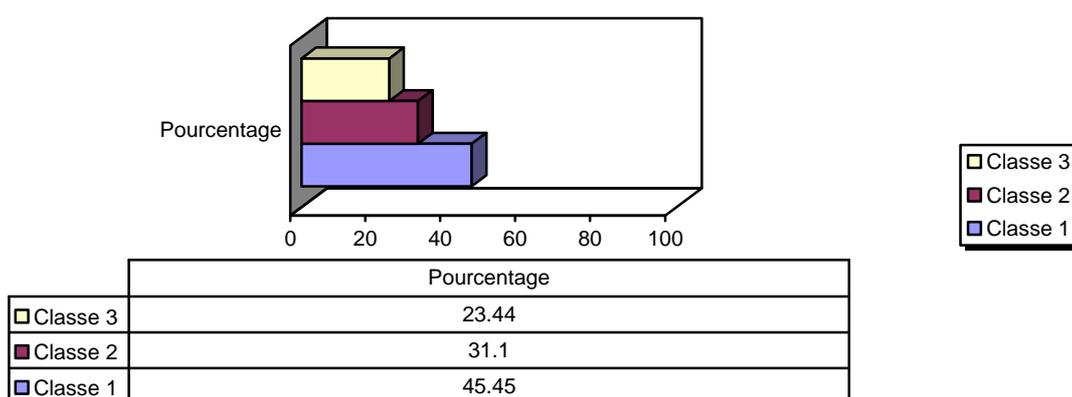
Tableau 38. Classement des erreurs selon le critère d'intelligibilité pour le corpus (J).

Les résultats obtenus dans le tableau 38 peuvent être visualisés sous forme de pourcentage dans le graphique 11. Le calcul des pourcentages s'est fait selon le calcul suivant :

(addition du nombre d'erreurs par catégorie) / (nombre d'erreurs total) *

100 = pourcentage d'erreur de la classe x

Dans l'idéal, pour obtenir une traduction acceptablement intelligible, sans aucun impact des erreurs, il faudrait que le taux d'erreurs de la classe 1 soit le plus élevé, c'est-à-dire 100%.



Graphique 11. Classement des erreurs selon le critère d'intelligibilité pour le corpus (J)

Au vu des chiffres du graphique 11, nous pouvons remarquer que près de 76% de l'information parvient à l'utilisateur final de manière presque intelligible ou du moins de manière acceptable ; cela constitue une amélioration de 10% par rapport au corpus (I) (sous-section 7.4.2.2). Notre objectif est atteint puisque, contrairement aux résultats du corpus (I) (sous-section 7.4.2.2), les erreurs de la classe 1 représentent le score le plus élevé. Le taux d'erreurs de la classe 2 et 3 présentent une baisse significative de près de 10% chacun. De plus, les catégories (ordre des mots et vocabulaire) qui nous avaient posé un souci de positionnement dans le classement (sous-section 7.4.2.4) n'ont pas nécessité cette fois-ci une réflexion puisque les erreurs sujettes à discussion ont disparues.

7.5.2.2 Application et résultats des mesures de degré d'amélioration pour le corpus (J)

Nous avons repris les mesures de degré d'amélioration (sous-section 7.4.2.2) et avons classé les erreurs trouvées dans le corpus (J) selon la gravité et la possibilité d'amélioration du texte. La classe 1 représentant les fautes qui peuvent le plus facilement être corrigées et la classe 3, celles qui nécessitent un travail de post-édition conséquent voire de traduction (tableau 39):

Classe 4	Classe 5	Classe 6
Catégorie grammaticale	Accord	
Expression	Article	
Chiffre	Préposition	
Nom propre (trad.littérale)	Ordre des mots	
Casse	Vocabulaire	
	Ponctuation	
	Pronom	

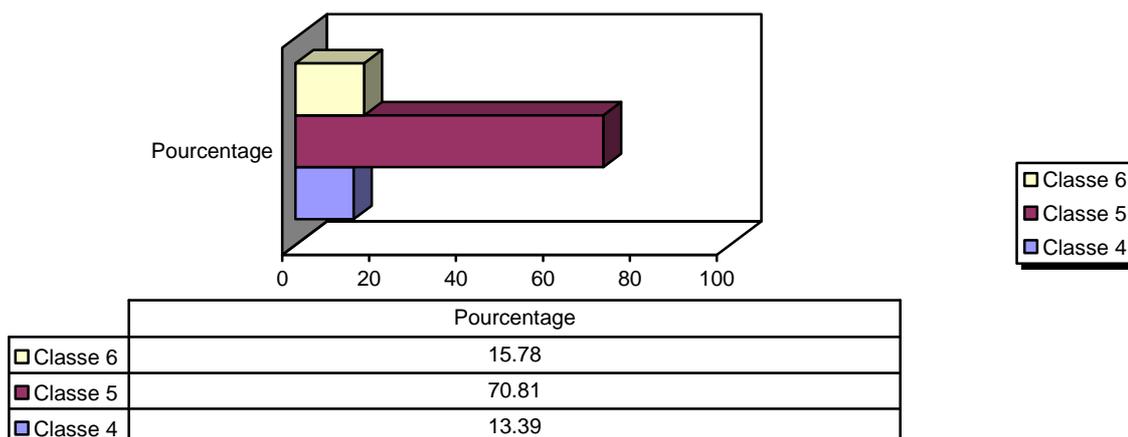
Tableau 39. Classement des erreurs selon le critère de degré d'amélioration pour le corpus (J).

Les résultats obtenus dans le tableau 39 peuvent être visualisés sous forme de pourcentage dans le graphique 12. Le calcul des pourcentages s'est fait selon le calcul suivant :

$$(\text{addition du nombre d'erreurs par catégorie}) / (\text{nombre d'erreurs total}) *$$

$$100 = \text{pourcentage d'erreur de la classe } x$$

Dans l'idéal, pour obtenir une traduction parfaitement améliorable en post-édition, il faudrait que le taux d'erreurs de la classe 5 soit le plus élevé, c'est-à-dire 100%. Puisque le travail de pré-édition a été effectué sur ce corpus, les taux de la classe 4 et 6 devraient obtenir des scores très bas.



Graphique 12. Classification des erreurs selon le critère de degré d'amélioration pour le corpus (J)

Comme nous le voyons dans le graphique 12, les fautes de la classe 1 obtiennent le plus petit score et donc tout ce qui pouvait être amélioré au niveau de la pré-édition a pratiquement disparu. Il reste toutefois des phrases ou des mots qui, pour une raison ou une autre, n'ont pas été traités par le système nécessitant donc une traduction humaine. Néanmoins, nous avons atteint notre objectif puisque le taux de post-édition (classe 5) est le plus élevé : il reste les fautes liées au vocabulaire (et expression) dont le contexte se détermine uniquement par une connaissance du monde et des usages ainsi que quelques erreurs facilement corrigibles. Au vu des résultats, nous pouvons dire que notre corpus (J) se prête volontiers à la post-édition rapide.

Somme toute, des graphiques 8 et 9 aux graphiques 11 et 12, nous remarquons un renversement des résultats : pour le corpus (I), les erreurs (de la plus grave à la moins grave) se partagent de manière égale l'ensemble du texte. Mais si le corpus (I) est donné comme présenté dans l'annexe 14, le post-éditeur se retrouve face à un nombre tel de fautes répétitives que cela l'oblige à considérer, en terme de temps, une traduction humaine. Par la spécialisation des dictionnaires et l'application du LC, le corpus (J) n'obtient plus un nombre incalculable d'erreurs. On constate dès lors le besoin de la post-édition (rapide) pour obtenir un texte 100% correct. La surprise reste au niveau de l'intelligibilité puisque l'information principale a toujours été plus ou moins véhiculée et que l'application du LC n'a fait qu'en augmenter encore plus le taux, déjà élevé. Dans la section qui suit, nous avons comparé les résultats obtenus tout

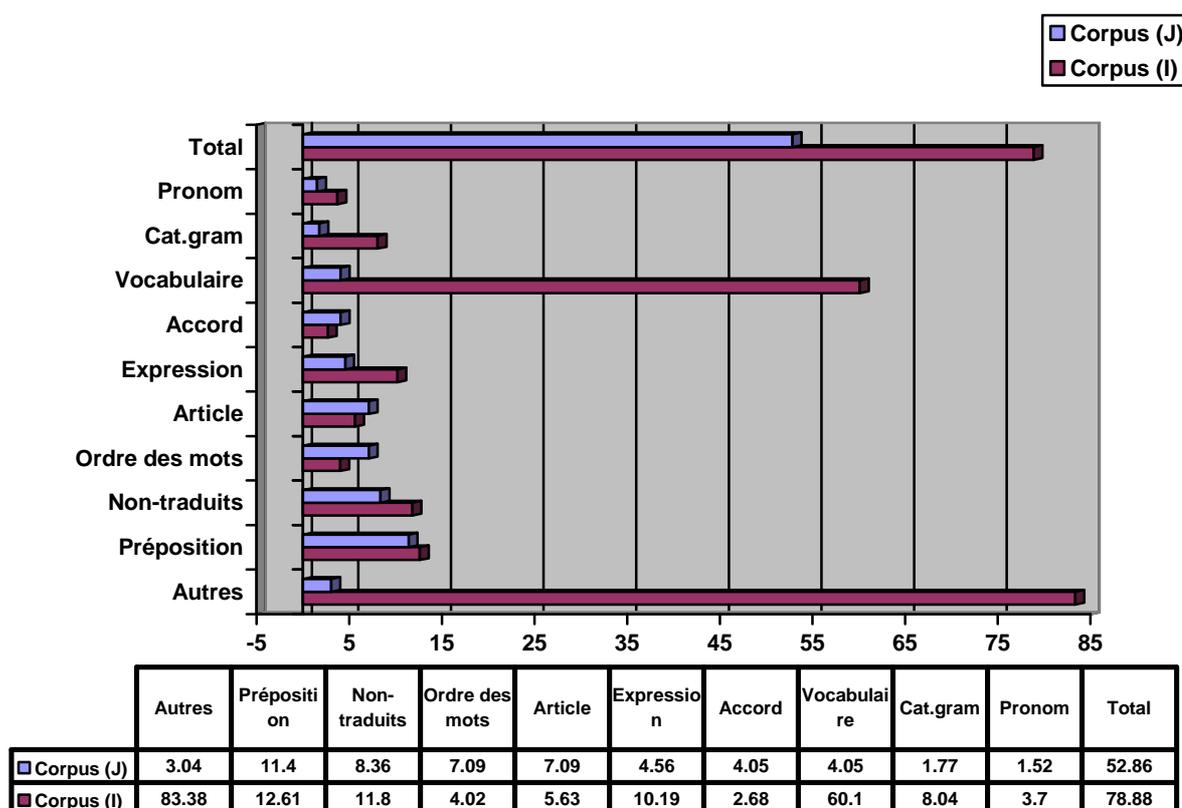
au long de l'évaluation afin de vérifier l'impact du travail commun de pré-édition et si l'effort de post-édition s'en trouve influencé.

7.6 Discussion

Après avoir fait le relevé des erreurs et leur classement selon les mesures de qualité choisies, intelligibilité et degré d'amélioration, pour chacun des corpus, nous pouvons comparer les résultats obtenus. Pour ce faire, nous avons tout d'abord réalisé un graphique 13 représentant le pourcentage d'erreurs pour le corpus (I) et (J) selon le calcul suivant :

$$(\text{nombre d'erreurs total}) / (\text{nombre de mots total}) * 1000 = \text{pourcentage d'erreur de la catégorie x}$$

Nous avons repris les catégories de fautes du corpus (J) et avons classé toutes les autres apparaissant dans le tableau des erreurs du corpus (I) (tableau 34) sous la catégorie « Autres ».



Graphique 13. Pourcentage d'erreurs corpus (I) versus corpus (J)

Au vu des résultats du graphique 13, nous pouvons constater que l'objectif de diminuer le nombre d'erreurs faites par la machine lors du processus de traduction, grâce à l'application du LC, est atteint. De fait, dans 7 catégories présentes, nous pouvons noter une réduction notable du nombre de fautes relevées une fois le LC appliqué, sans compter toutes les catégories classées sous « Autres » qui n'ont plus lieu d'être. En moyenne, le total d'erreur a diminué d'environ 26%. Cependant, le réarrangement de certaines phrases du texte source pour correspondre aux règles de LC a pu engendrer des fautes qui n'existaient pas ; ce qui pourrait expliquer la légère augmentation de certains types d'erreurs dans le corpus (J).

Néanmoins, pour obtenir un texte publiable, il reste un effort de post-édition à fournir. Pour rendre compte de ce dernier et suivant la méthodologie d'évaluation manuelle de la qualité de la sortie du système, nous avons comptabilisé tous les changements effectués sur le corpus (I) (Federman, 2011 ; De Almeida & O'Brien, 2010) à partir de la traduction post-éditée (K) que nous avons proposée. Nous avons divisés les changements comme suit :

- les corrections comptabilisent toutes les entrées faites au niveau des dictionnaires utilisateurs.
- les opérations, toutes les modifications effectuées au moment de la post-édition ; nous avons ainsi le moyen de mesurer les efforts fournis pour cette dernière.

Nous avons fait de même avec le corpus (J) et la traduction post-éditée (L). Selon le calcul suivant : $(\text{nombre d'opérations de PE}) / (\text{nombre de mots total}) * 1000$ nous pouvons obtenir le pourcentage de PE à effectuer. Les résultats obtenus sont présentés sous le tableau 40 ci-dessous.

Corpus (I) – 166 phrases :

Nombre d'opérations effectuées de PE : 269

Nombre de corrections effectuées dans les dictionnaires : 297

⇒ **72% de post-édition + temps de correction sur les dictionnaires**

Corpus (J) – 225 phrases :

Nombre d'opérations effectuées de PE : 209

Nombre de corrections effectuées dans les dictionnaires : 0

⇒ **53% de post-édition**

Tableau 40. Comptabilisation des modifications du corpus (I) et (J).

Nous remarquons dans le tableau 40 que le nombre d'opérations effectuées sur le corpus (J) s'avoisine à celui du corpus (I). À ce niveau là, les efforts de post-édition varient de très peu. Cependant, pour notre texte, l'impact des dictionnaires sur la sortie détermine l'effort de post-édition. Ainsi, nous avons dû créer 297 entrées (sans compter la mémoire de traduction) pour des mots qui se répètent plus d'une fois dans l'ensemble de nos corpus. Face à une telle quantité de post-édition manuelle concernant en priorité un vocabulaire et une terminologie très répétitifs, il est clair que les avantages de la spécialisation de dictionnaires sont démontrés.

7.7 Conclusion

Dans ce chapitre, nous avons comme objectif de travailler sur les sorties de systèmes produites suite à un travail commun effectué sur le texte source. Pour ceci, nous avons travaillé au niveau de la post-édition sur le type de document qui est le nôtre, touristique, et la paire de langue français-italien. À travers un relevé manuel des erreurs, nous avons pu également rendre compte de la performance du système choisi pour effectuer la traduction.

Nous avons pu vérifier que l'information principale a été le plus souvent véhiculée malgré les erreurs. Nous avons obtenu des traductions qui restaient compréhensibles et nous pouvons dire qu'une post-édition rapide est tout à fait acceptable pour ce type de texte. L'ambition des textes touristiques est avant tout d'informer l'utilisateur final et la notion de qualité est sans aucun doute moins importante. Nous pouvons également souligner que l'appartenance linguistique de la paire de langues au même groupe (latin) a eu un certain impact sur le niveau d'altération du sens. Les deux langues étant assez proches, les experts du test de compréhensibilité et de celui pour la post-édition, ont facilement retrouvés la signification du texte source (sans avoir ce dernier sous les yeux). Toutefois, le travail de pré-édition et de spécialisation des dictionnaires s'est avéré essentiel afin d'obtenir une diminution significative de l'effort de post-édition. Ainsi, plus du double des phrases du corpus (J) ont été correctement traduites et l'effort de post-édition a diminué de près de 25%.

Nous pouvons donc conclure que le travail que nous avons effectué en commun a eu l'impact espéré sur la post-édition des sorties de langue italienne.

VIII. Systran vs. Lucy LT: une évaluation comparative des sorties des systèmes

8.1 Introduction

Dans ce chapitre, nous comparerons les sorties de Systran et de Lucy LT, du point de vue de la qualité linguistique de la traduction automatisée des textes de type touristique. A cette fin, nous avons mis au point plusieurs tests dont les paramètres sont inspirés des mesures proposées par Dabbadie *et al.* (2002), Miller et Vanni (2001), Yuste-Rodrigo et Braun-Chen (2001) ainsi que Vilar et Banchs (2007). Nous avons également fait une évaluation automatique : nous avons calculé le Word Error Rate (WER) des sorties des systèmes. Nous avons décidé de soumettre à Lucy LT la version du corpus source contrôlée pour Systran afin de déterminer à quel point le LC est indépendant des systèmes de TA.

Les paramètres, les tests et les résultats sont exposés plus en détail dans les sections 8.4, 8.5 et 8.6 de ce chapitre. Toutefois, avant de présenter les mesures et les tests, nous allons introduire la méthodologie utilisée pour ce chapitre (8.2), et nous pencher sur la spécialisation de Lucy LT (8.3).

8.2. Méthodologie : corpus, experts et outils

8.2.1 Les corpus

Les corpus que nous avons choisis pour mener à bien cette évaluation correspondent à la version brute et à la version contrôlée des cellules 142 à 342 du corpus source. La version brute correspond au corpus E du chapitre V (5.2.1). Ce dernier totalise 1564 mots. Le mini corpus source contrôlé, que nous appellerons (F), en totalise 1893.

A partir du corpus source E, nous avons produit ces deux sorties (voir annexe 17 et 18) :

(E_Lucy) brut sans dictionnaire

(B_Lucy) brut avec dictionnaire

A partir du corpus (F), nous avons produit ces deux sorties (voir annexe 19 et 20):

(C_Lucy) contrôlé sans dictionnaire

(D_Lucy) contrôlé avec dictionnaire

Nous avons décidé de garder les appellations B pour la sortie brute avec dictionnaire, C et D pour les sorties contrôlées avec et sans dictionnaire afin de montrer que ces sorties sont directement liées aux corpus (B_Systran), (C_Systran et (D_Systran), dont nous avons parlé dans le chapitre V (5.2.1).

Il est important de préciser que la traduction a été faite du français à l'anglais uniquement. Puisqu'il s'agit d'une évaluation comparative entre les sorties de Systran et Lucy, nous avons aligné dans le tableau 41 ci-dessous les correspondances entre les sorties des systèmes :

Sorties	SOURCE	SYSTRAN	LUCY
brute	(E)	(E_Systran)	(E_Lucy)
brute + dictionnaire		(B142-342)	(B_Lucy)
contrôlé	(F)	(C142-342)	(C_Lucy)
contrôlé + dictionnaire		(D142-342)	(D_Lucy)

Tableau 41. Récapitulatif des sorties de Systran et de Lucy LT

Pour des raisons pratiques, nous avons dû créer pour les tests 3 et 4 deux sous-corpus (annexes 21 et 22) à partir des sorties présentées dans le tableau 41 ci-dessus. Les sous-corpus sont détaillés dans le tableau suivant (tableau 42) :

	TEXTE SOURCE	TEXTE DE RÉFÉRENCE	SYSTRAN	LUCY
<u>Test 3.</u> Évaluation comparative : <i>ranking test</i>	(G) : <ul style="list-style-type: none"> ▪ 39 phrases brutes (E) ▪ 33 phrases contrôlées (F) 	→	(G_Systran)	(G_Lucy)
<u>Test 4.</u> WER	→	(H) : 23 phrases provenant du corpus (D142-342) de Systran <u>post-éditées</u>	(H_Systran) : 23 phrases provenant du corpus (D142-342)	(H_Lucy) : 23 phrases provenant du corpus (D_Lucy)

Tableau 42. Les sous-corpus

8.2.2 Les experts

A l'exception de l'évaluation comparative : « *ranking test* » (voir 8.5.2), nous serons à la fois les auteurs et les juges des tests que nous proposons dans ce chapitre. Nous sommes de langue maternelle française et avons une excellente connaissance de l'anglais, avec une expérience non professionnelle de la traduction automatique.

Pour le « *ranking test* », nous avons soumis un questionnaire à un groupe de cinq personnes :

1. L'évaluateur n° 1 est assistante à l'ETI. Elle possède d'excellentes connaissances de l'anglais, et une bonne connaissance du vocabulaire du ski.
2. L'évaluateur n°2 est étudiante en traduction à l'ETI et possède également une excellente connaissance de l'anglais. Étant originaire de la région, elle maîtrise également le vocabulaire spécifique au ski.

3. L'évaluateur n°3 est également étudiante en traduction à l'ETI. Elle possède de très bonnes connaissances de l'anglais et maîtrise parfaitement le vocabulaire du ski, étant une pratiquante assidue de ce sport.
4. L'évaluateur n°4 est elle aussi étudiante à l'ETI. Originnaire d'Afrique du Sud, elle est de langue maternelle anglaise, mais n'est pas très expérimentée au sujet du ski.
5. L'évaluateur n°5 est également étudiante à l'ETI. Elle possède de très bonnes connaissances de l'anglais et du vocabulaire du ski.

8.2.3 Les outils

Comme nous l'avons mentionné précédemment, les outils qui feront l'objet de la comparaison sont les systèmes de TA Systran et Lucy LT, dont nous avons parlé plus en détail dans la section 3.1. Systran peut être considéré comme un système direct évolué (Hutchins, 1986) qui a été pensé pour les besoins d'une information rapide (Gerber & Yang, 1997 :211). Lucy LT, quant à lui, est un système indirect de transfert (sous-section 3.2.2) développé à partir de METAL.

8.3 Spécialisation des dictionnaires de Lucy LT

Avant de commencer l'évaluation comparative des sorties de Lucy LT et de Systran, il convient de dire quelques mots sur le travail de personnalisation des lexiques avec Lucy LT. En effet, afin de produire les sorties (B_Lucy) et (D_Lucy), nous avons créé trois lexiques (deux monolingues et un de transfert) contenant les termes inconnus ou mal traduits par le système, les termes techniques et les noms propres. Le lexique de transfert contient 71 termes, dont 35 noms propres, 17 termes techniques et 19 mots inconnus ou mal traduits. Le lexique monolingue français compte 60 termes, dont 35 noms propres, 13 termes techniques et 12 mots inconnus ou mal traduits. Le lexique monolingue anglais comporte également 60 termes dont 35 noms propres, 11 termes techniques et 14 mots inconnus ou mal traduits.

Comme nous l'avons expliqué dans le chapitre III (3.2.2), il est possible d'améliorer la traduction en associant des tests et/ou des actions aux entrées du lexique de transfert. En utilisant uniquement la sortie (D_Lucy), nous avons relevé trois situations problématiques que nous avons tenté de corriger :

1. Les prépositions :

« Pour les séjours du samedi 18/12/2010 au dimanche 26/12/2010(...) »

« For the stays **of** the Saturday 18/12/2010 **at** the Sunday 26/12/2010(...) »

Dans le lexique de transfert, nous avons ajouté un test et deux actions sur les compléments du nom pour l'entrée séjour/stay:

Test : (XFR-NST-TEST :TEST '((N1 (PREP "de" "à") (TYN TMP))))

Actions : (XFR-NST-MAP :SL-PREPCAN ("de") :TL-PREPCAN "from")

(XFR-NST-MAP :SL-PREPCAN ("à") :TL-PREPCAN "to")

Nous voulions associer les prépositions au nom, et préciser que pour les compléments temporels (samedi, dimanche) venant après « séjour », la traduction des prépositions « de » et « à » devenaient « from » et « to ». Nous avons donc ajouté les compléments ((N1 (PREP "de" "à") (TYN TMP)) et (N1(PREP "from" "to") (TYN TMP)) dans les lexiques monolingues correspondant à l'entrée « séjour » et « stay ». Les actions ont eu pour but d'influencer la traduction des prépositions. Cela s'est révélé efficace pour la première préposition, qui a été rattachée au mot « séjour », mais pas pour la seconde :

« For the stays **from** the Saturday 18/12/2010 **at** the Sunday (...) »

2. L'analyse du mot « versant » :

« (...) d'un côté le versant de la Lombarde (...) »

« (...) of a side **pouring it** of the Lombarde(...) »

Malheureusement, la seule solution consistait à effacer l'entrée « verser » de notre lexique. Nous avons estimé qu'il valait mieux corriger le problème à l'étape de post-édition.

3. L'analyse du segment « des pistes de ski » dans la phrase :

« Le forfait est à retirer lors de l'ouverture des pistes de ski.»

« The ski pass is **ski slopes** to be withdrawn during the opening. »

Nous avons observé dans l'arbre syntaxique généré pour l'analyse (annexe 23) que « des pistes de ski » n'était pas correctement rattaché à « ouverture » comme il aurait dû l'être. Nous avons tenté d'ajouter le complément « piste de ski » à ouverture dans les lexiques monolingue, mais cela n'a pas fonctionné. Nous avons donc entré « ouverture des pistes de ski » dans le lexique de transfert.

Des trois situations, une seule n'a pas pu être résolue complètement par la spécialisation des lexiques. Toutefois, nous avons tout de même pu effectuer quelques améliorations dans le reste de la sortie (D_Lucy), par exemple en corrigeant la position de l'adjectif « mains-libres » en ajoutant une action à l'entrée. En résumé, nous avons tenté de rendre les sorties (B_Lucy) et (D_Lucy) les plus correctes possible, afin de les confronter aux sorties de Systran.

8.4 Définition des critères de qualité

Étant donné qu'il s'agit d'une évaluation comparative de la qualité des sorties des systèmes d'un point de vue linguistique, nous avons décidé de ne pas prendre en compte les caractéristiques techniques et/ou économiques des systèmes testés. En effet, comme nous l'avons dit dans l'introduction, les critères que nous avons retenus, proposés par Dabbadie *et al.* (2002), Miller et Vanni (2001), Yuste-Rodrigo et Braun-Chen (2001) et Vilar et Banchs (2007), déterminent surtout la qualité linguistique des sorties des systèmes. Il convient de préciser qu'ils proviennent pour la plupart de la classification ISLE (*International Standards for Language Engineering*). Outre les mesures nécessitant une évaluation humaine, nous avons estimé que le calcul du WER de la sortie (D_Lucy) à l'aide de l'outil de mesure automatique proposé par Maria Georgescu se révélerait intéressant. Les tests seront décrits dans la section suivante (8.5).

8.4.1 Couverture lexicale : termes et vocabulaire

Par cette mesure, nous vérifions la précision de la terminologie au niveau de la sortie du système. En effet, les termes, qu'ils soient techniques ou généraux, doivent être retransmis fidèlement par le système de TA. Comme pour la sous-section 5.3.1, ce critère comporte deux caractéristiques: le nombre de mots inconnus ou mal traduits et le nombre de termes techniques correctement traduits (Dabbadie *et al.*, 2002 : 12). Dans le cadre de la comparaison entre les deux systèmes de TA, nous avons mesuré le nombre de mots inconnus ou mal traduits ainsi que le nombre de termes techniques correctement traduits des sorties concernées et confronté les résultats (voir 8.5.1 et 8.6.1).

8.4.2 Évaluation comparative : « Sentence-based ranking »

L'approche proposée par Vilar et Banchs (2007) se distingue des critères généralement utilisés pour l'évaluation « humaine » de la TA, bien que le(s) test(s) se fassent aussi au niveau des phrases d'une sortie donnée. La différence se situe dans le fait qu'il n'est pas demandé à l'évaluateur de noter une première fois la compréhensibilité des phrases de la sortie, puis de noter une seconde fois, le plus souvent à l'aide d'une échelle différente, la fidélité de ces mêmes phrases. Au contraire, Vilar et Banchs proposent à l'évaluateur de choisir parmi les différentes sorties de systèmes la traduction qu'il préfère. Les auteurs avancent deux raisons justifiant l'utilisation d'un tel critère: d'un côté, cela simplifie le processus de décision pour les évaluateurs, puisqu'ils font leur choix en prenant en compte tous les paramètres (compréhensibilité, fidélité, lisibilité) de manière globale. De l'autre côté, la préférence de l'évaluateur même permet de ne pas imposer de contraintes artificielles sur le processus d'évaluation (étant donné que l'intention principale d'une telle évaluation est de montrer l'utilité d'un système de TA pour des utilisateurs humains) (Vilar & Banchs, 2007 : 97).

8.4.3 WER

Le WER est une unité de mesure automatique des systèmes de reconnaissance vocale, qui opère au niveau du mot. Développé à partir de la distance de Levenshtein⁶, le WER calcule donc le nombre minimum d'insertions, de suppressions et de substitutions pour transformer la sortie du système *hyp* en texte référence *ref* (Popovic & Ney, 2007).

8.5 Description des tests

Ayant exposé ci-dessus les critères que nous avons sélectionnés pour mener à bien cette évaluation comparative, nous allons maintenant présenter les tests :

8.5.1 Couverture lexicale : termes et vocabulaire

Les tests 1 et 2 ont été réalisés avec les sorties (E_Systran) vs. (E_Lucy). Pour le test 2, nous avons en outre utilisé les sorties (C142-342) vs. (C_Lucy).

8.5.1.1 Test 1 : Couverture lexicale –Vocabulaire

Tout d'abord, nous avons relevé manuellement le nombre de mots inconnus ou mal traduits pour les sorties (E_Systran) et (E_Lucy), puis nous avons calculé le pourcentage de mots inconnus ou mal traduits pour chaque sortie, comme nous l'avons fait pour le test 2 de la sous-section 5.4.1 (chapitre V).

⁶ La distance de Levenshtein permet de calculer le nombre minimum d'insertions, de suppressions et de substitutions entre deux chaînes afin de rendre la première chaîne identique à la seconde.

in "Multiple Dimension Levenshtein Edit Distance Calculations for Evaluating Automatic Speech Recognition Systems During Simultaneous Speech" J. Fiscus *et al.*, LREC 2006. (Traduction personnelle)

8.5.1.2 Test 2 : Couverture lexicale –Terminologie

Dans un premier temps, nous avons extrait la liste des termes techniques contenus dans le corpus (E). Ensuite, en prenant pour référence le dictionnaire des termes techniques de Systran réalisés par nous (annexe 6), nous avons vérifié les traductions et aligné les équivalents proposés par les corpus (E_Systran) et (E_Lucy). Nous avons fait la même chose pour les corpus (C142-342) et (C_Lucy). Dans un deuxième temps, nous avons attribué des points à chaque équivalent : 1 point si la traduction était exacte, 0.5 point si la traduction était très similaire ⁷ et 0 point si le terme était mal traduit ou inconnu. Après avoir additionné les points, nous avons déduit le pourcentage de termes techniques correctement traduits pour les sorties de Systran et de Lucy LT.

8.5.2. Évaluation comparative : « Sentence-based ranking »

8.5.2.1 Test 3 : ranking des systèmes

Pour ce test, nous avons constitué un fichier Excel avec les corpus (G), (G_Systran) et (G_Lucy). Nous avons aligné chaque phrase (G) avec ses traductions (G_Systran) et (G_Lucy). Nous avons remplacé le nom de Systran et Lucy LT par « Système 1 » et « Système 2 », et mélangé les phrases brutes et contrôlées. Ensuite, nous avons demandé à cinq personnes d'indiquer par « 1 » ou « 2 » dans la colonne qui leur était réservée la traduction qu'elles préféraient. Le but de ce test n'était pas de confronter les versions brutes aux versions contrôlées des deux systèmes, mais bien de découvrir quel système a été le mieux perçu par les évaluateurs. Avec les résultats, nous avons pu établir la moyenne de classement des systèmes (sous-section 8.6.2.1).

⁷ C'est le cas notamment de « enneigement » = « snow cover » : Lucy LT l'a traduit « snow coverage », qui signifie « taux d'enneigement » (annexe 25).

8.5.3 WER

8.5.3.1 Test 4 : WER

Ce test a pour objectif de confronter la sortie (D_Lucy) à un texte de référence afin de calculer le WER de Lucy LT. Pour des raisons de limite de temps, nous n'avons pas pu utiliser le corpus (D_Lucy) en entier. Nous avons donc sélectionné 23 phrases dans le corpus (D142-342), puis nous les avons post-éditées. C'est le sous-corpus (H). En suivant les instructions de Maria Georgescu, nous avons enregistré ces phrases sous *reference_ad.txt*. Nous avons ensuite sélectionné les mêmes phrases dans la sortie (D_Lucy) que nous avons enregistrées sous *hyp_lu.txt*. C'est le sous-corpus (H_Lucy). Nous avons fait la même chose avec le corpus (D142-342) non post-édité, que nous avons appelé (H_Systran). Enfin, nous avons soumis les fichiers à l'outil proposé par Maria Georgescu, qui a calculé de manière automatique la précision, le rappel et le WER de chaque phrase.

8.6 Résultats

Cette section présente les résultats des différents tests, que nous présenterons sous forme de tableaux et/ou de graphiques.

8.6.1 Couverture lexicale : termes et vocabulaire

8.6.1.1 Test 1 : Couverture lexicale – Vocabulaire

Le corpus (E_Systran) est composé de 1741 mots au total et de 317 occurrences uniques. Le corpus (E_Lucy) est composé de 1767 mots au total et de 307 occurrences uniques (annexe 24). Nous avons compté 33 mots inconnus ou mal traduits pour Systran, contre 32 pour Lucy LT, en laissant de côté les noms propres.

Nous avons ensuite calculé le pourcentage de couverture lexicale incorrecte de chaque sortie, exposé dans le tableau ci-dessous (tableau 43) :

	Systran	Lucy LT
nombre de mots inconnus ou mal traduits	33/317	32/307
Pourcentage de couverture lexicale incorrecte	10.41%	10.42%

Tableau 43. Pourcentage de couverture lexicale incorrecte des sorties du système

Étonnamment, le pourcentage de couverture lexicale incorrecte est très proche entre Systran et Lucy LT. C'est un résultat très intéressant : le fait que les chiffres soient presque identiques (une différence de 0.0133 %) montrent que les sorties brutes (E_Systran) et (E_Lucy) ont un taux d'erreur équivalent, du moins au niveau lexical.

8.6.1.2 Test 2 : Couverture lexicale –Terminologie

Tout d'abord, nous avons extrait manuellement 25 termes techniques du corpus source (E) ainsi que leur équivalent tiré du corpus (D142-342), qui fait ici office de référence. Ensuite, nous avons alignés les équivalents tirés du corpus (E_Systran) et du corpus (E_Lucy) et calculé le pourcentage de termes techniques bien traduits. Nous avons fait de même avec les corpus (C142-342) et (C_Lucy) (voir annexe 25) pour voir si le LC amenait une amélioration quelconque à la traduction. Les résultats sont détaillés dans le tableau 44 ci-dessous :

	Systran	Lucy LT
sortie brute	4/25	6.5/25
pourcentage de termes techniques correctement traduits pour la sortie brute	16%	26%
sortie contrôlée	4/25	7.5/25
pourcentage de termes techniques correctement traduits pour la sortie contrôlée	16%	30%
Moyenne	16%	28%

Tableau 44. Pourcentage de termes techniques correctement traduits.

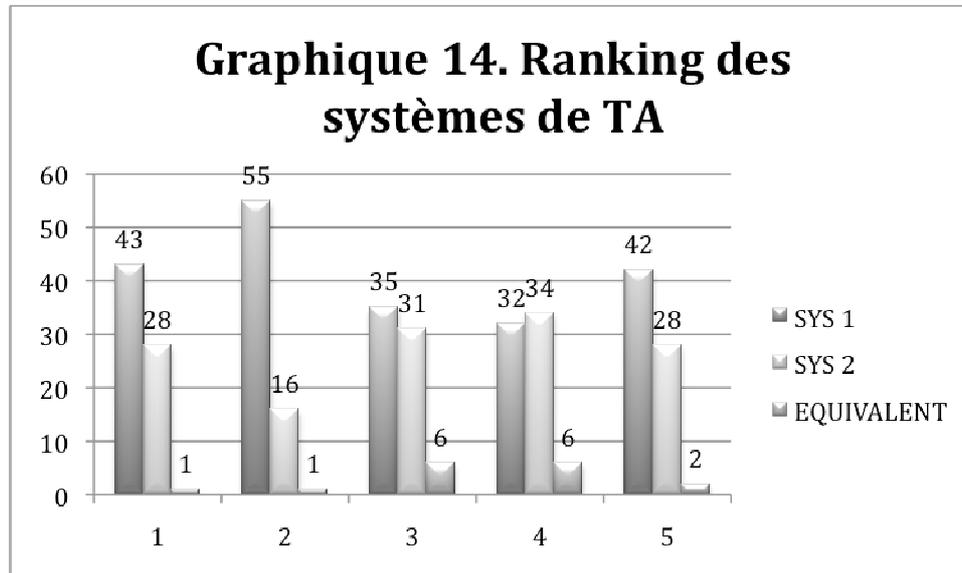
Lucy LT a un pourcentage de termes techniques correctement traduits de 12% supérieur à celui de Systran, comme nous pouvons le voir dans le tableau 45. Cela signifie que Lucy LT semble plus pointu au niveau de la terminologie « technique ». De plus, nous pouvons noter que le taux de termes techniques bien traduits augmente de 4% entre la version brute et la version contrôlée de Lucy LT alors qu'elle reste stable pour Systran. Lucy LT présente donc une amélioration à deux niveaux.

8.6.2 Sentence-based ranking

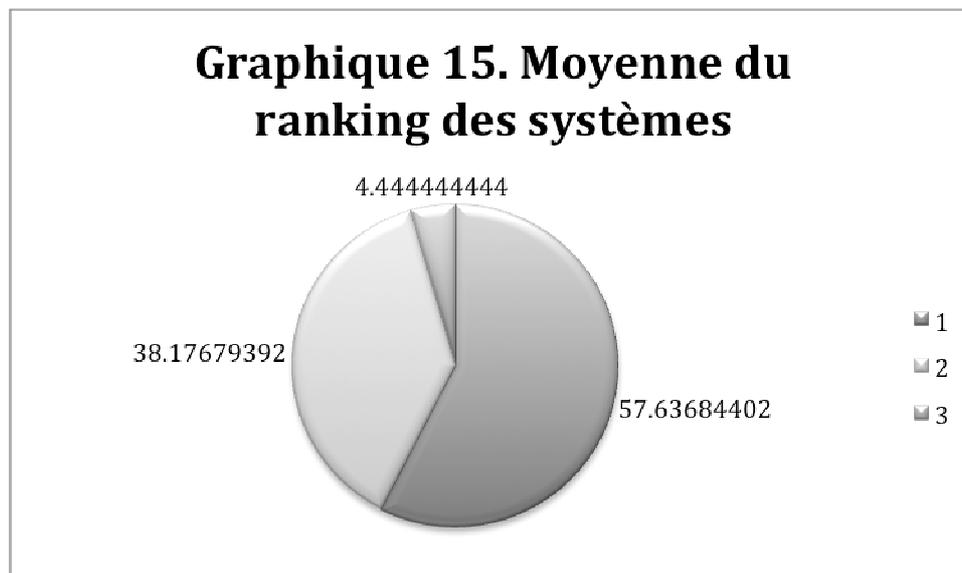
8.6.2.1 Test 3 : ranking des systèmes

Comme nous l'avons expliqué dans la sous-section 8.5.2.1, nous avons soumis le fichier Excel contenant les corpus (G), (G_Systran) et (G_Lucy) à cinq personnes. Nous leur avons demandé de choisir entre deux traductions candidates, celle du système 1 ou celle du système 2, en écrivant « 1 » ou « 2 » dans la colonne appropriée. Les phrases ayant été mélangées, les évaluateurs ne savaient pas

lesquelles étaient brutes et lesquelles étaient contrôlées. Nous pouvons voir dans les graphiques 14 et 15 ci-dessous les résultats de cette évaluation :



Graphique 14. Ranking des systèmes de TA



Graphique 15. Moyenne du ranking des systèmes

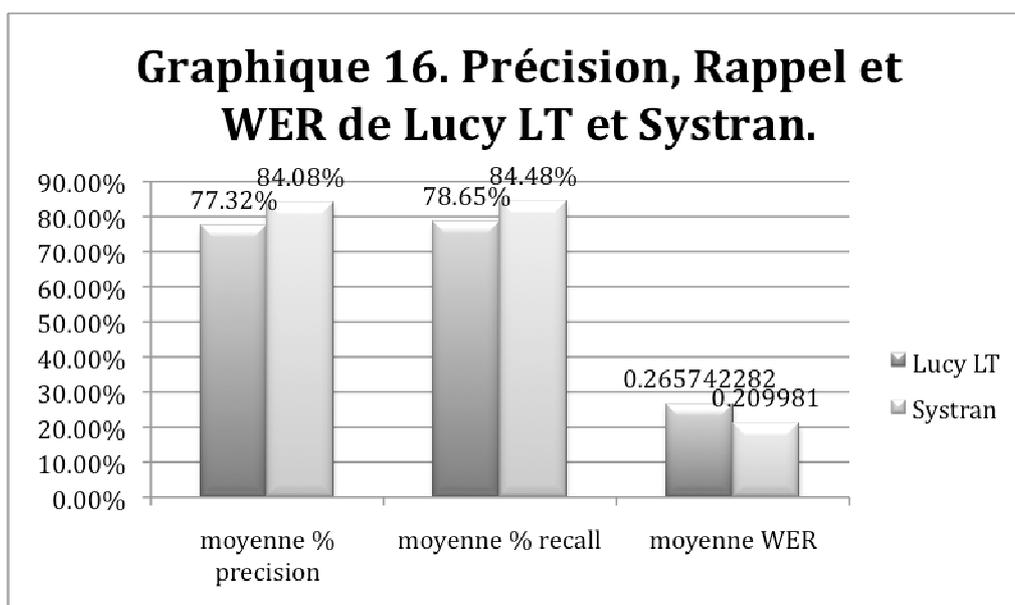
Les chiffres 1 et 2 représentent les systèmes 1 et 2, le chiffre 3 correspond à la valeur « équivalent ». En observant les résultats du « *ranking test* », il ressort que les

évaluateurs préfèrent Systran dans environ 57.5 % des cas, Lucy LT dans environ 38% des cas, et n'ont pas de préférence dans environ 4,5 % des cas, que ce soit pour la version brute ou contrôlée. Il est néanmoins intéressant de préciser que la seule personne du groupe de langue maternelle anglaise, et, incidemment, la seule à ne pas maîtriser le vocabulaire du ski, est également la seule à avoir légèrement préféré Lucy LT à Systran.

8.6.3 WER

8.6.3.1 Test 4 : WER

L'outil de Maria Georgescu a calculé la précision, le rappel et le WER de chaque phrase du sous corpus (H_Lucy) par rapport aux phrases du sous-corpus (H) dans un premier temps, puis, dans un second temps, des phrases de (H_Systran) par rapport au sous-corpus (H). Après avoir relevé les chiffres (voir annexe 26), nous avons établi la moyenne de précision, rappel et WER des sous-corpus (H_Lucy) et (H_Systran), comme nous pouvons le voir dans le graphique 16 ci-dessous :



Graphique 16. Précision, Rappel et WER de Lucy LT et Systran

Nous pouvons le voir clairement, Lucy LT présente en moyenne entre 75% et 80% de précision et de rappel par rapport au texte de référence, et un WER d'environ 25%. Les résultats, comme nous pouvons le voir, se révèlent très intéressants, compte tenu du fait que la sortie a été comparée à la version post-éditée du corpus (D142-342). Certes, les résultats de la précision, du rappel et du WER du corpus (H_Systran) sont prévisibles, puisque la seule différence entre (H_Systran) et (H) se situe au niveau de la post-édition.

Les résultats obtenus par (H_Lucy) sont d'autant plus intéressants que l'outil calcule le nombre d'insertions, de suppressions et de substitutions sans tenir compte du fait que certaines traductions alternatives ne sont pas forcément incorrectes.

8.7. Conclusion

En résumé, nous pouvons dire que l'évaluation comparative a montré des résultats très intéressants. Du point de vue purement lexical, Lucy LT semble avoir l'avantage sur Systran, en faisant preuve d'une plus grande précision terminologique. En revanche, les résultats du test 3 sont sans appel : les traductions de Systran sont perçues plus favorablement que celles de Lucy LT pour 4 personnes sur 5. En effet, Systran obtient environ 20% de résultats positifs de plus que Lucy LT. Nous pouvons voir également dans le calcul du WER (test 4) que Lucy LT parvient à un résultat plus qu'honorable par rapport à Systran. Si nous tenons compte du fait que la comparaison est légèrement biaisée puisque (H) est la version post-éditée de (H_Systran), le fait que Lucy LT produise un WER d'environ 5% supérieur à Systran (graphique 16) atteste de la qualité de Lucy LT.

Nous avons voulu savoir si le LC était indépendant des systèmes. Au vu des résultats du « *ranking test* », il semblerait que ce ne soit pas le cas. En effet, certaines des règles de contrôles choisies et/ou élaborées pour Systran n'ont pas eu le même effet sur Lucy LT, ce qui a selon toute probabilité amené les évaluateurs à préférer Systran. Pour aller plus loin dans la comparaison, il serait intéressant de voir à quel point les résultats puissent changer si nous adaptons spécialement des règles de contrôle à Lucy LT et refaisons les tests.

IX. Conclusion

Comme nous l'avons mentionné auparavant, l'objectif de notre travail consistait à tester sur un corpus de textes de type touristique en vue d'une traduction automatique multilingue les hypothèses suivantes :

- La personnalisation du système de TA doit permettre de produire une sortie du système conforme à la terminologie du texte source.
- Le langage contrôlé améliore la traduction automatique.

Pour le premier point, nous avons tout de suite constaté l'effort important et la quantité non négligeable de temps passé à rechercher la terminologie et à l'implémenter dans le système durant la phase de pré-édition. En effet, nous avons estimé que le temps consacré aux dictionnaires s'est élevé à plus ou moins 12 jours (non consécutifs). L'effort s'est toutefois révélé profitable puisqu'au vu des résultats obtenus pour l'évaluation (voir chapitre VI) les dictionnaires ont permis une nette amélioration de la sortie du système. Le fait de pouvoir ajouter des termes propres à la terminologie de notre corpus nous a donc permis d'améliorer la fidélité de la sortie au texte source, et ainsi répondre aux attentes de l'utilisateur final. Par exemple, nous pouvons voir cet impact dans les résultats du test 5 de la sous-section 6.2.3.2 : la partie B, qui correspond au corpus brut avec dictionnaire (B_Systran) a été mieux reçue par les experts que la partie C, tiré du corpus contrôlé sans dictionnaire (C_Systran).

En ce qui concerne le second point, l'effort de pré-édition s'est révélé moindre par rapport aux dictionnaires, et cela s'explique par le fait que notre corpus est composé d'un seul type de texte, avec une construction souvent répétitive et facilement identifiable. Certaines règles de contrôle sont venues à nous de manière assez évidente, telles que l'orthographe et la ponctuation à titre d'exemple, tandis que d'autres nous ont posé plus de problèmes comme la règle xii (sous-section 4.2.2) concernant les verbes réfléchis. En effet, nous n'avons pas réussi à éliminer tous les verbes réfléchis du corpus puisque nous n'avons pas pu recourir à chaque fois à un verbe synonyme qui ne soit pas pronominal. Si cette règle n'a pas eu d'effet probant dans le cadre de notre travail, elle pourrait prouver son efficacité au moment de la rédaction du texte source. Dans l'ensemble, les règles que nous avons sélectionnées et

appliquées ont permis une meilleure saisie du texte source par la machine ainsi qu'une meilleure compréhension du texte cible par l'utilisateur final. L'utilisation des règles et des dictionnaires a également engendré des conséquences positives puisque l'effort de post-édition a été considérablement réduit.

Ce dernier constat s'est vu confirmé dans le chapitre individuel (chapitre VII) sur la post-édition pour la langue italienne. Avant même le relevé manuel des erreurs, la sortie du système contrôlée (J) présentait 44% de phrases correctes contre 2.5% pour la sortie brute (I). L'application du LC et surtout la spécialisation des dictionnaires ont réduit le nombre de fautes de 789 à 209, sans parler du fait que leur degré de gravité s'est amoindri. De plus, en terme d'intelligibilité, le LC et les dictionnaires ont permis une augmentation de 10% par rapport au texte brut (corpus (I)) alors même qu'il reste des fautes. Notre type de texte se prête donc volontiers à la post-édition rapide, si un travail préalable tel que nous l'avons mené est réalisé.

Nous avons constaté l'effet manifeste du LC tout au long de travail, ce qui nous amène à la question suivante : Le LC s'avère-t-il efficace indépendamment des systèmes de TA ? C'est en partie pour répondre à cette question que nous avons mené l'évaluation comparative qui fait l'objet du chapitre VIII. Après examen des sorties, il apparaît que le LC est plus efficace lorsqu'il est adapté à un système de TA particulier : certaines règles que nous avons sélectionnées pour le système direct évolué Systran n'ont eu aucune incidence sur le système indirect de transfert Lucy LT. Au contraire, si nous prenons par exemple la règle n°ix (éviter l'utilisation des clitiques) nous avons constaté que Lucy LT n'analysait pas la phrase correctement, alors que si nous rajoutons le clitique, la traduction devenait juste. En conséquence, si les évaluateurs jugent les traductions de Systran meilleures que celles de Lucy LT, c'est selon toute probabilité parce que les règles de contrôle ont été spécialement sélectionnées pour Systran. Si nous nous référons aux résultats des tests de couverture lexicale uniquement, Lucy LT semble avoir un léger avantage sur Systran.

En résumé, tout le travail effectué jusqu'ici avait pour but de répondre à une question: « Ce type de texte se prête-t-il à la traduction automatique ? » Nous ne pouvons pas nous prononcer de manière catégorique mais nous pouvons donner quelques pistes :

- OUI car ce genre de texte présente une terminologie plutôt restreinte, peu de polysémie et une fois le(s) dictionnaire(s) compilé(s) le texte est suffisamment bon pour être post-édité. A l'exception des descriptions du ou des lieu(x), les informations

contenues dans ce type de texte sont suffisamment répétitives pour faciliter l'uniformisation du document.

- NON car le texte source, de sa nature publicitaire, se rapproche plus du langage littéraire (utilisation de figures de style, d'expressions et de locutions) que du langage technique supposé être plus simple et direct. L'application du LC, seul moyen de rendre ce type de texte compréhensible après une traduction automatique, tend à gommer l'aspect vendeur et convivial (le texte s'adresse directement au futur touriste) au profit de constructions neutres.

En somme, nous pouvons dire que ce type de texte se prête à la TA à condition de réaliser un travail de pré-édition important. L'idéal serait de rendre un texte suffisamment neutre pour la machine tout en étant assez vendeur pour l'être humain. Toutefois ces considérations devraient être soulevées au moment de la rédaction du texte source, ce qui allégerait considérablement le volume de pré et de post-édition. La discussion s'insère en réalité dans un plus grand débat: les limitations de la TA en général. Bernth et Gdaniec (2001) proposent d'agir à deux niveaux : la rédaction des textes et le développement des systèmes de TA. En effet, comme ils le disent eux-mêmes, le processus de traduction automatisée/automatique serait grandement facilité si les rédacteurs prêtaient une attention particulière au fait d'écrire des phrases grammaticales, avec une orthographe correcte, une ponctuation adéquate, etc. Le second niveau concerne les systèmes de TA : ils devraient être capables de reconnaître les ambiguïtés et ainsi de proposer plusieurs traductions à choix. Le travail de pré-édition et de post-édition s'en trouverait ainsi facilité. Il n'est pas sûr que la TA brute puisse un jour égaler la TH, puisque comme nous l'avons vu dans la sous-section 2.2.4, les problèmes principaux des systèmes de TA proviennent du manque de contexte et de connaissance du monde, et donc de la capacité à utiliser ces deux éléments pour résoudre les problèmes d'ambiguïtés (Bernth & Gdaniec 2001). Si un jour les systèmes intègrent cela, la traduction automatique prendra un nouveau tournant.

X. Bibliographie

AECMA Document. *Simplified English Standard, a guide for the preparation of Aircraft Maintenance Documentation in the International Aerospace Maintenance Language*, 1995.

ALLEN, J. « Post-editing » in *Computers and Translation: A Translators Guide* de Harold Somers, Benjamins Translation Library, Amsterdam, 2003, pp.297-317.

ALLEN, J. & HOGAN, C. « Toward the development of a post-editing module for Machine Translation raw output: a new productivity tool for processing controlled language » in *CLAW 2000: Third International Controlled Language Applications Workshop*, Seattle, Washington, 29-30 April 2000, pp.62-71.

ARNOLD, D. [et al.], *Machine translation: an introductory guide*, NCC Blackwell, Manchester, Oxford, Cambridge, 1994, 240 p.

BARTHE, K., « GIFAS Rationalised French: Designing one Controlled Language to Match Another » in *CLAW 98: proceedings of the second International Workshop on Controlled Language Applications: 21-22 May 1998*, Language Technologies Institute, Carnegie Mellon University, Pittsburg, Pennsylvania, pp.87-102.

BERNTH, A. & GDANIEC, C., « MTranslatibility » in *Machine Translation*, 2001, vol. 16, n°3, pp. 175-218.

BIRCHER, Beatrice, « Systran versus Google translate (GT) – the performance of two MT Systems in Query Translation », Ecole de Traduction et d'Interprétation, Université de Genève, 2009, 110 p.

BOUILLON, P. & CLAS, A., *La Traductique, Études et recherches de traduction par ordinateur*, Les Presses de l'Université de Montréal, Québec, 1993, 507 p.

BOWKER, L. & PEARSON, J., *Working with specialized language: a practical guide to using corpora*, London/New York, Routledge, 2002, 256 p.

CARL, M. « Introduction à la traduction guidée par l'exemple » in *TALN-RECITAL 2003 : 10^e conférence sur le traitement automatique des langues naturelles*, 11-14 juin 2003, Batz sur Mer, 15 p.

DABBADIE, M. et al. « A Hands-On Study of the Reliability and Coherence of Evaluation Metrics » in *LREC-2002: Third International Conference on Language Resources and Evaluation*. Workshop: Machine translation evaluation: human evaluators meet automated metrics, Las Palmas Canary Islands, 27 May 2002, pp.8-16.

De ALMEIDA, G. & O'BRIEN, S., « Analysing post-editing performance: correlations with years of translation experience » in *EAMT 2010: Proceedings of the 14th Annual conference of the European Association for Machine Translation*, Saint-Raphaël, France, 27-28 May 2010, 8p.

DESMET, I., « Terminologie, culture et société. Éléments pour une théorie variationniste de la terminologie et des langues de spécialité » in *Cahiers du Rifal*, 2007, pp.3-13.

DRUETTA, R. « Types de communication et types de textes », 18 p., consulté le 22 novembre 2011 : http://www.francaisunivers.unito.it/documents/types_druetta.pdf

EAGLES evaluation work group, *EAGLES Evaluation of Natural Language Processing Systems, Final report*, Center for Sprogteknologi, Denmark, 1996.

FEDERMANN, C., « How can we measure machine translation quality? » in *Proceedings of the Tralogy Conference*, Paris, France, Mars 2011.

FERNANDEZ GUERRA, A., *Machine Translation : capabilities and limitations*, Lengua Inglesa Universitat de Valencia, Valencia, 2000, 159 p.

FLANAGAN, M., « Error Classification for MT Evaluation », in *Proceedings of AMTA 94*, 1994, pp. 65-72.

FURLANI, D., *Translation technologies and medical translation*, Université de Genève, Ecole de Traduction et d'Interprétation, Genève, 2009, 104 p.

GERBER, L. & YANG, J., « SYSTRAN MT Dictionary Development. Machine Translation: Past, Present and Future » in *Proceedings of Machine Translation Summit VI*, Octobre/Novembre 1997. San Diego, USA, 1997, pp. 211-218.

GIFAS (Groupement des Industries Françaises Aéronautiques et Spatiales), Guide du rédacteur : Français Rationalisé, Paris, 1996.

GREEN, R., « The MT errors which cause most trouble to posteditors », in *Veronica Lawson*

(ed.) *Practical Experience of Machine Translation*, Amsterdam, 1982, pp. 101-104.

GUZMAN, R. « Manual MT Post-editing: if it's not broken, don't fix it! » in *Translation Journal*, vol. 11, No. 4, 2007, consulté le 15 novembre 2011: <http://translationjournal.net/journal//42mt.htm>

HUTCHINS, J., « Machine Translation : past, present, future » in *Ellis Horwood Series in Computers and their Applications*, Chichester, Ellis Horwood, 1986, 382p.

HUTCHINS, J. & SOMERS, H., *An introduction to machine translation*, Academic Press, London, San Diego, 1997, 362 p.

HUTCHINS, J. « Why computers do not translate better » in *Translating and the Computer 13: the theory and the practice of machine translation - a marriage of convenience?* Papers presented at a conference jointly sponsored by Aslib, The Association for Information Management; Aslib Technical Translation Group; Institute of Translation and Interpreting, 18-19 November 1991. London: Aslib, 1992. p.3-16. Reprinted in: **Aslib Proceedings** 44 (10), October 1992, p.351-359.

HUTCHINS, J., « The development and use of machine translation systems and computer-based translation tools » in *International Symposium on Machine Translation and Computer Language Information Processing*, 26-28 June 1999, Beijing, China, pp. 1-16.

HUTCHINS, J., «Machine Translation: History» in *Encyclopedia of Language & Linguistics, Second Edition*, volume 7, 2000, pp. 375-383.

HUISJENS, W-O., «Controlled Language – An Introduction», in *CLAW 98: proceedings of the second International Workshop on Controlled Language Applications: 21-22 May 1998*, Language Technologies Institute, Carnegie Mellon University, Pittsburg, Pennsylvania, pp.1-15.

HUANG, H., «Researches on MT and language information processing» in *Translingual Europe*, Berlin, Allemagne, 7 juin 2010, 37p.

KITTREDGE, R., «The Significance of Sublanguage for Automatic Translation» in *Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Colgate University, Hamilton, New York, August 14-16 1985, pp.154-166.

LALAUDE, M., LUX, V. & REGNIER-PROST, S., « Modular Controlled Language design » in *CLAW 98: proceedings of the second International Workshop on Controlled Language Applications: 21-22 May 1998*, Language Technologies Institute, Carnegie Mellon University, Pittsburg, Pennsylvania, pp. 107-112.

LEHMANN, S., «Acrolinx IQ Suite, Quality Assurance and Controlled Languages in Technical Documentation» présentation, Genève, 2007.

LexShop 2.2 User Manual, consulté le 14 décembre 2011.

LexShop : Introduction à LexShop – Lexique de Transfert – Tests et Actions, 2009, 14 p.

LOFFLER-LAURIAN, A.-M., *La Traduction automatique*, Presses universitaires du Septentrion, 1996, 156 p.

MCENERY, K., et al., « Evaluating machine translation output for an unknown source language: report of an ISLE-based investigation » in *MT Summit VIII*, Santiago de Compostela, Spain, 18-22 September 2001, pp.15-20.

MILLER, K. , et al. « Evaluating machine translation output for an unknown source language: report of an ISLE-based investigation » in *MT Summit VIII*, Santiago de Compostela, Spain, 18-22 September 2001, pp.15-20.

MILLER, K & VANNI, M., « Scaling the ISLE Taxonomy : Development of Metrics for the Multi-Dimensional Characterisation of Machine Translation Quality » in *MT Summit VIII: Machine Translation in the Information Age*, Proceedings, Santiago de Compostela, Spain, 18-22 September 2001; pp. 229-233.

MILLER, K. & VANNI, M., « Scaling the ISLE Framework : Validating Tests of Machine Translation Quality for Multi-Dimensional Measurement » in *MT Summit VIII: Machine Translation in the Information Age*, Proceedings, Santiago de Compostela, Spain, 18-22 September 2001; pp. 21-27.

MUEGGE, U., « Règles pour la traduction automatique » in *Traduction automatique, langage contrôlé, normes de traduction*, consulté le 02 juin 2011, <http://www.muegge.cc/langage-controle.htm>

National Institute of Standards and Technology, Linguistic Data Consortium. Post Editing Guidelines for GALE Machine Translation Evaluation, Version 3.0.2, 2007, consulté le 6 décembre 2011:

http://projects ldc.upenn.edu/gale/Translation/Editors/GALEpostedit_guidelines-3.0.2.pdf

O'BRIEN, S., «Controlling Controlled English - An Analysis of Several Controlled Language Rule Sets» in *Proceedings of the Joint Conference combining the 8th International Workshop of the European Association for Machine Translation and the 4th Controlled Language Applications Workshop, (EAMT-CLAW 03)*, Dublin City University, Ireland, 2003, pp. 105-114.

O'BRIEN, S., « Controlled Language and post-edition » in *Multilingual* , Octobre/Novembre 2006, pp.17-19.

O'BRIEN, S., « Introduction to Post-Editing: Who, What, How and Where to Next? » in *Ninth Conference of the Association for Machine Translation in the Americas*, Denver, USA, 31 Octobre – 4 Novembre 2010.

O'BRIEN, S. & ROTURIER, J., « How Portable are Controlled Language Rules ? A comparison of Two Empirical MT Studies » in *Machine Translation Summit XI* 10-14 Septembre 2007, Copenhagen, Denmark, pp. 345-352.

PEARSON J., *Terms in context*, John Benjamins B.V., Amsterdam, 1998, 249 p.

ROTURIER, J., « Assessing a set of Controlled Language rules : Can they improve the performance of commercial Machine Translation Systems ? » in *Translating and the Computer* 26, Novembre 2004, pp. 1-13.

RYAN, R, « Les langues contrôlées : une valeur ajoutée pour le traducteur » in *Traduire*, 220, 2009, pp. 57-67.

SCHNEIDER, T., «The METAL system» in *MT Summit III* Proceedings, July 1-4, 1991, Washington, DC, USA, pp. 41-44.

SENEILLART, J. *et al.*, «New generation SYSTRAN translation system» in *MT Summit VIII*, Santiago de Compostela, Espagne, 2001, pp. 311-316.

SENEILLART, J. *et al.*, « SYSTRAN Intuitive Coding Technology » in *MT Summit IX*, New Orleans, Louisiana, USA, 2003, pp.346-353.

SHEI, C., «Teaching Mt through Pre-editing : Three Case Studies » in *Sixth EAMT Workshop "Teaching machine translation"*, 14-15 Novembre 2002, UMIST, Manchester, England, pp.89-98.

SOMERS, H. *et al.*, *Computer and Translation : A Translator's Guide*, Amsterdam, Philadelphie, J. Benjamins, 2003, 349 p.

VILAR, D. & BANCHS, R. E « Human Evaluation of Machine Translation Through Binary System Comparisons » in *Proceedings of the Second Workshop on Statistical Machine Translation*, Association for Computational Linguistics, 2007, Prague, pp.96-103.

WHITELOCK P. & KILBY K., *Linguistic and Computational Techniques in Machine Translation System Design*, Second Edition, UCL Press, Londres, 1995, 208 p.

WOODSEND, K.& LAPATA, M., « WikiSimple: Automatic Simplification of Wikipedia Articles» in *Proceedings of the National Conference on Artificial Intelligence*, 7-11 août 2011, San Francisco, USA, pp.927-932.

YUSTE-RODRIGO, E. & BRAUN-CHEN, F., « Comparative Evaluation of the Linguistic Output of MT Systems for Translation and Information Purposes » in *MT Summit VIII*, Santiago de Compostela, Spain, 18-22 September 2001. Workshop on MT Evaluation; pp.9-14.

Index des figures

Figure 1. Les systèmes de TA.....	17
Figure 2. Exemple de mots autorisés (en majuscule) et non autorisés (en minuscule).....	31
Figure 3. Règles lexicales du GIFAS.....	32
Figure 4. Exemples de Règles de style du GIFAS	33
Figure 5. Processus de traduction d'un système direct évolué – Systran (Hutchins : 1986) ..	41
Figure 6. L'architecture d'Acrolinx IQ.....	47
Figure 7. Les erreurs les plus fréquentes relevées par Acrolinx.....	47

Index des graphiques

Graphique 1. Progression de la compréhensibilité des experts italiens.	109
Graphique 2. Progression de la compréhensibilité des experts anglais.....	110
Graphique 3. Niveau de lisibilité des phrases tests pour l’italien.	111
Graphique 4. Niveau de lisibilité des phrases tests pour l’anglais.	111
Graphique 5. Progression de la compréhensibilité des experts italiens avec Acrolinx.....	113
Graphique 6. Progression de la compréhensibilité des experts anglais avec Acrolinx.	114
Graphique 7. Pourcentage d’erreurs du corpus (I).	127
Graphique 8. Pourcentages du classement des erreurs selon le critère d’intelligibilité.....	131
Graphique 9. Pourcentages du classement des erreurs selon le critère de degré d’amélioration.	134
Graphique 10. Pourcentages d’erreurs du corpus (J).....	141
Graphique 11. Classement des erreurs selon le critère d’intelligibilité pour le corpus (J) ...	142
Graphique 12. Classification des erreurs selon le critère de degré d’amélioration pour le corpus (J)	144
Graphique 13. Pourcentage d’erreurs corpus (I) versus corpus (J).....	145
Graphique 14. Ranking des systèmes de TA.....	161
Graphique 15. Moyenne du ranking des systèmes	161
Graphique 16. Précision, Rappel et WER de Lucy LT et Systran.....	162

Index des tableaux

Tableau 1. Exemple de Basic English	27
Tableau 2. Exemple de règle grammaticale du SE.....	33
Tableau 3. Exemple de cellules du corpus avec ses trois parties.	53
Tableau 4. Fonctions et caractéristiques d'organisation du corpus.....	55
Tableau 5. Exemple de segments répétitifs du corpus.	56
Tableau 6. Caractéristiques lexicales et grammaticales pour les parties injonctives du corpus.	56
Tableau 7. Caractéristiques observables pour les parties descriptives et explicatives du corpus.	58
Tableau 8. Exemple de sortie de système brute – nom propre	64
Tableau 9. Sorties de systèmes avec et sans dictionnaires utilisateurs.	66
Tableau 10 . Entrées et sorties de systèmes avant et après application de la règle i.....	69
Tableau 11. Entrées et sorties de systèmes avant et après application de la règle i – ex. 2 ..	71
Tableau 12. Entrées et sorties de système avant et après application de la règle ii.	72
Tableau 13. Entrées et sorties de système avant et après application de la règle iii.	73
Tableau 14. Entrées et sorties de système avant et après application de la règle iv.....	74
Tableau 15. Entrées et sorties de système avant et après application de la règle v.	76
Tableau 16. Entrées et sorties de système avant et après application de la règle vi.....	78
Tableau 17. Entrées et sorties de système avant et après application de la règle vii – ex. 1.	80
Tableau 18. Entrées et sorties de système avant et après application de la règle vii – ex2. ..	81
Tableau 19. Entrées et sorties de système avant et après application de la règle viii.....	84
Tableau 20. Entrées et sorties de système avant et après application de la règle ix.....	85
Tableau 21. Entrées et sorties de système avant et après application de la règle x.	86
Tableau 22. Entrées et sorties de système avant et après application de la règle xi.....	87
Tableau 23. Entrées et sorties de système avant et après application de la règle xii.....	88

Tableau 24. Récapitulatif des corpus	94
Tableau 25. Récapitulatif des tests, des critères et des corpus	104
Tableau 26. Couverture lexicale technique du corpus.....	105
Tableau 27. Couverture lexicale incorrecte du corpus.....	106
Tableau 28. Pourcentage d'amélioration de la sortie du système par le LC.....	107
Tableau 29. Résultats pour l'italien Tableau 30. Résultats pour l'anglais.....	108
Tableau 31. Résultats pour l'italien Tableau 32. Résultats pour l'anglais.....	113
Tableau 33. Récapitulatif des corpus.	121
Tableau 34. Tableau des erreurs de la sortie du système (I).....	126
Tableau 35. Classement des erreurs pour le critère d'intelligibilité.	130
Tableau 36. Classement des erreurs selon le critère de degré d'amélioration.	133
Tableau 37. Tableau des erreurs de la sortie du système (J).	139
Tableau 38. Classement des erreurs selon le critère d'intelligibilité pour le corpus (J).....	141
Tableau 39. Classement des erreurs selon le critère de degré d'amélioration pour le corpus (J).	143
Tableau 40. Comptabilisation des modifications du corpus (I) et (J).	147
Tableau 41. Récapitulatif des sorties de Systran et de Lucy LT.....	150
Tableau 42. Les sous-corpus.....	151
Tableau 43. Pourcentage de couverture lexicale incorrecte des sorties du système.....	159
Tableau 44. Pourcentage de termes techniques correctement traduits.	160

Annexes

Pour obtenir les annexes, nous vous invitons à nous contacter aux adresses suivantes :

celine.chiodi@gmail.com

alexandra.dupont@ymail.com