

Je ne l'obtiens pas vraiment - I don't really get it  
Faiblesses de fonctionnement de deux outils de traduction automatique

Mémoire de maîtrise  
en philologie romane  
Institut des langues  
romanes et classiques  
Université de Jyväskylä

Ritva Tiihonen  
Novembre 2001

Tiedekunta HUMANISTINEN	Laitos ROMANISTEN JA KLASSISTEN KIELTEN LAITOS
Tekijä Ritva Tiihonen	
Työn nimi JE NE L'OBTIENS PAS VRAIMENT – I DON'T REALLY GET IT FAIBLESSES DE FONCTIONNEMENT DE DEUX OUTILS DE TRADUCTION AUTOMATIQUE	
Oppiaine Romaaninen filologia	Työn laji pro gradu
Aika Marraskuu 2001	Sivumäärä 64 + 2
<p><b>Tiivistelmä - Abstract</b></p> <p>Tutkimuksessa esitellään kaksi käännöskonetta, Babel Fish ja Reverso, ja selvitetään niiden toimivuutta tarkastelemalla niiden tekemiä käännöksiä englannista ranskaan ja ranskasta englantiin. Työn pääasiallisena tarkoituksena on selvittää, millaisia ongelmia näiden kielten ominaisuudet voivat aiheuttaa konekäännöksessä ja miten koneet ratkaisevat nämä ongelmatilanteet. Korpus koostuu useista lyhyistä, erityyillisistä teksteistä, jotka on otettu enimmäkseen sanomalehtien www-sivuilta ja käännetty molemmilla koneilla.</p> <p>Tutkimus osoittaa että kääntäminen ei ole suoraviivaista ”koodinpurkamista”, vaan että kielten rakenteellisista eroista johtuen käännettävien lauseiden syntaksiin täytyy usein puuttua muuttamalla sanajärjestystä (esim. adjektiiviattribuuttien paikka suhteessa substantiiviin ei yleensä ole sama englannissa ja ranskassa) tai sovittamalla aikamuotojen käyttöä kohdekielelle tyypilliseksi (erityisesti menneen ajan aikamuodot eivät vastaa toisiaan englannissa ja ranskassa).</p> <p>Toinen vaikea alue on monimerkityksisten sanojen ja rakenteiden tulkinta. Molemmat koneet tulkitsevat väärin monia homografeja, erityisesti imperatiiveja. Englannin pitkien nominiketjujen viittaussuhteet eivät ole aina itsestäänselviä, minkä vuoksi niiden ohjelmallinen analysointi on vaikeaa. Myös pronominit, erityisesti ranskan, ovat ongelmallisia analysoida, koska niiden korrelaattien selvittämiseksi täytyisi voida ottaa useamman lauseen laajuinen konteksti huomioon, mihin nämä käännöskoneet eivät pysty.</p> <p>Lisäksi tekstin koneellinen käsittely tuo mukanaan tiettyjä teknisiä ongelmia, jotka eivät välttämättä riipu käsitellyistä kielistä. Näitä ovat joustamattomuus oikeinkirjoituksen suhteen ja erikoismerkkien (esim. ranskan œ-ligatuurin) koodaukseen ja tulkintaan liittyvät ongelmat.</p>	
Asiasanat traduction automatique, ambiguïtés, traitement automatique de langues	
Säilytyspaikka	
Muita tietoja	

# Table des matières

<b>1. Introduction .....</b>	<b>3</b>
1.1. But, méthode et corpus.....	3
1.2. Traduction automatique .....	4
1.2.1 <i>Histoire de la traduction automatique</i> .....	4
1.2.2 <i>Une répartition possible des systèmes de traduction</i> .....	6
1.2.3 <i>Traduction automatique sur Internet</i> .....	8
1.2.3.1 Remarques préliminaires.....	8
1.2.3.2 Babel Fish .....	8
1.2.3.3 Reverso.....	9
1.3. Ambiguïtés .....	10
1.3.1 <i>Remarque préliminaire</i> .....	10
1.3.2 <i>Homographes</i> .....	10
1.3.3 <i>Ambiguïtés syntaxiques</i> .....	11
1.3.4 <i>Anaphoriques</i> .....	12
1.4. Différences entre le français et l'anglais.....	12
1.4.1 <i>Remarques préliminaires</i> .....	12
1.4.2 <i>Les temps du passé en français et en anglais</i> .....	13
1.4.3 <i>La modification des noms en anglais</i> .....	15
<b>2. Analyse.....</b>	<b>17</b>
2.1. Remarques préliminaires.....	17
2.2. Orthographe .....	17
2.2.1 <i>Remarques préliminaires</i> .....	17
2.2.2 <i>La ponctuation</i> .....	17
2.2.3 <i>L'orthographe du français</i> .....	18
2.3. Ambiguïtés .....	19
2.3.1 <i>Ambiguïtés en anglais</i> .....	19
2.3.1.1 Homographes .....	19
2.3.2 <i>Ambiguïtés en français</i> .....	23
2.3.2.1 Homographes .....	23
2.3.2.2 Pronoms .....	26
2.4. Lexies figées .....	28
2.4.1 <i>Remarques préliminaires</i> .....	28
2.4.2 <i>Lexies figées de l'anglais</i> .....	28
2.4.2.1 Verbes à particule ( <i>phrasal verbs</i> ).....	28

2.4.2.2	Mots composés à tiret.....	30
2.4.2.3	Locutions.....	31
2.4.3	<i>Lexies figées du français</i> .....	33
2.4.3.1	Locutions.....	33
2.4.3.2	Structures grammaticales .....	35
2.5.	<b>L'ordre des mots</b> .....	36
2.5.1	<i>L'ordre des mots en anglais</i> .....	36
2.5.1.1	Noms modifiés par d'autres noms.....	36
2.5.1.2	Adjectifs épithètes.....	39
2.5.1.3	La place des prépositions dans la phrase.....	41
2.5.2	<i>L'ordre des mots en français</i> .....	43
2.5.2.1	Compléments des noms.....	43
2.5.2.2	Autres compléments.....	45
2.5.2.3	D'autres cas.....	46
2.6.	<b>Les temps</b> .....	47
2.6.1	<i>Les temps de l'anglais</i> .....	47
2.6.1.1	Remarques préliminaires.....	47
2.6.1.2	La traduction du prétérit anglais.....	47
2.6.1.3	Les temps composés anglais.....	49
2.6.2	<i>Les temps du français</i> .....	51
2.6.2.1	Les temps du passé.....	51
2.6.2.2	Impératif.....	53
2.7.	<b>Noms propres</b> .....	54
2.7.1	<i>Remarques préliminaires</i> .....	54
2.7.2	<i>Fautes dues aux noms propres dans les textes en anglais</i> .....	54
2.7.3	<i>Fautes dues aux noms propres dans les textes en français</i> .....	55
2.8.	<b>Différences quantitatives entre le français et l'anglais</b> .....	56
2.8.1	<i>Remarques préliminaires</i> .....	56
2.8.2	<i>Omissions en anglais</i> .....	56
2.8.3	<i>Expressions françaises contenant moins de mots que celles anglaises</i> .....	57
2.8.4	<i>La négation en français</i> .....	58
<b>3.</b>	<b>Conclusion</b> .....	<b>60</b>
<b>4.</b>	<b>Bibliographie</b> .....	<b>62</b>
<b>5.</b>	<b>Appendices</b> .....	<b>65</b>

# 1. Introduction

## 1.1. But, méthode et corpus

Le but de ce travail est de présenter deux outils de traduction automatique accessibles sur le World Wide Web (Babel Fish de Systran<sup>1</sup> et Reverso de Softissimo<sup>2</sup>) et d'étudier certaines faiblesses dans leur fonctionnement. Le corpus consiste, d'une part, en divers textes en anglais et de leurs traductions en français générées par les deux machines, et d'autre part en textes en français et de leurs traductions en anglais.

Babel Fish (=BF) reçoit plus d'un million de demandes de traduction par jour<sup>3</sup>. C'est peut-être l'une des plus connues parmi les machines de traduction en accès gratuit sur Internet, car elle est gérée par Altavista, un moteur de recherche populaire. Reverso, l'outil de Softissimo est accessible depuis plusieurs adresses web<sup>4</sup>.

Dans le choix des documents du corpus, la variation stylistique a prévalu comme critère, les textes anglais du corpus contenant la page web contenant des conseils d'emploi de BF (=FAQ dans les exemples), deux chroniques humoristiques (=HAL et SPI) et plusieurs articles de journal (=GEN, OIL, TIM et TOR), ces derniers étant conseillés comme tests textuels par les rédacteurs de la FAQ de BF. Les textes français contiennent également un mode d'emploi, celui de Reverso (=AID), deux chroniques (=BOY et TEL) et des articles de journaux (=BIO, PRE, SHA et SIB). Le total des mots du corpus est de 12 390 mots environ (5970 mots en anglais, 6420 en français), sans compter les mots des traductions.

Les textes ne sont pas très longs, compte tenu de la capacité de traduction de ces outils (à peu près 800 mots à la fois pour BF, à peu près 100 mots pour Reverso, quand il s'agit du texte). Tous les documents étudiés sont ou ont été accessibles sur

---

<sup>1</sup> Accessible à <http://babelfish.altavista.com>.

<sup>2</sup> Accessible à <http://www.reverso.net>.

<sup>3</sup> Anon. 2000a.

<sup>4</sup> Notamment <http://www.softissimo.com>, <http://www.translate-online.com> et <http://www.reverso.net>.

Internet (voir la bibliographie pour plus de détails). La plupart des textes ont donc été traduits dans le format HTML (page web), mais comme certains articles de journaux ne sont accessibles gratuitement que pendant le jour de leur publication, une partie des textes a été traduite en format texte. Les différences de format n'auront pas influencé la qualité de la traduction.

## 1.2. Traduction automatique

### 1.2.1 Histoire de la traduction automatique

L'un des «pères» de la traduction automatique est l'Américain Warren Weaver<sup>5</sup>, un mathématicien qui pensait que le russe était simplement de l'anglais écrit avec des symboles bizarres. Pour le traduire, il suffisait donc de décoder ces symboles<sup>6</sup>, comme

Weaver l'avait déjà vu faire avec les messages codés militaires pendant la Deuxième Guerre mondiale. Les premiers systèmes produisirent des traductions mot à mot ou à peine plus évoluées<sup>7</sup>. Les mauvaises connaissances en linguistique et l'optimisme naïf des informaticiens travaillant sur la TA étaient cependant les raisons pour laquelle le développement de la TA ne fut pas abandonné. Très vite, pourtant, les chercheurs dans le domaine comprirent que traduire n'était pas remplacer un mot russe par un mot anglais, mais qu'il s'agissait d'une opération bien plus complexe.

La TA prit de l'ampleur grâce à la guerre froide. Les Américains, ayant peur des Soviétiques, investirent des millions de dollars dans son développement pour pouvoir traduire le plus vite possible un grand nombre de textes scientifiques écrits en russe.<sup>8</sup>

En Europe, la politique linguistique de la Communauté Européenne, qui demande que les documents officiels soient traduits dans toutes les langues des états membres, a créé un besoin toujours croissant de traductions. Dans les années 1970, les

---

<sup>5</sup> Weaver (1894-1978) fut le vice-président de la Fondation Rockefeller dans les années 1940, Anis 1994 : 111.

<sup>6</sup> Admiral 1994 : 7.

<sup>7</sup> Anis 1994 : 112.

<sup>8</sup> Anis 1994 : 112.

institutions de la CEE (la Communauté Économique Européenne) achetèrent les systèmes anglais – français, français – anglais et anglais – italien développés par SYSTRAN (*System Translation*). Ces systèmes, améliorés, sont utilisés depuis 1981 pour traduire des documents officiels.<sup>9</sup>

Les principales raisons de l'intérêt pour le développement de la TA seraient donc militaires ou scientifiques (russe – anglais), politiques (entre les langues européennes) et avant tout, financières, la TA permettant un important gain de temps et d'argent.

Encore aujourd'hui, plusieurs projets de développement de la TA et de la Traduction Assistée par Ordinateur (=TAO<sup>10</sup>) sont en cours, aussi bien des projets universitaires (University of Maryland, New Mexico State University, Carnegie-Mellon, EAMT) que des projets commerciaux (Globalink, Lingua Vision, Systran)<sup>11</sup>. L'idée originale de FAHQT (*Fully Automatic High Quality Translation* – traduction de grande qualité entièrement automatisée), qui remplacerait la traduction dite humaine, a été abandonnée en faveur de systèmes qui demandent un certain travail de prédiction du texte source et de postédition du texte cible<sup>12</sup>. Cependant, certains systèmes de traduction automatique sont utilisés quotidiennement et avec beaucoup de succès. En plus de SYSTRAN susmentionné, il y a lieu de noter METEO, qui traduit depuis 1977 des bulletins météorologiques de l'anglais vers le français au Canada, et SPANAM avec ENGSPAN, qui sont utilisés depuis 1976 par l'Organisation panaméricaine de la Santé pour traduire toute sorte de documents entre l'anglais et l'espagnol<sup>13</sup>.

---

<sup>9</sup> Anis 1994 : 116-117.

<sup>10</sup> Ce sigle peut venir aussi des mots *Traduction Automatisée par Ordinateur*, la signification étant plus ou moins la même : la traduction n'est pas entièrement automatisée mais demande une intervention humaine. Ce terme est utilisé notamment par Boitet et Blanchon du GETA. Boitet - Blanchon 1994 : 22.

<sup>11</sup> Ball 1999 : 1.

<sup>12</sup> Ladmiral 1994 : 7-8.

<sup>13</sup> Arnold *et al.* 1994 : 7 ; Anis 1994 : 117.

### 1.2.2 Une répartition possible des systèmes de traduction

Les systèmes de traduction automatique peuvent être répartis de la manière suivante<sup>14</sup> :

1) La traduction directe (une étape) : la syntaxe et le vocabulaire de la langue source ne sont analysés qu'au point où il est possible de choisir le bon ordre de mots et les expressions correspondantes dans la langue cible. Les premiers systèmes utilisaient cette méthode, qui rappelle l'idée de Warren Weaver, selon laquelle traduire c'est décoder.

2) La traduction avec une *interlingua* (deux étapes) : le texte source est analysé et représenté en une *interlingua*, qui peut être une langue artificielle ou simplement un ensemble de traits sémantiques universels. Puis le texte en *interlingua* est synthétisé en langue cible (voir fig. 1.).

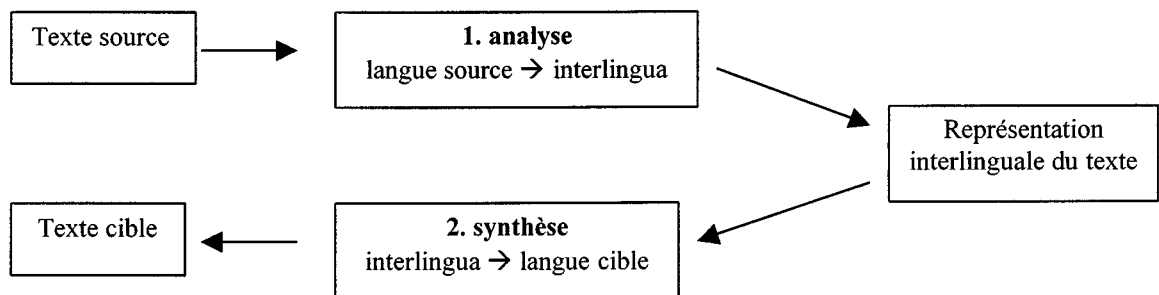


Fig. 1. Traduction par interlingua<sup>15</sup>

Cette méthode est particulièrement économique quand il s'agit de traduire un texte vers plusieurs langues différentes, car il suffit d'analyser le texte source une seule fois. Cependant, cette méthode est difficile à réaliser, car il faut trouver des *primitifs*, des représentations abstraites (indépendantes des langues) qui soient en concordance avec toutes les différentes manières possibles de décrire les concepts et les phénomènes de la réalité<sup>16</sup>.

<sup>14</sup> Anis 1994 : 112-114. Cette répartition des systèmes de traduction provient de Hutchins 1992 = Hutchins – Somers, *An Introduction to Machine Translation*. S.I, que nous n'avons malheureusement pas pu consulter.

<sup>15</sup> Une adaptation de Arnolds *et al.* 1994 : 79.

<sup>16</sup> Le choix des traits sémantiques pertinents pour la description des concepts n'est pas évident, car chaque langue décrit la réalité de sa propre manière. Par ex. le français distingue *le fleuve* (un cours



3) La traduction par transfert (trois étapes) : le texte source est transformé en représentations de transfert de la langue source, qui sont transformées en représentations de transfert de la langue cible, qui, elles, sont synthétisées en texte en langue cible (voir fig.2.).

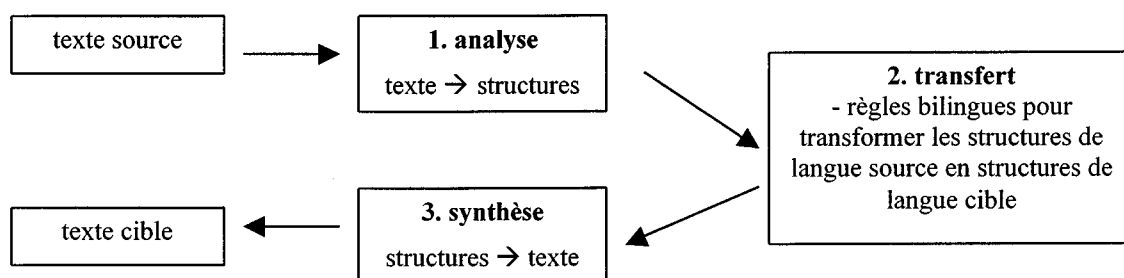


Fig. 2. Traduction par transfert<sup>17</sup>

L'avantage de ce système est que les grammaires pour l'analyse et la synthèse peuvent être développées indépendamment l'une de l'autre et peuvent également être utilisées avec n'importe quelle autre langue. Par contre, chaque paire de langues demande un module de transfert, une grammaire bilingue. Contrairement aux grammaires des systèmes de traduction directe, les grammaires de l'analyse et de la synthèse sont de vraies grammaires monolingues, si bien que la grammaticalité des traductions est bien supérieure à celle des traductions produites par les autres systèmes<sup>18</sup>.

La méthode interlinguale est davantage un sujet intrigant de recherche scientifique qu'une vraie solution pratique pour la traduction automatique. Néanmoins, un système de traduction par transfert peut contenir des éléments propres à un système

---

d'eau important qui se jette dans la mer) de *la rivière* (un cours d'eau moins important) tandis qu'en anglais le mot *river* désigne les deux, cf. Arnold *et al.* 1994 : 76, 78.

<sup>17</sup> Une adaptation de Arnold *et al.* 1994 : 68.

<sup>18</sup> Arnold *et al.* 1994 : 67.

interlingual. Plus les éléments de la langue sont analysés profondément et qu'ils soient représentés d'une manière abstraite, plus la phase de transfert devient facile.<sup>19</sup>

### 1.2.3 Traduction automatique sur Internet

#### 1.2.3.1 Remarques préliminaires

Internet offre une variété époustouflante de services de traduction automatique. Il y a lieu remarquer, pourtant, que tous les services sur Internet présentés sous le nom de traduction automatique ne sont pas des services de traduction proprement dits. Souvent il s'agit d'un simple dictionnaire électronique qui ne fait que de la traduction mot à mot. Les deux services que nous avons choisis comme sujets de notre étude sont parmi ceux qui méritent, malgré leurs défauts, d'être appelés des outils de traduction.

#### 1.2.3.2 Babel Fish

La machine à traduire appelée Babel Fish<sup>20</sup> est accessible par Internet depuis le 9 décembre 1997. C'était le premier service de traduction automatique offert au public sur Internet<sup>21</sup>. D'utilisation gratuite, elle traduit entre l'anglais et neuf autres langues (le français, l'allemand, l'italien, l'espagnol, le portugais, le japonais, le coréen et le chinois) dans les deux sens<sup>22</sup>. En plus de cela, elle traduit du russe vers l'anglais et encore entre le français et l'allemand dans les deux sens. D'autres paires de langues sont constamment en voie de développement<sup>23</sup>.

---

<sup>19</sup> Arnold *et al.* 1994 : 75, 79.

<sup>20</sup> Le nom vient d'un roman de science fiction humoristique, *Le guide galactique (The Hitch-Hiker's Guide to the Galaxy)* par Douglas Adams. Selon Adams (1982 : 69), « le poisson Babel -- est petit et jaune -- Il vit en effet de l'énergie des ondes cérébrales émises non pas par son hôte mais par tous ceux qui l'environnent. -- il vous suffit de glisser un poisson Babel dans votre oreille pour instantanément comprendre tout ce que l'on vous dit et ce, dans n'importe quelle langue. Les structures linguistiques effectivement entendues sont le décodage de la matrice d'ondes cérébrales injectées dans votre esprit par le poisson Babel. » Le choix du nom indiquerait que le service ne doit pas être pris trop au sérieux.

<sup>21</sup> Yang – Lange 2000 : 1.

<sup>22</sup> Les sens de traduction proposés en juillet 2001.

<sup>23</sup> [http://live.altavista.com/s?spage=help/babel\\_faq.htm](http://live.altavista.com/s?spage=help/babel_faq.htm) .

Babel Fish a été développée par SYSTRAN, le pionnier de la traduction automatique (=TA). BF utilise un système de traduction par transfert<sup>24</sup>, où le texte source est transformé en représentations de transfert de la langue source, qui sont transformées en représentations de transfert de la langue cible, qui, elles, sont synthétisées en texte en langue cible. L'avantage de ce système est que la synthèse vers la langue cible est indépendante de la langue source<sup>25</sup>, ce qui devrait éviter une bonne partie des interférences entre la langue source et la langue cible.

BF propose deux manières d'utilisation : l'utilisateur peut faire traduire des textes normaux en les tapant (ou en les copiant à partir d'un document déjà tapé) dans un champ réservé à cet effet ; ou bien il peut faire traduire une page web en tapant son adresse dans un autre champ (voir appendice 1). Dans ce dernier cas, la page conserve sa forme originale (avec les images et la mise en page), seul le texte ayant changé.

### 1.2.3.3 Reverso

Le logiciel de traduction Reverso a été développé par Softissimo et PROject MT<sup>26</sup>. Sa version *online* fonctionne en gros comme celle de Babel Fish : il permet de traduire soit des textes tapés dans une fenêtre<sup>27</sup>, soit des pages web entières (voir appendice 2). En plus de cela, il propose la traduction de courrier électronique, que nous n'avons pas essayé. La première version du logiciel (traduction du français vers l'allemand et l'anglais et l'inverse) fut publiée en 1998<sup>28</sup> et la version *online* quelque temps après.

Reverso *online* propose la traduction entre l'anglais et trois langues (le français, l'allemand et l'espagnol) dans les deux sens, entre le français l'allemand) dans les deux sens et encore de l'anglais et du français vers le russe<sup>29</sup>. Les sélections de langues des deux services ne sont pas identiques : BF propose quelques paires de

---

<sup>24</sup> Anon 2000b.

<sup>25</sup> Anon. 2000b.

<sup>26</sup> Anon. 2001a : 3.

<sup>27</sup> Selon Softissimo, la longueur des textes est limitée à 5000 caractères, mais en pratique Reverso ne traduit qu'une centaine de mots à la fois.

<sup>28</sup> Anon 2000c : 2.

<sup>29</sup> Voir par ex. <http://www.reverso.net/url/default.asp>.

langues en plus (notamment des langues asiatiques), mais Reverso propose la traduction vers le russe, ce qui vient du fait qu'un de ses deux fabricants, PROject MT, est russe.

Selon le *Livre blanc traduction*<sup>30</sup>, les étapes dans la traduction de Reverso sont : « 1. l'analyse du texte source ; 2. application de règles linguistique de transformation et de désambiguïsation ;3. génération de la traduction ». Cet outil pourra donc être classé parmi les systèmes de traduction par transfert.

### 1.3. Ambiguïtés

#### 1.3.1 Remarque préliminaire

Selon Greenfield-Cardey, une bonne partie des erreurs en traduction automatique résultent de l'échec de l'analyse des ambiguïtés du texte source. Il existe plusieurs types d'ambiguïtés, notamment celles lexiques (les homographes), celles syntaxiques et encore les anaphoriques.<sup>31</sup>

#### 1.3.2 Homographes

On appelle homographie le phénomène où une même graphie peut représenter deux ou plusieurs signifiés<sup>32</sup>. Par ex. *or* peut signifier un certain métal jaune mais aussi une conjonction introduisant un argument ou une objection. Dans cet exemple, les deux mots sont d'origine étymologique différente et ont fini par avoir la même forme par coïncidence. On peut également parler d'homographie dans des cas de polysémie, c'est-à-dire quand une forme graphique donnée prend deux ou plusieurs sens nettement différents, souvent par un élargissement de sens durant l'évolution diachronique<sup>33</sup>. Une distinction claire et pertinente entre la polysémie et l'homographie étant difficile à faire<sup>34</sup>, et cette problématique ne nous intéressant pas

---

<sup>30</sup> Anon. 2001a : 3.

<sup>31</sup> Greenfield-Cardey 1996 : 316.

<sup>32</sup> Dubois *et al.* 1973 : 245.

<sup>33</sup> Dubois *et al.* 1973 : 245.

<sup>34</sup> Dubois *et al.* 1973 : 382.

particulièrement dans ce travail, nous ne distinguerons pas les cas de polysémie des cas d'homographie lors de l'analyse.

Les formes homographiques peuvent appartenir à différentes parties de discours. En anglais, en particulier, il est possible de former des verbes à partir de noms (*jail* n 'prison', *jail* vt 'mettre en prison') ou l'inverse (*cry* v 'pleurer, crier', *cry* n 'cri') sans rien changer à la forme graphique (et phonétique) du mot. La même opération est également possible avec les adjectifs (*free* a 'libre', *free* vt 'libérer'). La reconnaissance de la fonction de telle et telle forme se fait, bien sûr, grâce aux indices syntaxiques, mais si la syntaxe de la phrase est très compliquée, le résultat de l'analyse automatique de la phrase risque de rester insatisfaisant.

### 1.3.3 Ambiguïtés syntaxiques

L'ambiguïté syntaxique (aussi appelée homonymie de construction) est due au fait que la structure syntaxique d'une phrase donnée peut avoir deux ou plusieurs interprétations sémantiques possibles<sup>35</sup>. Par ex. dans la phrase *Georges aime Marie autant que Jean*<sup>36</sup> il est clair que Georges aime Marie, mais est-ce qu'il l'aime autant que Jean aime cette femme, ou est-ce qu'il l'aime avec la même intensité qu'il aime son ami Jean ? Le contexte aidera éventuellement à le décider.

Les homographes peuvent évidemment être à la source de ce genre de problèmes. *La municipalité a fait de nos concitoyens les plus démunis des assurés sociaux*<sup>37</sup>. Ici, selon la signification de *des*, la municipalité a été soit prête à aider les pauvres (*des* interprété comme un déterminant indéfini), soit sans pitié (*des* vu comme la contraction de la préposition *de* avec le déterminant défini *les*). Au fond, le problème est donc de savoir auquel groupe nominal le groupe adjectival *les plus démunis* est attaché, *nos concitoyens* ou *[l]es assurés sociaux*.

Les phrases de ce genre sont rares, surtout dans des textes formels et bien rédigés. Pourtant, une phrase intelligible pour un lecteur humain peut perturber une machine

---

<sup>35</sup> Dubois *et al.* 1973 : 29.

<sup>36</sup> L'exemple provient de Dubois *et al.* 1973 : 29.

<sup>37</sup> Cet exemple provient de Laporte 1996 : 291.

qui doit l'analyser. Par ex. la phrase *Nous avons convenu avec la dame de l'agence de voyage de l'heure de départ du car* contient un tel nombre de prépositions *de* qu'il n'est pas clair lequel des *de* va avec *convenir*<sup>38</sup>.

### 1.3.4 Anaphoriques

Les anaphoriques sont des pronoms personnels ou démonstratifs qui font référence à un terme antérieur (=l'antécédent)<sup>39</sup>. L'antécédent est souvent un syntagme nominal (*Christine est intelligente. Je l'aime bien*) mais il peut aussi être un adjectif (*Il est gentil ? – Il l'est*), un verbe (*Partir, c'est mourir*), une phrase complète ou, moins souvent, un adverbe (*Vous êtes là ? – J'y suis*)<sup>40</sup>. L'emploi anaphorique des pronoms s'oppose à leur emploi déictique, dans lequel cas les pronoms ne font pas référence à un terme du texte<sup>41</sup>.

Les ambiguïtés causées par les pronoms anaphoriques sont difficiles, parfois même impossibles, à résoudre. Dans la phrase *She dropped the plate on the table and broke it* 'Elle a fait tomber l'assiette sur la table et elle l'a cassée' le texte seul ne nous permet pas de savoir si c'est l'assiette ou la table qui s'est cassée<sup>42</sup>. Dans le traitement automatique du texte, l'analyse des liens de référence se fait mécaniquement. Une multitude de pronoms anaphoriques dans une phrase rend l'analyse difficile et augmente le risque de fautes.

## 1.4. Différences entre le français et l'anglais

### 1.4.1 Remarques préliminaires

Certaines particularités des langues naturelles peuvent gêner la traduction automatique. Parmi celles-ci figurent la manière différente de « découper » le temps et les questions d'ordre de mot, notamment la modifications des noms.

---

<sup>38</sup> Un exemple de Greenfield-Cardey 1996 : 316.

<sup>39</sup> Dubois *et al.* 1973 : 34.

<sup>40</sup> Grevisse 1993 : 956-957.

<sup>41</sup> Déictique = qui donne les coordonnées spatio-temporelles de la situation de l'énonciation, Dubois *et al.* 1973 : 34, 137.

<sup>42</sup> Un exemple de Greenfield-Cardey 1996 : 316.

### 1.4.2 Les temps du passé en français et en anglais

Quand on compare les temps grammaticaux de deux langues différentes, on se rend compte qu'ils ne coïncident jamais parfaitement. L'expression des temps du passé est l'un des facteurs qui font que deux langues ne soient pas identiques entre elles. Le français exprime les différents aspects en variant l'imparfait et le « parfait » (passé composé et passé simple), tandis qu'en anglais le prétérit et le *present perfect* occupent des sphères de temps différents, comme on peut le voir au tableau 1. Le prétérit (*He ate*) peut donc correspondre à trois temps différents du français.

Français	Anglais
Il a mangé une pomme. (passé composé)	He has eaten an apple. (present perfect)
Il mangea une pomme. (passé simple)	He ate an apple. (prétérit)
Il mangeait une pomme. (imparfait)	He was eating an apple. (prétérit + aspect progressif)

Tableau 1. Différences de temps entre le français et l'anglais

En français, on utilise le passé simple ou le passé composé pour exprimer un fait considéré comme achevé et qui s'est déroulé à un moment dont on connaît le début et la fin précis. On utilise le passé simple si l'on ne veut pas considérer le contact qu'il peut avoir avec le présent (*Napoléon **naquit** en 1769*), et le passé composé si on veut considérer le rapport avec le moment de la parole. Ce rapport peut être de deux natures : soit le fait a eu lieu dans une période non encore entièrement écoulée (*je **me suis réveillé** tard ce **matin***), soit ce fait a eu des conséquences dans le moment présent (*il **a oublié** son manuel à la maison (et il ne peut pas travailler maintenant)*).<sup>43</sup> Dans la langue parlée, le passé composé peut également exprimer des faits sans rapport avec le moment de la parole, fonction traditionnellement conservée pour le passé simple.<sup>44</sup>

<sup>43</sup> Grevisse 1993 : 1252, 1254.

<sup>44</sup> Grevisse 1993 : 1255.

La valeur générale de l'imparfait est celui de montrer un fait en train de se dérouler dans le passé, sans en montrer les limites dans le temps (*Je lisais*). Il convient également pour indiquer des faits qui se répètent (*Elle allait à l'école tous les jours*), ainsi que pour la description des circonstances et du milieu des événements (*Il pleuvait*).<sup>45</sup>

En anglais, *the past tense* 'le passé' s'utilise pour se référer à des faits situés à un moment donné du passé. Il a trois utilisations différentes. Premièrement, il peut faire référence à un seul événement défini du passé (*The plane took off at noon.* 'L'avion a décollé à midi.'). Deuxièmement, il peut se référer à des faits qui se répètent (*We went to my aunt's every Sunday.* 'Nous allions chez ma tante chaque dimanche'). Finalement, utilisé avec des verbes statifs, il se réfère à un état des affaires ininterrompu dans le passé (*I once liked watching romantic films.* 'A une époque, j'aimais regarder des films romantiques').<sup>46</sup>

Il est possible de varier le sens du passé en y ajoutant un ou deux aspects. L'aspect est une catégorie grammaticale qui reflète le rapport que le sens du verbe a avec le temps. L'anglais a deux aspects, *perfect* et *progressif*. Le premier, combiné au présent, se rapproche du « parfait » du français. Il est possible de distinguer trois utilisations différentes. Premièrement, *the present perfect* s'utilise pour faire référence à un état des choses qui a commencé dans le passé et qui continue dans le présent (*They have lived in Paris for five years.* 'Ils habitent à Paris depuis cinq ans'). Deuxièmement, il peut se référer à des événements ayant eu lieu à une période qui s'étend jusqu'au présent, soit des événements très récents (*There has been an accident.* 'Il y a eu un accident'), soit des événements plus lointains qui ont néanmoins une influence sur le présent (*My daughter has had measles.* 'Ma fille a eu la rougeole' (sous-entendu : elle ne l'aura plus) ). Dernièrement, *the present perfect* s'utilise pour parler des événements qui se répètent encore aujourd'hui (*Plato has influenced many philosophers (till now).* 'Platon a influencé de nombreux philosophes (et aujourd'hui encore)').<sup>47</sup>

---

<sup>45</sup> Grevisse 1993 : 1250.

<sup>46</sup> Greenbaum – Quirk 1990 : 50.

<sup>47</sup> Greenbaum – Quirk 1990 : 51-52.



Ces différences entre le français et l'anglais provoquent des difficultés de traduction. La phrase : *As-tu entendu les nouvelles ce matin ?* se traduit en anglais selon le contexte extra-langagier soit par *Have you heard the news this morning?* (si c'est encore la matinée), soit par *Did you hear the news this morning?* (si c'est déjà l'après-midi).<sup>48</sup> D'un autre côté, la phrase *we went to the theatre* se traduit soit par le passé composé *nous sommes allés au théâtre* s'il s'agit d'une seule fois (*that night we went to the theatre* 'ce soir-là, nous sommes allés au théâtre'), soit par l'imparfait *nous allions au théâtre*, s'il s'agit d'une action répétée (*we often went to the theatre* 'nous allions souvent au théâtre'). Ici, ce sont les éléments de la phrase qui aident à décider de quoi il s'agit.<sup>49</sup>

En plus de cela, le français exprime une action qui a commencé dans le passé et qui se déroule depuis un certain temps avec la construction PRESENT + *depuis*, par ex. *Il habite à Paris depuis cinq ans*, ce qui se dirait en anglais avec la construction HAVE BEEN + -ING + *for*, *He has been living in Paris for five years*.

### 1.4.3 La modification des noms en anglais

Un nom peut être modifié par un nombre d'éléments qui se placent soit devant soit après le nom. La post-modification (modification par les éléments suivant le nom) est la manière la plus courante, mais nous nous concentrerons sur la pré-modification, qui pose plus de problèmes au passage de l'anglais vers le français.

En anglais, un nom peut être modifié par un autre nom (*machine translation* 'traduction automatique'), un adjectif (*intelligent person* 'personne intelligente') ou une nominalisation (*reception committee* 'comité de réception (qui reçoit)') qui se place devant le nom. La phrase nominale qui résulte de cette opération peut devenir un mot composé ou rester un élément libre dans la langue. Il y a des tests purement structurels pour décider si une phrase de ce genre est un mot composé. Si le premier élément ne peut pas être modifié indépendamment, il s'agit d'un mot composé.<sup>50</sup>

---

<sup>48</sup> Pilard *et al.* 2000 : A48.

<sup>49</sup> Pilard *et al.* 2000 : B28.

<sup>50</sup> Adams 1973 : 57.

Voyons l'exemple de *wet day* 'une journée pluvieuse' et *small talk* 'conversation mondaine'. Il est possible de dire *wetter day*, *very wet day* and even *the day is wet*. Par contre, il semble étrange de dire *\*smaller talk*, *\*very small talk* et *\*the talk is small*. Nous pouvons donc en conclure que *wet day* est une formation libre et *small talk* est un mot composé.<sup>51</sup>

Le même test s'applique aux phrases composées de deux noms. Des phrases comme *\*outer door knob* 'poignée de la porte extérieure' et *\*rare book case* 'étagère de livres rares' semblent étranges parce qu'on ne peut pas modifier le premier élément d'un mot composé indépendamment de la tête. Par contre, *brass door knob* 'poignée (de la porte) en laiton' et *mahogany book case* 'étagère en acajou' sont acceptables car le premier élément modifie la tête de la phrase nominale et donc toute la phrase.<sup>52</sup>

Quand ces phrases composées de deux ou plusieurs noms sont traduites en français, elles sont souvent rendues à l'aide d'une préposition, notamment *de* (*milk tooth*, 'dent de lait') ou avec un adjectif, qui remplace le premier de ces noms, celui qui modifie le deuxième (*tear gas* 'gaz lacrymogène'). L'ordre des mots est inversé, car en français, un complément qui définit un mot le suit plus souvent qu'il ne le précède. Certaines de ces phrases nominales doivent pourtant être traduites avec plus de liberté, parfois en explicitant les liens de modification, qui sont clairs en anglais mais qui ne le seraient pas en français, par ex. *Galapagos oil spill captain* 'le capitaine (du bateau qui a provoqué/subi) le déversement (accidentel) de pétrole (à) Galapagos' se traduirait mécaniquement par *\*le capitaine de déversement de pétrole de Galapagos*.

---

<sup>51</sup> Ce test est souvent fiable, mais il y a des exceptions, comme par exemple *bad loser* 'mauvais perdant' qui a un statut de composé malgré ses formes modifiées comme *a very bad loser*, Adams 1973 : 57.

<sup>52</sup> Adams 1973 : 57-58.

## 2. Analyse

### 2.1. Remarques préliminaires

Nous allons présenter certaines difficultés auxquelles les systèmes de traduction automatique doivent faire face. Parmi celles-ci figurent les questions d'orthographe, d'ambiguïtés et de locutions figées, et d'une autre part les problèmes liés aux temps du passé, à la modification des noms et à l'ordre des mots en général.

Dans les références au corpus, les mots et les passages qui nous intéressent le plus seront écrits en caractères gras. Les exemples tirés du corpus seront suivis d'un code qui permet de les identifier, p.ex. *GEN\_a, 60* veut dire qu'il s'agit de la ligne 60 dans l'article sur les gènes, en anglais. *GEN\_BF* indique la traduction en français de Babel Fish du même document et *GEN\_REV* la traduction de Reverso. De même, *SHA\_f* indique le document en français d'où est tiré l'exemple en question, et *SHA-BF* et *SHA\_REV* indiquent les traductions en anglais.

### 2.2. Orthographe

#### 2.2.1 Remarques préliminaires

L'ordinateur ne peut faire autre que ce pour quoi il a été programmé. Il ne peut donc pas reconnaître des mots s'ils ont une orthographe incorrecte. Nous avons repéré quelques cas où la traduction a échoué à cause des problèmes liés à l'orthographe.

#### 2.2.2 La ponctuation

Pour l'ordinateur, même les espaces (les blancs entre les mots) sont importants. Si les espaces sont omis, l'ordinateur ne sait pas définir les limites entre les mots. Nous avons identifié un cas de ce type de faute d'orthographe dans le corpus.

1. Click in the frame that contains the text you want to save.
2. At the File menu, choose Save Frame As.

3. **Name** the file, choose a folder on your hard disk, and click the Save button. (FAQ\_a, 203-205)

1. **Click** dans la trame qui contient le texte vous veulent sauvegarder.  
 2. **At** le menu de fichier, choisissent économiser la vue As.  
 3. **Name** le fichier, choisissent une chemise sur votre disque dur, et cliquent le bouton.  
 (FAQ\_BF, 213-215)

1. **Le clic** dans l'encadrement qui contient le texte vous veut sauver(économiser).  
 2. **Au** menu de Fichier, choisissez Sauvent(économisent) l'Encadrement Comme.  
 3. **Nommez** le fichier, choisissez un dossier sur votre disque dur et cliquez sur le bouton de Save. (1FAQ\_REV, 80-182)

Les premiers mots de chaque ligne de cette liste numérotée n'ont pas été traduits parce que le numéro et le point sont collés au mot. Ces mots (*1.Click*) ne se trouvent pas dans le dictionnaire. Dans ce cas précis, l'échec de la traduction des premiers mots n'a pas eu beaucoup d'effet sur la traduction du reste des phrases.

### 2.2.3 L'orthographe du français

Nous avons repéré quelques problèmes concernant la ligature œ, qui est utilisée dans un bon nombre de mots français, dont certains sont très courants comme *cœur*, *sœur*, *œil*, *vœux*.

au **cœur** de la Sibérie orientale (SIB\_f, 9-10)  
 with the **c?ur** of Eastern Siberia (SIB\_BF, 8-9)  
 in the **cœur** of the Eastern Siberia (SIB\_REV, 9)

... quand le candidat que nous appelons de nos **vœux** aura été élu... (PRE\_f, 28-29)  
 ... when the candidate that we call of our **v.ux** will have been elected... (PRE\_BF, 28-29)  
 ... when the candidate whom we call of our **vœux** will have been an elected member ...  
 (PRE\_REV, 30-31)

**l'œil** rivé sur les études d'opinion, le nez reniflant la moindre bonne nouvelle,... (PRE\_f, 62-63)  
 the **il** rivetted on the studies of opinion, the nose reniflant the least good news, ... (PRE\_BF, 62-63)  
 the **il** riveted on the studies of opinion, the nose sniffing at the slightest good short story(piece of news), ... (PRE\_REV, 68-70)

Aucune des deux machines n'a reconnu la ligature. Le fait que cette lettre soit remplacée (dans la traduction de BF) tantôt par un point d'interrogation, tantôt par un point, tantôt par rien du tout, indique que le problème est éventuellement lié à un mauvais codage en HTML<sup>53</sup> de cette lettre.

---

<sup>53</sup> HTML = Hyper Text Markup Language, langage utilisé pour produire des pages web.

un sujet d'actualité dans un pays étranger : traduisez la page web des meilleurs sites de ce pays pour être au **coeur** de l'information (AID\_f, 61-62)  
 a subject of topicality in a foreign country: translate the Web page of the best sites of this country to be in the **heart** of information (AID\_BF, 59-60)  
 A subject of current events in a foreign country: please , translate the web page the best sites of this country to be in the **heart** of the information (AID\_REV, 61-62)

besoin de garder un **oeil** sur la concurrence étrangère ? (AID\_f, 65)  
 need to keep an **eye** on foreign competition? (AID\_BF, 63)  
 Need to keep(guard) an **eye** on the foreign competition? (AID\_REV, 65)

Dans ces exemples les mots *coeur* et *oeil* ont été traduits correctement par les deux systèmes, malgré leur graphie erronée (il faudrait bien sûr écrire *cœur*, *œil*). Ces formes sont donc attestées dans le dictionnaire du système de traduction. Nous pouvons donc supposer que d'autres mots où la lettre *æ* est remplacée par *oe* se traduisent correctement.

## 2.3. Ambiguïtés

### 2.3.1 Ambiguïtés en anglais

Les homographes sont nombreux dans les langues telles que l'anglais et le français. Les homographes peuvent même appartenir à deux catégories syntaxiques différentes, si bien qu'un verbe peut, par erreur, être traduit par un nom, ce qui peut résulter en une traduction erronée du reste de la phrase.

#### 2.3.1.1 Homographes

**Copy and paste** in small sections and **translate** each section separately. (FAQ\_a, 105-106)  
**La copie et la pâte** dans de petites sections et **traduisent** chaque section séparément.  
 (FAQ\_BF, 111-112)  
**La copie et la pâte** dans de petites sections et **traduisent** chaque section séparément.  
 (FAQ\_REV, 95-96)

Comme en anglais la plupart des verbes au présent ont la même forme pour toutes les personnes (sauf la III<sup>e</sup> personne du singulier, qui prend *-s*), et que cette même forme s'utilise également à l'impératif, et qu'elle peut en plus être identique à une forme nominale, il y a de fortes chances qu'une machine se trompe dans l'analyse des formes ambiguës. Dans cet exemple, les trois impératifs ont été mal analysés : *Copy* et *paste* comme noms et comme sujet du verbe *translate*, et *translate* comme indicatif présent de la III<sup>e</sup> personne au pluriel.

Les impératifs, ainsi que les autres homographes, semblent poser des problèmes partout dans le FAQ, le texte qui contient le plus d'impératifs.

Just **type a period** in front of any proper name or other word that should not be translated. For example, **notice** what happens when we **type** two **sentences** in English and **translate** them into French, Italian, and German (FAQ\_a, 157-159)

**Type juste par période** devant tout nom propre ou tout autre mot qui ne devraient pas être traduits. Par exemple, **notification** ce qui se produit quand nous **type** deux **condamne** en anglais et les **traduit** en français, italien, et Allemand (FAQ\_BF, 165-167)

**Tapez juste une période** devant n'importe quel nom propre ou d'autre mot qui ne doit pas être traduit. Par exemple, **l'avis** ce qui arrive quand nous **tapons** deux **phrases(sentences)** en anglais et les **traduit** dans le français, l'italien et l'allemand (FAQ\_REV, 136-138)

Un groupement d'homographes mal interprétés a mené à une traduction incompréhensible. Dans la traduction de BF deux impératifs (le premier *type* 'tapez', *notice* 'remarquez, faites attention à'), ainsi qu'une autre forme verbale (*we type* 'nous tapons') ont été traduits par des noms. D'autre part, BF a analysé un nom en tant que verbe (*sentences* 'phrases'). De plus, un des noms de langue (*Allemand*) a gardé la majuscule. Les deux machines ont mal analysé le sujet de *translate* 'traduisons', qui est donc *nous* et non la troisième personne du singulier, comme le propose la forme verbale choisie, *traduit*.

Il est difficile de découvrir les raisons pour lesquelles BF réussit parfois la traduction des impératifs et pourquoi il se trompe dans d'autres cas, bien plus nombreux.

**View** the AltaVista France (or Germany or Italy) Search Results translated into English. (FAQ\_a, 43)

**Visualisez** les résultats de recherche d'AltaVista France (ou l'Allemagne ou l'Italie) traduits en anglais. (FAQ\_BF, 46-47)

**Considérez** AltaVista la France (ou l'Allemagne ou l'Italie) des Résultats de Recherche traduits dans l'anglais. (FAQ\_REV, 40-41)

**View** AltaVista.com Search Results in French, German, Spanish, Portuguese, or Italian. (FAQ\_a, 45)

Résultats de recherche **de la vue** AltaVista.com en français, allemand, espagnol, portugais, ou l'Italien. (FAQ\_BF, 48-49)

**Considérez** des Résultats de Recherche d'AltaVista.com en français, l'allemand, l'espagnol, le portugais, ou l'italien. (FAQ\_REV, 41-42)

Ces deux derniers exemples ne diffèrent que dans quelques mots, le verbe à l'impératif est le même dans les deux phrases, et pourtant le résultat final est complètement différent. Il est possible que le point dans *AltaVista.com* gêne

l'analyse de la phrase. Reverso a produit une traduction grammaticalement juste des deux phrases bien que son choix lexical soit discutable.

Selon le FAQ, BF traduit mieux si les phrases sont complètes. Il ne fait pas d'exception pour les titres. Dans l'exemple suivant, la forme *asked* est interprétée comme un verbe conjugué, et non comme un adjectif épithète qui spécifie le nom *questions*. De même, le nom *Babel Fish* est considéré comme le sujet de la phrase, et non comme un complément de *questions*.

BABEL FISH FREQUENTLY ASKED QUESTIONS (FAQ\_a, 1)  
 LES POISSONS DE BABEL ONT FRÉQUEMMENT POSÉ DES QUESTIONS (FAQ\_BF, 1)  
 QUESTIONS FRÉQUEMMENT POSÉES DE POISSON DE BROUHAHA(BABYLONE)  
 (FAQ\_REV, 1)

Ici, le contexte est le seul facteur désambiguïsant, car l'interprétation de BF serait tout à fait légitime pour cette phrase hors contexte. La prise en compte du contexte (la phrase est un titre, et les titres sont souvent des phrases incomplètes) aurait pu contribuer à l'amélioration de l'analyse syntaxique, ainsi que la conscience du fait que *Frequently Asked Questions* est un concept (dont une traduction possible est *Foire Aux Questions*) et qu'il doit être analysé comme un tout.

L'organisation des phrases d'une manière inattendue est une source de ce genre de fautes, car il empêche le bon déroulement de l'analyse de la phrase.

DNA of people of different races is unexpectedly alike, new genome findings show.  
 (GEN\_a, 2-3)  
 L'cAdn des personnes de différentes unités de feuillets magnétiques est inopinément des résultats semblables et nouveaux de génome montrent. (GEN\_BF, 6-7)  
 l'ADN des gens de courses(races) différentes est l'exposition de découvertes de génome inopinément semblable, nouvelle. (GEN\_REV, 2-3)

Cette phrase peut se traduire ainsi : *L'ADN des gens appartenant à des races différentes est inopinément semblable, montrent les nouvelles découvertes sur le génome*. La phrase doit être coupée en deux au niveau de la virgule : le début est une phrase subordonnée, la fin est la principale. Si elles étaient dans l'ordre inverse (*New genome findings show that...* 'Nouvelle découvertes sur le génome montrent **que...**'), l'identification de la subordonnée serait plus facile grâce à *that*, qui marquerait la subordonnée. Ne pouvant pas profiter de cet indice, les deux machines ont analysé la phrase d'une manière erronée. Reverso l'a prise pour une phrase

simple ayant la structure sujet [*DNA of people of different races*] + verbe [*is*] + attribut du sujet [*unexpectedly alike, new genome findings show*], comme si la phrase disait que l'ADN était [*an*] *unexpectedly alike, new genome findings show* 'une nouvelle exposition inopinément semblable de découvertes sur le génome'. La mauvaise analyse de *show* (dans cette phrase 'montrent' et non 'exposition') contribue à l'échec de la traduction du reste de la phrase. BF a bien traduit le mot *show*, mais il a coupé la phrase ainsi : sujet [*DNA of people of different races*]+ verbe [*is*] + attribut du sujet [*unexpectedly alike, new genome findings*] + verbe [*show*]. De plus, il a traduit le mot *races* 'races' par le terme technique *unités de feuillets magnétiques* dont l'origine nous reste mystérieuse<sup>54</sup>.

**Stuart Wheeler, chairman and chief executive of the IG Index**, which specialises in sport and financial betting, **gave** the Conservative Party £1 million yesterday and **promised** a further £4 million before the election. In so doing **he became** probably the biggest one-off donor in political history. (TOR\_a, 11-15)

**Le cycliste Stuart, le Président et le cadre supérieur de l'incrément d'Ig**, qui se spécialise dans le sport et le pari financier, **ont donné** à la partie conservatrice £1 million hier et **ont promis** un autre £4 million avant l'élection. Ce faisant **il est devenu** probablement le plus grand donateur unique dans l'histoire politique. (TOR\_BF, 11-15)

**Stuart Wheeler, le président et le directeur de l'Index IG**, qui se spécialise dans le pari sportif et financier, **ont donné** la Partie(parti) Conservatrice £ 1 million hier et **ont promis** un nouveau £ 4 millions avant l'élection. Ainsi **il est devenu** probablement le donateur(donneur) unique le plus grand dans l'histoire politique. (TOR\_REV, 11-15)

Cet exemple démontre encore une fois que les machines de traduction ne prennent pas le contexte en considération. Pour un humain il suffit de lire les deux phrases pour comprendre que Stuart Wheeler et *chairman and chief executive of the IG Index* sont la même personne. Pour une machine, ce fait est moins évident. Les mots *gave* 'a donné' et *promised* 'a promis' peuvent de par leur forme être liés à un sujet au pluriel et ils peuvent donc se traduire par *ont donné* et *ont promis*. Or, *Stuart Wheeler, chairman and chief executive of the IG Index* 'Stuart Wheeler, le président et le cadre supérieur de IG Index' n'est pas une énumération de trois personnes, mais une apposition du nom propre d'une personne et deux noms désignant la position de cette personne. Savoir cela permet de résoudre que *gave* et *promised* ont un sujet au singulier.

---

<sup>54</sup> Dans le langage de l'informatique, *race* désigne un état non-désiré provoqué par des circuits digitaux mal conçus, Green 1987 : 440. D'autre part, le *feuille* signifie une mince feuille plastique recouverte de matière aimantée servant de mémoire auxiliaire d'ordinateur, Robert 1985 : 488. Ces deux termes ont donc peu à voir l'un avec l'autre.



## 2.3.2 Ambiguïtés en français

### 2.3.2.1 Homographes

De petites erreurs se produisent constamment, pourtant **la vie continue** on ne sait trop comment. (BIO\_f, 91-93)  
 Small errors occur constantly, however **the life continues** one does not know too much how. (BIO\_BF, 87-89)  
 Small errors occur constantly, nevertheless **the continuous life** one does not too much know how. (BIO\_REV, 81-83)

Reverso a analysé le verbe *continue* en tant qu'adjectif, ce qui a rendu la phrase pratiquement incompréhensible. La traduction de BF est formellement correcte, mais il serait plus idiomatique d'écrire *life continues* (donc sans l'article défini) ou même *life goes on* (l'extrait parle des conditions de vie des cellules).

Je regarde **d'un air distrait** les objets que je suis censé acquérir la semaine prochaine. (BOY\_f, 55-57)  
 I **look of an air** distracts the objects which I am supposed to acquire the next week. (BOY\_BF, 54-56)  
 I **look of a distracted air(sight)** at the objects which I am supposed to acquire next week. (BOY\_REV, 55-57)

Dans cet exemple le phénomène inverse s'est produit. BF a analysé un adjectif, *distrain*, comme un verbe si bien qu'un groupe nominal est devenu une phrase contenant un sujet (*un air*) et un verbe (*distrain*). Reverso a bien analysé ce groupe nominal, mais une traduction fidèle, mot à mot, n'est pas la meilleure solution possible ici.

Elle regardait **son imbécile**. (TEL\_f, 95)  
 It looked at **its imbecile**. (TEL\_BF, 96-97)  
 She(it) looked **imbecile at sound**. (TEL\_REV, 99-100)

Hélas, mille fois hélas, **son fougueux amant** s'était distraitement abonné chez un opérateur concurrent. (TEL\_f, 118-120)  
 Alas, thousand times alas, **his/her impetuous lover** had abstractedly subscribed in a competitor operator. (TEL\_BF, 118-121)  
 Regrettably, one thousand time regrettably, **passionate sound lover** had absently subscribed at a rival operator's. (TEL\_REV, 125-127)

Le premier extrait parle d'un couple assis à un café, l'homme téléphone, la femme s'ennuie. *Son imbécile* fait donc référence à l'homme, *son* est un adjectif possessif, et *imbécile* est un nom. La traduction de BF serait correcte dans un contexte différent, mais comme *elle* fait référence à une femme et non à un objet, le sujet de cette phrase devrait se traduire par *she* et *son*, conséquemment, par *her*. Reverso a mal analysé le groupe nominal. Il est vrai que *son* peut être un nom, et *imbécile* est souvent utilisé en tant qu'adjectif, mais ici la manque d'article devant le nom conteste cette

possibilité. De plus, la phrase anglaise de Reverso permet une lecture qui veut dire qu'elle « regardait un son » et que, tout en le regardant, elle était imbécile. Une image très poétique mais fautive. Le même phénomène se reproduit dans la traduction du deuxième extrait. *Son fougueux amant*, un seul groupe nominal, est coupé en deux [*son fougueux*] et [*amant*], qui donnent respectivement *passionate sound* et *lover*. Ici, comme dans le premier exemple, le manque de l'article devant le nom n'a pas empêché Reverso de choisir cette solution.

Rien n'est plus déprimant pour le téléspectateur romantique qu'une publicité Itinérés.  
(TEL\_f, 103-105)

Nothing is depressing any more for the romantic televiewer than a Itinérés publicity.  
(TEL\_BF, 104-106)

Nothing is more depressing for the romantic televiewer than an advertisement Itinérés.  
(TEL\_REV, 108-110)

*Rien n'est plus déprimant ... qu'une publicité Itinérés.* Pour un lecteur humain, cette phrase n'a rien d'ambigu. Une machine, pourtant, hésite entre *davantage* et *désormais ne ... pas*, les deux sens de *plus*. Il est vrai qu'une personne guérie de dépression pourrait dire *nothing is depressing any more* 'rien n'est plus déprimant'. Par contre, dans la phrase *rien n'est plus déprimant que...* il est clair que *plus* est utilisé pour exprimer le comparatif. Reverso a su faire le rapprochement, BF non.

Je vous aurais bien raconté la suite, mais j'avais déjà zappé. (BOY\_f, 73-74)

I would have told you well the continuation, but I had already zappé. (BOY\_BF, 72-73)

I would have told well you the continuation(suite), but I had already channel-flicked.

(BOY\_REV, 73-75)

Ici *bien* n'est pas l'adverbe correspondant à l'adjectif *bon*, mais il est utilisé de manière emphatique, pour souligner la volonté du narrateur de partager son expérience avec les lecteurs<sup>55</sup>. Dans le parlé, il pourrait se traduire par une emphase sur le mot *would* (marque de conditionnel), *I would have told you the continuation...* A l'écrit le même effet pourrait être atteint par une légère modification de la phrase, *I would have liked to tell you... 'j'aurais aimé vous raconter...'*.

"Il faut nous préparer à organiser, à porter un mouvement qui sera le mouvement de ce président, qui sera le mouvement destiné à soutenir son action", a ajouté le maire de Bordeaux. "Qu'on ne se trompe pas d'adversaire". (PRE\_f, 32-35)

" We should be prepared to organize, to carry a movement which will be the movement of this president, who will be the movement intended to support his action ", added the mayor of Bordeaux " Which one is not mistaken in adversary ". (PRE\_BF, 32-35)

" It is necessary to prepare us to organize, to carry(wear) a movement which will be the movement of this president, which will be the movement intended to support its action "

<sup>55</sup> Robert 1985 : 971-973.

added the mayor of Bordeaux. " **That one does not make a mistake** about opponent " .  
(PRE\_REV, 34-37)

Il est particulièrement ennuyeux pour la traduction que certains mots grammaticaux doivent être des homographes d'autres mots ayant d'autres valeurs grammaticales. Tel est le cas de *que*, qui peut être un pronom relatif ou une conjonction. Dans la phrase originale, il est utilisé en tant que conjonction, pour marquer un impératif ('ne nous trompons pas...'). Les deux machines l'ont pris pour un pronom relatif, dont l'antécédent est peu clair. Le malheur avec cette solution est que la traduction n'est pas lue en tant que subordonnée relative mais qu'elle évoque une image totalement différente de celle transmise par le texte original. *Which one is not mistaken* 'lequel ne se trompe pas' est comme un début de question, car une phrase qui commence par *which* se lit plus facilement comme une interrogation que comme une subordonnée relative, qui, normalement, se place après la principale. *That one does not make a mistake* 'celui-là ne se trompe pas' a tout d'une constatation. La « mauvaise » lecture de la traduction de Reverso est due au fait que *that*, en plus des valeurs grammaticales correspondantes à celles de *que*, a la valeur de pronom démonstratif *celui-là*.

Le fait que ces outils ont été développés avant tout pour faciliter la traduction des documents spécialisés frappe un utilisateur dit « normal ». Les machines choisissent parfois parmi les significations d'un mot celui du vocabulaire spécialisé.

Cliquez sur l'**onglet** 'Texte' : (AID\_f, 3)  
Cliquez on **the mitre** ' Texte': (AID\_BF, 3)  
Please , click **the tab** 'Text': (AID\_REV, 1)

Le mot *onglet* a plusieurs significations, toutes plus ou moins spécialisées. Dans ce texte il signifie la partie qui dépasse une feuille de papier et qui permet de repérer la feuille bien qu'elle soit intercalée entre d'autres feuilles (dans ce cas une image numérique imitant une feuille, voir l'annexe 2). La traduction de Reverso, *tab*, a cette signification. Le terme choisi par BF est utilisé dans la menuiserie pour indiquer l'extrémité d'une planche ou d'une moulure qui forme un angle de 45 degrés, et d'une autre part, un assemblage de deux moulures (ou deux planches) de cette sorte se coupant à l'angle droit<sup>56</sup>. De plus, BF n'a pas traduit les mots *cliquez* et *texte*, le

---

<sup>56</sup> Murray *et al.* 1989 : 912.

premier parce qu'il n'est pas dans le dictionnaire du système, le deuxième probablement parce qu'il est entre parenthèses.

Quelques **conseils** pour améliorer les traductions fournies par Reverso Online (AID\_f, 69)  
 Some **consultings** to improve the translations provided by Reverso Online (AID\_BF, 67)  
 Some **advices** to improve the translations supplied by Reverso Online (AID\_REV, 70)

Le mot *conseil* signifie, dans le langage courant, un avis, proposition, recommandation ou suggestion donné pour aider un ami. Il signifie aussi une personne chez qui on cherche des conseils ou encore une réunion de personnes qui donnent leur avis sur des affaires publiques ou privées<sup>57</sup>. Aujourd'hui, *consulting* ainsi que *conseil* sont utilisés pour désigner un métier qui consiste à donner des conseils (financiers, par rapport au personnel etc.) par ex. à des entreprises. Le mot *advice* proposé par Reverso pourrait être la bonne solution ici, mais non au pluriel car ce mot n'en a pas.

### 2.3.2.2 Pronoms

Certains pronoms du français sont ambigus, soit dans la fonction (par ex. *te* peut être un COD, un COI ou un pronom réfléchi), soit dans la sémantique (les pronoms de la III<sup>e</sup> personne peuvent être utilisés pour référer non seulement à des humains mais aussi à des objets et des animaux)<sup>58</sup>.

La découverte du **faible nombre de gènes** qui gouvernent notre devenir biologique - un grain de riz **en** comporte 50 000 ! - implique que... (BIO\_f, 25-27)  
 The discovery of the **low number of genes** which control our to become biological - un grain of rices comprises **of them** 50 000 ! - implies that... (BIO\_BF, 25-27)  
 The discovery of the **weak number of genes** which govern our biologic future - a grain of rice contains 50 000 **of it!** - imply that... (BIO\_REV, 22-23)

Le pronom *en*, qui exprime le partitif ou un pluriel indéfini peut se traduire de plusieurs façons selon le groupe nominal qu'il reprend. Seul le contexte permet de choisir la bonne traduction. Ici BF a fait le rapprochement entre le pluriel *gènes* et le pronom *en*, en traduisant ce dernier par *of them* 'd'eux'. Reverso semble avoir considéré *en* comme la reprise du singulier *nombre*, ce qui explique pourquoi il a

---

<sup>57</sup> Robert 1985 : 836.

<sup>58</sup> Pour avoir une vue d'ensemble sur les pronoms personnels, voir Grevisse 1993 : 986.

traduit *en par of it* 'de lui, de cela'. Une traduction idiomatique pourrait être *a grain of rice contains 50 000 genes*.

Un jeune couple s'embrasse avec passion. Le train (ou le bus, je ne sais plus) va partir. Elle s'arrache à son étreinte, lui demande son numéro de portable. Il le griffonne, amoureuxment. (TEL\_f, 109-113)

A young couple is embraced with passion. The train (or the bus, I do not know any more) is leaving. It is torn off with its pressure, asks for its number of portable to him. It it griffonne, lovingly. (TEL\_BF, 109-113)

A young couple kisses(embraces) with passion. The train (or bus, I do not know any more) is going to leave. She(it) tears away(extracts) from her(its) embrace, asks him(her) her(its) number of mobile. It(he) scrawls him(it), lovingly. (TEL\_REV, 113-119)

Cet exemple nous présente un couple, un homme et une femme. Les anaphoriques de la III<sup>e</sup> personne du singulier sont donc nombreux. BF propose seulement deux traductions formellement correctes (*to him* 'lui' et *it* 'le, COD') mais bizarrement placées dans les phrases, ce qui fait que la traduction n'est pas de l'anglais idiomatique. Il préfère visiblement les pronoms démonstratifs aux pronoms personnels. Reverso propose, comme d'habitude, deux traductions pour chaque mot ambigu. Sur six pronoms de cet exemple, quatre (*she, him, he, it*) sont juste, bien que le bon choix ne soit pas toujours en première position. Pour un traducteur humain les liens de référence sont parfaitement clairs, car il peut imaginer la scène devant ses yeux. Faire comprendre la même situation à une machine, par contre, doit être très difficile, voire impossible.

Depuis, les couples ne sont plus ce qu'ils étaient. Pendant qu'il conduit, on la voit téléphoner à quelqu'un d'autre. (TEL\_f, 68-71)

Since, the couples are not any more what they were. While it leads, one sees it telephoning someone else. (TEL\_BF, 70-73)

Since, the couples are not any more what they were. While it(he) leads(drives), one sees her(it) phoning to someone else. (TEL\_REV, 72-74)

On rencontre les mêmes difficultés dans un deuxième exemple du même texte. Ici encore, BF a préféré le pronom démonstratif *it*, bien qu'il n'y ait dans la phrase aucun antécédent qui puisse y correspondre. Ici les pronoms personnels sont utilisés de manière déictique pour faire référence à deux personnes de sexes opposés auxquelles le texte ne fait pas référence explicitement (implicitement *un couple se forme d'une femme = elle et d'un homme = il*). La seule interprétation possible des mots *il* et *la* est donc celle de pronoms personnelles, mais il faut, encore une fois, être un humain pour voir la relation entre le mot *couples* et les pronoms *il* et *la*, ses composants. Nous ne connaissons pas les raisons pour lesquelles Reverso propose *it* (pronom impersonnel) plutôt que *he* (pronom personnel) pour traduire *il*, et, d'un

autre côté, il propose *her* (pronom personnel) plutôt que *it* (pronom impersonnel) pour traduire *elle*.

## 2.4. Lexies figées

### 2.4.1 Remarques préliminaires

Les formes idiomatiques, ne pouvant se traduire mot à mot, sont difficiles même pour un traducteur humain. Un système de TA ne peut les identifier et les traduire que si elles sont enregistrées dans son dictionnaire, ce qui demande un important travail préparatoire.

BF utilise le dictionnaire d'expressions de SYSTRAN, qui contient des expressions formées de plusieurs mots. Les expressions peuvent être simplement des mots qui sont fréquemment attestés ensemble, des mots composés ou des expressions qui contiennent des règles de traduction syntaxiques ou sémantiques.<sup>59</sup> De même, Reverso vante, sans pourtant en présenter le moindre détail technique, un système d'analyse qui permet la traduction des expressions de manière plus intelligente que le simple mot à mot<sup>60</sup>.

### 2.4.2 Lexies figées de l'anglais

#### 2.4.2.1 Verbes à particule (*phrasal verbs*)

Le vocabulaire anglais contient beaucoup de verbes dont le sens varie, parfois considérablement, selon la préposition, la postposition ou l'adverbe qui lui est associé. P. ex. à partir du verbe *put* 'mettre' ont été formés de nouveaux verbes ayant la signification de 'proposer, avancer' (*put forward*), 'reporter' (*put off*), 'loger' (*put up*) et même 'supporter' (*put up with*). Le sens d'un verbe à particule ne peut pas toujours être déduit à partir du sens du verbe et de celui de la particule qui le forment, et de plus, ces verbes à particule ont souvent plusieurs significations. Une combinaison de verbe et de préposition devrait donc toujours être analysée comme

---

<sup>59</sup> Anon. 2000b : 1.

<sup>60</sup> Anon 2001a : 5-6.

un seul mot, ce qui n'est malheureusement pas toujours le cas. Nous verrons d'abord un exemple de traduction réussie pour démontrer que la traduction de ces formes est à la portée des systèmes automatisés et ensuite des exemples de problèmes d'analyse que ces formes peuvent entraîner.

... and look for certain patterns of SNPs that **crop up** in the group with the disease.

(GEN\_a, 64-65)

... et rechercher certaines configurations de SNPs qui **surviennent** dans le groupe avec la maladie. (GEN\_BF, 85-86)

... et cherchent les certains modèles de SNPS qui **surgissent** dans le groupe avec la maladie. (GEN\_REV, 76-77)

Le verbe *crop* signifie 'couper ; cultiver'. *Crop up* de sa part veut dire 'se présenter, surgir'. Les deux machines l'ont bien analysé et traduit avec un seul mot.

He said that he had given the money in the hope of helping Mr Hague to have a chance of winning the election, rather than simply **putting up** a good show. (TOR\_a, 34-36)

Il a dit qu'il avait donné l'argent dans l'espoir d'aider M. Hague à avoir une chance de gagner l'élection, plutôt que simplement **mettant vers le haut** d'une bonne exposition. (TOR\_BF, 36-38)

Il a dit qu'il avait donné l'argent dans l'espoir d'aider M. Hague à avoir une chance de gagner l'élection, plutôt que le simplement **levage** d'une bonne exposition.

(TOR\_REV, 36-38)

*Put up* (verbe transitif) signifie, entre autres, 'relever, ouvrir (un parapluie) ; dresser, monter ; augmenter (prix) ; loger ; proposer, présenter (des arguments) ; fournir (une somme) ; baisser (une épée)'. Pour traduire cette phrase, BF a analysé les deux mots séparément et Reverso a choisi le sens de 'lever', qui est de première vue le plus naturel. Le sens recherché par l'auteur serait pourtant celui de 'proposer, présenter'. Une traduction possible de la fin de la phrase est *plutôt que simplement présenter (monter) un bon spectacle*.

... Mr Cook is expected to **fly on** to Syria and Jordan (TIM\_a, 43-44)

... on s'attend à ce que M. Cook **vole en fonction** en Syrie et en Jordanie. (TIM\_BF, 49-50)

... on s'attend à ce que M. Cook **vole sur** en Syrie et la Jordanie. (TIM\_REV, 45-46)

BF interprète *on* dans le sens 'en fonction, allumé' quand il s'agit d'une partie intégrante du verbe, comme dans *fly on* ('continuer le voyage en avion'). Reverso interprète cette préposition comme si elle ne faisait pas partie du verbe.

... is transferred from small treat bags into larger containers **on** top of the refrigerator. (HAL\_a, 25-26)

... est transférée à partir de petits sacs de festin dans de plus grands récipients **sur** le réfrigérateur. (HAL\_BF, 27-28)

... est transféré de petits sacs de plaisir dans de plus grands conteneurs **en plus du** réfrigérateur. (HAL\_REV, 27-28)

... put the peace process back **on** track. (TIM\_a, 30)  
 ... remettre le processus de paix **sur** la piste. (TIM\_BF, 34-35)  
 ... remettre le processus de paix **sur** la trace. (TIME\_REV, 32)

Dans le premier exemple, BF a traduit la préposition indépendante *on* correctement. Reverso l'a interprété comme faisant partie de la phrase *on top of* 'en plus de'. Dans le deuxième exemple toutes les deux machines ont choisi la bonne traduction.

#### 2.4.2.2 Mots composés à tiret

BF a un comportement qui semble illogique vis-à-vis des mots composés fournis d'un tiret. Des fois il ne les traduit pas du tout, bien qu'au moins une des parties du mot composé soit un mot connu par le système (p.ex. *English* dans *Non-English*).

How Can I Use My Physical U.S. Keyboard to Type Non-English Characters? (FAQ\_a, 174)  
 Comment est-ce que je puis employer mon clavier physique des États-Unis pour taper des caractères de Non-English? (FAQ\_BF, 184)  
 Comment puis-je Employer Mon Clavier américain Physique pour Taper Caractères non-anglais ? (FAQ\_REV, 154)

... choose World, English, French, German, Italian, Spanish, Portuguese, or Russian from the **drop-down** list. (FAQ\_a, 180-181)  
 ... choisissez le monde, l'anglais, le français, l'allemand, l'italien, l'espagnol, le portugais, ou le russe de la liste **drop-down**. (FAQ\_BF, 192-193)  
 ... pour choisir le Monde, l'anglais, le français, l'allemand, l'italien, l'espagnol, le portugais, ou le russe de la **liste déroulante** (FAQ\_REV, 159-160)

Dans d'autres cas, il traduit les mots composés, mais en traitant les deux parties séparément, comme s'il s'agissait de deux mots isolés écrits ensemble par hasard, plutôt que d'un mot composé ayant une signification à part.

the **drive-through** bank teller handouts (HAL\_a, 11)  
**piloter-à travers** des aumônes de caissier guichetier (HAL\_BF, 12-13)  
 la **commande - par** des aumônes(prospectus) de caissier de banque (HAL\_REV, 11-12)

the kindergarten class **hoe-down** (HAL\_a, 12)  
 la **houe-vers le bas** de classe de jardin d'enfants (HAL\_BF, 13)  
 la **danse** de classe de jardin d'enfants (HAL\_REV, 13)

Reverso semble avoir un dictionnaire plus complet, car il connaît *drop-down* et *hoe-down* et il en donne une traduction satisfaisante. Il a cependant traduit *drive-*



*through*<sup>61</sup> par *commande - par*, probablement parce qu'il n'est pas attesté dans son dictionnaire. Les espaces autour du tiret dans *commande - par* ont été ajoutés par Reverso, le texte d'origine n'en contenait pas. Il est vrai que *drive* (en tant que verbe) peut signifier 'commander (un programme)' dans le langage de l'informatique, et que la préposition *through* a parmi ses significations celle de 'par (le moyen de)'<sup>62</sup>.

Il n'y a pas de raison évidente qui explique pourquoi BF ne sait pas traduire *drop-down* ; à priori il devrait connaître le mot *drop*, qui est un mot courant, et *hoe-down*<sup>63</sup> prouve qu'il connaît *down*. Vu que les exemples où les mots avec le tiret n'ont pas été traduits proviennent du même document (FAQ), nous ne pouvons que poser qu'il s'agit d'une erreur d'analyse du tiret dans ces cas.

### 2.4.2.3 Locutions

Une locution est un groupe de mots ayant un sens autre que la somme des sens de ses composants. Par conséquent, elles doivent être traitées comme unités de sens. Une manière pour permettre à un système de traduction de reconnaître les locutions est de coder chaque mot<sup>64</sup> de chaque locution connue comme susceptible de faire partie d'une locution, ce qui demanderait énormément de travail. Les locutions figées constituent les structures se prêtant le moins bien à la traduction automatique.

... the Foreign Secretary -- flew to Egypt today **on the first leg** of his **day-long** diplomatic mission to try to salvage the Middle East peace process. (TIM\_a, 20-22)  
 ... le ministre des affaires étrangères -- a volé en Egypte aujourd'hui **sur la première jambe** de sa mission **toute la journée** diplomatique pour essayer de récupérer le processus de paix de le Moyen-Orient. (TIM\_BF, 23-26)  
 ... le Ministre des Affaires Etrangères, a volé en Egypte aujourd'hui **sur le premier pied** de sa mission diplomatique **de toute la journée** d'essayer de sauver le processus de paix du Moyen-Orient. (TIM\_REV, 22-24)

Ici il s'agit autant de polysémie que d'une locution figée, étant donné que *leg* a comme signification 'une partie d'un long voyage, souvent entre deux points où l'on s'arrête' et qu'il s'utilise normalement avec un ordinal et la préposition *of* (*the first*

<sup>61</sup> Un concept américain qui désigne des services (notamment banque, cinéma, restaurant) dont le client peut profiter depuis le volant sa voiture, se dit aussi *drive-in*, Murray *et al.* 1989 : 1061-1062.

<sup>62</sup> Pilard *et al.* 2000 : 280, 992.

<sup>63</sup> Ce mot signifie à l'origine une danse bruyante et passionnée qui conclut un bal. Murray *et al.* 1989 : 284. Ici, il signifierait une fête pour les enfants de la garderie.

<sup>64</sup> Sauf les mots les plus fréquents et ceux qui ne portent pas beaucoup de valeur sémantique, comme *être, avoir, faire*, les prépositions, les articles etc.

*leg of*)<sup>65</sup>. La traduction *sur la première jambe*, incompréhensible, n'est en aucun cas acceptable, car *jambe* n'a pas cette signification et il en va de même pour *sur le premier pied*. De plus, la traduction de *day-long* 'd'une journée' par *toute la journée* ne facilite pas la compréhension de cette phrase.

The smell is like being trapped in an ammonia factory after your nostril hairs **have caught on fire**. (SPI\_a, 44-45)

L'odeur est comme être emprisonné dans une usine d'ammoniaque après que vos poils de narine **se soient accrochés au feu**. (SPI\_BF, 48-50)

l'odeur ressemble être pris au piège à une usine d'ammoniac après que vos cheveux de narine **ont pris feu**. (SPI\_REV, 49-50)

La traduction fidèle de ces expressions peut évoquer des images surprenantes. On imagine mal des poils de nez qui *s'accrochent au feu* (*catch on fire*).

I don't really **get it**. (SPI\_a, 7)

Je ne **l'obtiens** pas vraiment. (SPI\_BF, 7)

Je ne **l'obtiens** pas vraiment. (SPI\_REV, 7)

Cette expression familière *to get it* 'comprendre, piger quelque chose' a le désavantage d'être formée uniquement de mots passe-partout. L'ensemble de ces deux mots, *get* 'obtenir, recevoir, avoir' et *it* 'il, ce' peut facilement être utilisé dans un autre contexte sans qu'il ne s'agisse de la locution *get it*. Il peut se rencontrer par ex. dans la réponse *No, I didn't get it* 'non, je ne l'ai pas reçu' à la question *Did you get my letter?* 'As-tu reçu ma lettre?'.

Parmi les locutions figées, les proverbes occupent un rôle particulier. Ils sont chargés de significations liées à la culture du locuteur et peuvent rarement se traduire tel quel.

We can't cry over **spilled milk** (OIL\_a, 115)

nous ne pouvons pas **pleurer lait plus de renversé** (OIL\_BF, 130)

Nous ne pouvons pas **pleurer sur le lait renversé** (OIL\_REV, 98)

Ici nous avons une transformation du proverbe *there's no point in crying over spilt milk* 'ce qui est fait est fait'<sup>66</sup> (litt. 'cela ne sert à rien de pleurer sur le lait renversé'). Le français exprime la même pensée d'une manière moins poétique, si bien que même une traduction fidèle aux sens des mots (comme celle de Reverso) serait fautive. La traduction de BF est également incompréhensible.

<sup>65</sup> Sinclair 1987 : 826.

<sup>66</sup> Pilard *et al.* 2000 : 590.

Le choix des locutions à inclure dans le dictionnaire d'un outil de traduction peut parfois sembler illogique.

In fact, baby spit-up is so **common place**, I'm surprised the nation's toy makers haven't **jumped on the bandwagon** and invented a Puking Patty doll we could give to our daughters for their birthdays. (SPI\_a, 33-35)

En fait, le bébé crachent-vers le haut est ainsi **endroit commun**, je suis étonné que les générateurs du jouet de la nation n'ont pas **branché dans le mouvement** et n'ont pas inventé une poupée de petit pâté de Puking que nous pourrions donner à nos filles pour leurs anniversaires.

(SPI\_BF, 36-39)

En fait, la salive en haut de bébé est **la place ainsi commune**, je suis étonné les fabricants de jouet nationaux n'ont pas **sauté sur le train en marche** et ont inventé une poupée de Petit pâté Puking que nous pourrions donner à nos filles pendant leurs anniversaires. (SPI\_REV, 36-39)

Les constructeurs de ces systèmes ont considéré que *jump on the bandwagon* 'sauter sur le train en marche' est une expression courante qui doit être incluse dans le dictionnaire de l'outil. En même temps, un bon nombre de verbes à particules reste inconnu. L'adjectif *commonplace* 'banal' n'a pas été reconnu à cause de sa graphie fautive<sup>67</sup>.

## 2.4.3 Lexies figées du français

### 2.4.3.1 Locutions

La proportion de ces expressions dans un texte dépend du style de l'auteur mais aussi du but du texte. Les textes journalistiques peuvent en contenir à volonté.

Mardi 22 mai, les autorités russes **ne donnaient pas cher de** la ville. (SIB\_f, 24)  
 Tuesday May 22, the Russian authorities **did not give a city expensive**. (SIB\_BF, 23)  
 On Tuesday, May 22, the Russian authorities **did not look expensively of** the city.  
 (SIB\_REV, 23-24)

Le texte parle d'une ville inondée. *Ne pas donner cher de qqch* signifie 'considérer que cette chose court de grands dangers'<sup>68</sup> et il pourrait se traduire par *to not give much for the chances of something*<sup>69</sup>. Ici, une solution raisonnable serait de formuler la phrase autrement, sans utiliser de locution figée, par ex. *the Russian authorities did not give the city a good prognosis*. La transformation de *la ville* (défini) en *a city* (indéfini) dans la traduction de BF est encore un des mystères de la logique de ces

<sup>67</sup> A l'origine, il s'écrivait en deux mots, puis avec tiret, mais aujourd'hui uniquement en un mot, Murray *et al* 1989 : 571.

<sup>68</sup> Robert 1985 : 546.

<sup>69</sup> Pilard *et al.* 2000 : 165.

machines. La transformation de *donner* en *look* ‘regarder’ dans la traduction de Reverso suggère que la locution ait été traduite par une autre locution, une locution qui n’a pourtant pas beaucoup de sens.

"Vive la France !" (PRE\_f, 43)  
 " Lives France! " (PRE\_BF, 43)  
 " Long live France! " (PRE\_REV, 47)

*Vive la France* est une locution idiomatique par excellence ; il faut la connaître pour savoir la traduire. L’évaluation de ces deux traductions est donc des plus simples, le dictionnaire de Reverso contient cette expression, celui de BF non.

Sharon appelle les Palestiniens à **faire le premier pas** (SHA\_f, 1)  
 Sharon invites the Palestinians to **take the first step** (SHA\_BF, 1)  
 Sharon calls the Palestinians to **make(do) the first step** (SHA\_REV, 1)

Le français contient un bon nombre d’expressions formées de *faire* + NOM (*faire du vélo, faire du pain, faire un cadeau*). Ces expressions se traduisent rarement telles quelles, mais elles demandent un autre verbe et parfois une autre structure syntaxique (*ride a bike, bake bread, give a gift*). *Faire un pas* se traduit par *take a step*.

M. Vlassov **s'en est pris** aux services de prévision, qui avaient... (SIB\_f, 94-95)  
 Mr. Vlassov **was caught some with** the services of forecast, which had... (SIB\_BF, 90-91)  
 Mr Vlassov **set himself it in** the services of forecast, which had... (SIB\_REV, 77)

*S'en prendre à (quelque chose ou quelqu'un)* signifie ‘incriminer, accuser’. Cette locution est analogue au verbe à particule propre à l’anglais dans la mesure où elle contient un verbe (*prendre* ou *se prendre*) et une particule (*en*) qui change complètement le sens du verbe. Les deux systèmes la traduisent comme ils peuvent. BF a analysé *s’est pris* comme un passif (*was caught* ‘a été pris’), ce qui n’est pas complètement faux. *En* a été traduit par *some*<sup>70</sup>, ce qui serait parfaitement acceptable dans un autre contexte. Et effectivement, la préposition *à* correspond parfois à *with*, notamment dans les descriptions d’objets : *a knife with a silver handle* ‘un couteau à manche d’argent’, *a girl with blue eyes* ‘une jeune fille aux yeux bleus’<sup>71</sup>. Dans la traduction de Reverso *s’est pris* se traduit par *set himself* ‘s’est mis’, ce qui se justifie par le fait que *se prendre* s’utilise parfois pour dire *se mettre*. Le choix de *in* pour traduire *à* doit venir du même fait. Le partitif *en*, qui, du point de vue de la machine à

<sup>70</sup> Quand *some* désigne une quantité, il se traduit souvent par le partitif.

<sup>71</sup> Pilard *et al.* 2000 : 1104.

traduire, ne devrait pas faire partie de la phrase, a été traduit par un simple nominatif *it*.

**Il en va de même** des réseaux qui indiquent la façon dont un gène agit (BIO\_f, 20-21)  
**The same applies to** networks which indicate the way in which a gene acts (BIO\_BF, 20-21)  
**It's the same** networks which indicate the way a gene acts (BIO\_REV, 18)

*Il en va de même de (quelque chose)* est également une locution figée contenant exclusivement des mots fréquents, la plupart ayant un sens qui varie selon le contexte. BF connaît cette expression et il l'a bien traduite. Reverso a analysé le mot *des* en tant qu'article indéfini pluriel (*des réseaux*) bien qu'il s'agisse ici de la forme amalgamée de la préposition *de* et de l'article défini pluriel (*de + les réseaux*).

**J'ai du mal** à voir pourquoi. (BOY\_f, 40-41)  
**I have evil** to see why. (BOY\_BF, 39-40)  
**I have difficulty** in seeing why. (BOY\_REV, 39-40)

*Avoir du mal* ne veut pas dire avoir des relations avec le Mal (*evil*) mais simplement avoir des difficultés à faire quelque chose.

#### 2.4.3.2 Structures grammaticales

Il existe un certain nombre de structures qui sont formées de plusieurs mots et qui ont une valeur avant tout grammaticale.

« après la pub, **il y aura** une surprise » (BOY\_f, 55)  
 " after the advertizing, **there will be** a surprise " (BOY\_BF, 53-54)  
 " after the advertisement, **there will be** a surprise " (BOY\_REV, 55)

**Il y a là** pêle-mêle un album de Gérard Lenorman, des crèmes au caramel et une voiture à un prix exceptionnel. (BOY\_f, 57-59)  
**It A there** pêle-mixes **there** an album with Gerard Lenorman, custards caramel and a car at an exceptional price. (BOY\_BF, 56-58)  
**There is there** confusedly Gérard Lenorman's album, crème caramel and car in an exceptional price. (BOY\_REV, 57-59)

L'utilisation de la structure *il y a* n'est pas recommandé dans la langue écrite soignée. Dans le premier exemple elle fait pourtant partie d'une citation, ce qui justifie son utilisation. Là elle est reconnue par les deux systèmes. Dans le deuxième exemple elle est combinée à *là*, ce qui ajoute de la redondance à la phrase. BF s'en trouve complètement bouleversé. Tous les éléments de la phrase sont traduits, sauf *a* (*It 'il', there 'y', there 'là'*), mais l'analyse de l'ensemble n'est pas réussie, ce qui fait que le résultat final est incompréhensible. Reverso traduit fidèlement tous les mots, *il y a* par *there is* et *là* également par *there*. Ce qu'il ne prend pas en compte

c'est qu'il faudrait traduire *il y a* par *there are* car le sujet réel de la phrase (*un album de Gérard Lenorman, des crèmes au caramel et une voiture*) est une énumération qui contient un élément au pluriel. De toute manière, la traduction de cette phrase est peu idiomatique. Il faut cependant reconnaître que même la phrase originale est assez maladroite.

## 2.5. L'ordre des mots

### 2.5.1 L'ordre des mots en anglais

#### 2.5.1.1 Noms modifiés par d'autres noms

Quand BF rencontre deux ou plusieurs noms adjacents qui ne sont pas notés comme mots composés dans son dictionnaire<sup>72</sup> il applique une règle très simple : il inverse l'ordre des noms et ajoute des *de* entre eux. Reverso applique en gros la même règle, comme on pourra le voir dans les exemples suivants.

... Galapagos Oil Spill Captain (OIL\_a, 4-5)  
 ... Le Capitaine De Flaque De Pétrole De Galapagos (OIL\_BF, 4-5)  
 ... Capitaine de Déversement accidentel de pétrole des Galapagos (OIL\_REV, 3-4)

Pour une bonne partie des phrases, c'est la bonne solution, mais cela peut aussi mener à des traductions trompeuses, voire incompréhensibles. Nous en avons quatre exemples.

... The currents are still very dangerous," **Coast Guard spokesman Tod Lyons** said. (OIL\_a, 66-67)  
 ... Les courants sont toujours très dangereux, " **porte-parole Tod de la garde côtière Lyon** dit. (OIL\_BF, 75-77)  
 ...Les courants sont toujours très dangereux," **le porte-parole de Garde-côte Tod Lyons** a dit. (OIL\_REV, 55-57)

Dans le premier exemple, BF a confondu une apposition avec la pré-modification, et de plus, il a mal interprété le mot *Lyons* (ici un nom de famille et non le nom anglais de la ville de Lyon). Normalement, *Coast Guard Spokesman* devrait se traduire par

---

<sup>72</sup> Le système de SYSTRAN utilise un dictionnaire d'expressions (*dictionary of multiple word expressions*) qui contient notamment les mots composés et qui lui permet de traduire p. ex. *sea lion* par *otarie* plutôt que de le traduire mécaniquement par *lion de mer* (voir OIL\_a, 27-29), Anon. 2000 : 1.

quelque chose comme (le) *porte-parole des garde-côtes*, et le nom *Tod Lyons* devrait rester tel quel, éventuellement précédé d'une virgule. Les composants sont donc bien traduits (sauf *Lyons*) mais le tout est mal analysé. Reverso n'est pas tombé dans ce piège.

... I want no say whatsoever in **Conservative Party policy or appointments**. (TOR\_a, 58-60)

... je ne veux aucune parole quelque dans **la politique ou les rendez-vous conservateurs de partie**. (TOR\_BF, 64-65)

... je ne veux pas disent que dans **la politique(police) de Partie(parti) Conservatrice ou des rendez-vous(nominations)**." (TOR\_REV, 61-63)

Dans le deuxième cas, toutes les deux machines ont mal analysé les liens des modifications, mais non au même endroit. Ce n'est pas *la politique* ni *les rendez-vous* qui sont conservateurs, mais bien *le parti* (et non *la partie*) lui-même, donc une traduction correcte serait *la politique ou les nominations du parti conservateur*. En fait, le groupe *Party policy* est un exemple de prémodification qui correspond à une forme plus explicite avec une préposition, c'est-à-dire *party policy = policy of the party*<sup>73</sup>. Reverso a bien analysé le lien entre l'adjectif *Conservative* 'conservateur' et le nom *Party* 'parti', que ce premier définit, mais par contre, il n'a pas vu que le groupe *Conservative Party* 'le parti conservateur' modifie non seulement le nom *policy* 'la politique' mais aussi le nom *appointments* 'nominations'.

The whole science of genetics, covering more than a century, has been based on **tracking differences** in inherited traits... (GEN\_a, 53-54)

La science entière de la génétique, couvrant plus qu'un siècle, a été basée sur des **différences de cheminement** dans des traits hérités... (GEN\_BF, 71-72)

La science entière de génétique, couvrant plus d'un siècle, a été basée **en dépistage de différences** de traits hérités... (GEN\_REV, 63-64)

Dans le troisième exemple, ainsi que dans le dernier, il s'agit également de difficultés dans le découpage des unités de sens. BF a analysé le participe présent *tracking* du verbe *track* 'suivre à la piste, traquer' en tant que nom (d'où *cheminement*) qui définit un autre nom, c'est-à-dire, *differences* 'différences'. Si la phrase avait été formulée d'une manière plus explicite (...*based on the tracking of differences* 'basée sur la recherche de différences'), elle serait plus facile à traduire, mais moins idiomatique. Reverso a su traduire *tracking differences*, car il connaît visiblement une traduction indépendante pour le mot *tracking* 'dépistage'.

---

<sup>73</sup> Greenbaum – Quirk 1990 : 387.

... ocean currents have taken most of the diesel and bunker fuel, **used to power tour boats**, away from the islands ... (OIL\_a, 69-70)

... les courants d'océan ont pris la majeure partie du diesel et du carburant de soute, **utilisé pour actionner des bateaux d'excursion**, loin à partir des îles (OIL\_BF, 79-81)

Des courants Jusqu'ici océaniques ont pris la plupart du diesel et le carburant de soute, **ont employé aux bateaux de tour de pouvoir(puissance)**, loin des îles ... (OIL\_REV, 59-60)

Ici c'est Reverso qui a mal analysé un groupe de mots. Dans cette phrase *power* 'puissance' n'est pas un substantif mais un verbe ('actionner, alimenter'). Il ne s'agit donc pas de *\*[power tour] boats* 'bateaux de tour de puissance' mais de *tour boats* 'bateaux d'excursion, de croisière' qui sont *powered* 'actionnés, alimentés' par le carburant. La confusion peut résulter en partie d'une mauvaise analyse de *used to* 'utilisé pour', participe passé qui remplace une phrase relative *which are used to* 'qui sont utilisés pour'.

Assez souvent on rencontre des traductions maladroites mais compréhensibles de la forme *nom + nom*.

... and needed **hospital treatment** for... (OIL\_a, 34)

... et a eu besoin du **traitement d'hôpital** pour ... (OIL\_BF, 39)

... et a eu besoin du **traitement d'hôpital** pour... (OIL\_REV, 28-29)

The Galapagos are home to iguanas, flightless cormorants, sea lions and **the famous Galapagos giant tortoises**. (OIL\_a, 27-29)

le Galapagos sont à la maison aux iguanas, aux cormorants incapables de voler, aux otaries et aux **tortues célèbres de géant de Galapagos** (OIL\_BF, 31-34)

Les Galapagos sont domestiques aux iguanes, flightless des cormorans, des otaries et les **tortues de géant des Galapagos célèbres** (OIL\_REV, 23-24)

**Shoreline impact**, however, had been minimal. (OIL\_a, 68)

**l'impact de shoreline** de , cependant, avait été minimal. (OIL\_BF, 78-79)

**L'impact de littoral**, cependant, avait été minimal. (OIL\_REV, 58)

Le premier exemple, bien que peu idiomatique, est encore facile à comprendre, tandis que le deuxième évoque d'abord l'image d'un géant des Galapagos qui possède des tortues célèbres (BF), ou encore que les célèbres Galapagos sont habitées par un géant qui a des tortues (Reverso). Dans le premier exemple, une meilleure solution serait d'utiliser une autre préposition (*traitement à l'hôpital*). Dans le deuxième cas, il serait préférable de considérer *giant* comme adjectif (*tortues géantes des Galapagos*) et déplacer l'adjectif *célèbre* vers le début du groupe nominal, pour mieux distribuer le poids des adjectifs (*les célèbres tortues géantes des Galapagos*). Dans le dernier cas également une autre préposition éviterait un faux sens ; le littoral ne cause pas l'impact mais il le subit, donc *l'impact sur le littoral*.



### 2.5.1.2 Adjectifs épithètes

En anglais, l'adjectif épithète est normalement placé devant le nom qu'il définit<sup>74</sup>. En français, la place de l'épithète dépend de l'adjectif, mais la plupart des adjectifs sont postposés par rapport au nom<sup>75</sup>. Voilà pourquoi, en traduisant, il faut souvent changer l'ordre du nom et de l'adjectif.

- ... it serves as a **very useful** signpost for that gene. (GEN\_a, 52)
- ... il sert de poteau indicateur **très utile** à ce gène. (GEN\_BF, 69)
- ... il sert comme un poteau indicateur **très utile** pour ce gène. (GEN\_REV, 61-62)
  
- ... an **unexpected** boost ... (TOR\_a, 6)
- ... une **poussée inattendue** ... (TOR\_BF, 6-7)
- ... une **augmentation inattendue** (TOR\_REV, 6)

Quand, pour des raisons d'organisation de la phrase, l'adjectif se trouve après le nom, BF a des difficultés de trouver le nom qu'il modifie.

- Today, because of the genome effort, places in our DNA where those differences occur have been cataloged to a **degree** of detail **unimaginable** just a few years ago. (GEN\_a, 22-23)<sup>76</sup>
- Today, en raison de l'effort de génome, endroits en notre ADN où ces différences se produisent ont été catalogués à un **degré** de juste **inimaginable** de détail il y a quelques années. (GEN\_BF, 30-31)
- Aujourd'hui, à cause de l'effort de génome, les places dans notre ADN où ces différences arrivent ont été cataloguées à un **degré de détail inimaginable juste** il y a quelques années. (GEN\_REV, 25-27)

Dans le texte anglais, c'est *a degree of detail* 'un degré de détail' qui était *unimaginable just a few years ago* 'inimaginable il y a seulement quelques années'. Dans la traduction de BF, par contre, *just* 'seulement' a été analysé en tant que nom<sup>77</sup> défini par l'adjectif *unimaginable* 'inimaginable', qui le précède. L'ordre des phases de l'analyse est important. Il serait possible d'imaginer que les petits groupes nominaux qui forment la phrase sont analysés d'abord. Ensuite, ces groupes sont remis à leur place dans la phrase. BF ne prend pas en considération l'adjectif *inimaginable* qui définit *just* mais il considère seulement le groupe *\*detail just* (deux noms), qui se traduit, selon les règles, *juste de détail*.

<sup>74</sup> Greenbaum – Quirk 1990 : 364.

<sup>75</sup> 95% des adjectifs préfèrent la postposition, mais les adjectifs privilégiant l'antéposition (beau, grand, bon) sont parmi les plus courants, Grevisse 1993 : 498.

<sup>76</sup> En anglais, le verbe *cataloguer* a deux graphies différentes : *catalogue* (britannique) et *catalog* (américain), Pilard *et al.* 2000, 142.

<sup>77</sup> En anglais le mot *just* peut s'utiliser comme adjectif, adverbe (le plus courant) et enfin comme nom, mais seulement soit dans un contexte littéraire : *to sleep the sleep of the just* 'dormir du sommeil du juste', soit au pluriel *the just* 'les justes', Pilard *et al.* 2000, 512. En anglais on peut former des noms pluriels désignant un groupe simplement en ajoutant l'article défini devant l'adjectif correspondant, par ex. *the blind* 'les aveugles', *the old* 'les vieux'.

Il est possible que ces difficultés sont liées au fait que BF ne connaît pas le mot anglais *unimaginable*. Reverso, lui, le connaît et, en conséquence, a mieux traduit cette phrase. Il est également possible que BF connaisse cet adjectif mais qu'il n'ait pas été analysé correctement uniquement à cause de sa place peu habituelle dans la phrase.

Parfois, la modification d'un adjectif par un adverbe peut poser des problèmes, surtout s'il y a plusieurs adverbes dans la phrase.

Chimps, for instance, are **about** two to three times **more genetically variable** than we are (GEN\_a, 104-105)

Chimps, par exemple, sont **variable d'environ** deux à trois fois **plus génétiquement** que nous sommes (GEN\_BF, 137-138)

Les chimpanzés, par exemple, sont **environ** deux à trois fois **plus génétiquement la variable** que nous sommes (GEN\_REV, 124-125)

La phrase veut dire, à la base, que les chimpanzés sont plus variables que nous, *chimps ... are... more ... variable than we are*. L'analyse de la phrase devient plus difficile quand les adverbes *genetically* 'génétiquement' et *about* 'environ' sont ajoutés. Vu sa position, *more* semble définir davantage l'adverbe *genetically* que l'adjectif *variable*, comme si la variation chez les chimpanzés était plus génétique que chez les êtres humains. La chaîne des modifications (aussi bien *about two to three times more* que *genetically* modifient *variable*) est difficile à découper et à réorganiser, mais il faut le faire pour rendre la phrase compréhensible en français. Une possibilité serait *Les chimpanzés ... sont environ deux à trois fois plus variables génétiquement que nous (le sommes)*.

... if there are too many genes acting in **too complicated a manner**, it may never be possible to find them... (GEN\_a, 80-81)

...s' il y a trop de gènes agissant dans **trop compliqué une façon**, il peut jamais être possible de les trouver... (GEN\_BF, 105-106)

... s'il y a trop de gènes agissant dans **trop compliqué une façon**, il ne peut jamais être possible de les trouver... (GEN\_REV, 93-95)

Si un adjectif épithète est modifié par l'adverbe *so* 'tellement' ou *too* 'trop', ce groupe adjectival doit se placer devant l'article<sup>78</sup>. Ex. *it's too interesting a movie to miss* 'c'est un film trop intéressant pour le manquer', *it's so interesting a movie that...* 'c'est un film tellement intéressant que...' Cela a apparemment perturbé les

---

<sup>78</sup> Greenbaum – Quirk 1990 : 334.

deux machines, qui n'ont pas su faire le rapport entre l'adjectif *compliqué* et le nom *façon*.

The fact that we are **so similar** implies that... (GEN\_a, 106)

Le fait que nous sommes **si le semblable** implique que... (GEN\_BF, 138-139)

Le fait que nous sommes si **semblables** implique que... (GEN\_REV, 126-127)

Un *so* devant un adjectif attribut a également gêné BF, *similar* ayant été analysé en tant que nom. Il est possible que la phrase ait été découpée ainsi : [*the fact that we are so*][*similar implies that ...*] '[le fait que nous sommes ainsi] [le semblable implique que]', le sujet original de la phrase, *the fact*, étant trop loin du verbe *implies*. Cette fois-ci, donc, le problème ne serait pas lié à l'adverbe *so*, mais à la construction de la phrase en générale. Notons encore qu'il est tout à fait acceptable de laisser *so* tout seul à la fin d'une phrase<sup>79</sup>, mais ici cela semble peu opportun, la phrase d'origine voulant dire : *le fait que nous soyons tellement semblables (entre nous) implique que....*

### 2.5.1.3 La place des prépositions dans la phrase

En anglais il est possible de séparer la préposition du nom qu'elle définit. Cela se fait couramment dans le cas des verbes ditransitifs. Les verbes ditransitifs sont des verbes qui exigent deux compléments : le complément d'objet direct (COD) et le complément d'objet indirect (COI). *To give* 'donner' est un verbe ditransitif par excellence. En anglais les compléments peuvent être dans n'importe quel ordre mais normalement le COI est avant le COD et dans ce cas la préposition qui définit le COI est omise, p. ex. *John gave Jack a book* = *John gave a book to Jack* 'John a donné un livre à Jack'.<sup>80</sup> En français, par contre, la préposition ne peut normalement pas être séparée de son régime<sup>81</sup>.

WILLIAM HAGUE was given an unexpected boost... (TOR\_a, 6)

WILLIAM HAGUE a été donné à une poussée inattendue... (TOR\_BF, 6)

WILLIAM on a donné à la HAYE une augmentation inattendue... (TOR\_REV, 6)

La phrase française produite par BF est grammaticalement correcte, mais elle n'est pas exactement la traduction de la phrase anglaise ; dans cette phrase ce n'est pas « une poussée inattendue » qui a eu William Hague, mais bien M. Hague qui a eu

<sup>79</sup> *I'm afraid so* 'J'en ai bien peur', *that's so* 'c'est bien vrai', Pilard et al. 2000 : 895.

<sup>80</sup> Greenbaum – Guirk 1990 : 205-206.

une poussée. BF a apparemment ajouté l'élément qu'il s'attendait à rencontrer dans la phrase anglaise, à savoir la préposition *to*, ce qui lui a permis d'ajouter à la phrase française. Reverso, de sa part, a parfaitement saisi la phrase *HAGUE was given* 'on a donné à la HAYE', mais il n'a pas vu que le mot *William* fait également partie du sujet de cette phrase. Le sujet n'est donc pas *(The) Hague* 'La Haye', une ville néerlandaise.

Cette phrase pourrait aussi se dire : *An unexpected boost was given to William Hague* et cet ordre de mot serait plus facile à analyser pour une machine, et surtout plus facile à respecter dans une phrase française : 'une poussée inattendue a été donnée à William Hague', où tous les éléments (à part l'inversion nom – adjectif) occupent le même endroit que dans la phrase anglaise. L'auteur du texte a probablement choisi cet ordre de mots pour pouvoir mettre le nom de la personne à l'évidence, en tête de phrase (et de tout l'article).

This should give **researchers clues** to the movements of different peoples during history.  
(GEN\_a, 94-95)

Ceci devrait donner à **des chercheurs des indices** aux mouvements de différents peuples pendant l'histoire. (GEN\_BF, 124-125)

Cela doit donner **des indices de chercheurs** aux mouvements de peuples différents pendant l'histoire. (GEN\_REV, 112-114)

Dans cette phrase à l'actif BF a réussi à analyser les compléments comme il le faut, *researchers* 'chercheurs' étant le COI et *clues* 'indices' étant le COD. Il est donc probable que le problème de l'exemple précédent est dû au fait que la phrase anglaise est au passif. Reverso a analysé *\*researchers clues* 'indices de chercheurs' comme un groupe nominal formé de deux noms adjacents et il a attribué à ce groupe la fonction de COD dans cette phrase. Cette interprétation est tout à fait compréhensible, vu que plus loin dans la phrase il y a un groupe prépositionnel *to the movements* 'aux mouvements', qui, vu sa structure, pourrait être le COI du verbe *give* 'donner'. La traduction est presque une phrase compréhensible mais elle n'a pas la même signification que la phrase d'origine.

---

<sup>81</sup> Grevisse 1993 : 1486.

Le même phénomène est également observable dans le cas de verbes demandant un complément indirect. Ce complément peut être par ex. le COI ou un complément de lieu, par ex. *aller à Paris*.

**The more parents I talk to**, however, the more I learn I wasn't alone. (SPI\_a, 24-25)  
**Plus que je parle à parents**, cependant, plus j'apprends que je n'étais pas seul. (SPI\_BF, 26)  
**Plus parents à qui je parle**, cependant, plus j'apprennent que je n'étais pas seul. (SPI\_REV, 26-27)

Le choix de cet ordre de mots est imposé par la structure *the more ... the more* 'plus... plus'. Si l'auteur n'avait pas détaché la préposition du nom qu'elle définit, la comparaison se ferait sur le verbe *talk* et non sur le nom *parents* : *The more I talk to parents* (ce à quoi correspond ici la traduction de BF). La traduction de Reverso a le mérite de ne pas donner de faux sens, mais seulement parce qu'elle n'en a aucun de compréhensible. En français la comparaison *plus... plus* ne peut se faire que sur les verbes (*plus on lit, plus on apprend*) ou les adjectifs (*plus ce vin est vieux, plus il est délicieux*)<sup>82</sup>. Le contenu de la phrase originale serait donc en français *Plus nombreux sont les parents auxquels je parle, plus j'apprends...*

## 2.5.2 L'ordre des mots en français

### 2.5.2.1 Compléments des noms

En anglais il est possible de modifier des mots simplement en les antéposant. En français la même chose se fait à l'aide d'une préposition, le plus souvent *de* (pour les compléments de relation, parfois *à*, pour les compléments de matière, parfois *en*)<sup>83</sup>. En générale, les compléments du nom se placent après le nom qui modifie (la tête du groupe nominale)<sup>84</sup>.

Ça commençait bien pourtant ; comme dans **un clip de David et Jonathan**. (TEL\_f, 105-107)  
 That well however started ; as in **a clip of David and Jonathan**. (TEL\_BF, 106-107)  
 That began well nevertheless ; as in **David's clip and Jonathan**. (TEL\_REV, 110-112)

Dans cet exemple il s'agit d'un complément de relation. BF évite le problème de l'analyse des liens de référence en adoptant le même ordre de mots que dans la phrase originale. Il n'y a aucun moyen pour savoir si la phrase anglaise veut dire [a

<sup>82</sup> Grevisse 1993 : 368.

<sup>83</sup> Grevisse 1993 : 531, 538.

*clip*] of [David and Jonathan] ou bien [a clip of David] and [Jonathan]. Reverso a choisi de traduire N<sub>1</sub> de N<sub>2</sub> ( [un clip] de [David et Jonathan] ) par N<sub>2</sub>'s N<sub>1</sub> ( [David and Jonathan] 's [clip]), ce qui nécessite une analyse approfondie de la phrase. Dans ce cas, l'analyse est manquée : N<sub>2</sub> est formé de *David et Jonathan* et non seulement de *David*.

En aval de Lensk, un "bouchon" de glace de 50 kilomètres de long... (SIB\_f, 50-52)  
 Downstream from Lensk, a "stopper" of 50 kilometers length ice... (SIB\_BF, 49-50)  
 Downstream to Lensk, a "cork" of ice(mirror) 50 kilometres long... (SIB\_REV, 46-47)

Dans cet exemple le mot *bouchon* a été redéfini avec le mot *glace*, ce qui nous a donné un nouveau concept, un *bouchon de glace*, *glace* étant dans ce cas un complément de caractérisation<sup>85</sup>. Ce mot peut être traduit mécaniquement, comme l'a fait Reverso, par *a cork of ice*. Cette solution est probablement la meilleure possible, vu qu'il s'agit d'un terme inventé pour les besoins de ce texte particulier et que ce mot n'existe pas dans le vocabulaire de manière permanente. L'analyse de cette phrase est rendue difficile par la présence d'un deuxième complément de caractérisation, *de 50 kilomètres de long*. BF a découpé la phrase ainsi : [un "bouchon"] [(de) (glace) (de 50 kilomètres) (de long)], ce qui a donné respectivement [a "stopper "] [(of) (50 kilometers) (length) (ice)]. Pour BF donc la tête du syntagme nominal n'est pas *bouchon*, qui est complètement séparé du reste de la phrase, mais *glace*. Conséquemment, *50 kilometers* et *length* ont été transférés devant le nom *ice*. Nous ne savons pourtant pas pourquoi l'adjectif *long* n'a pas été traduit par un adjectif mais par le nom correspondant *length*. Dans la traduction de Reverso, le contenu du groupe adjectival *50 kilometers long* est juste mais son emplacement est discutable. Il serait préférable d'écrire *a 50-kilometer long cork of ice* ou encore *a cork of ice which was 50 kilometers long*.

- 2) Choisissez ensuite la direction de traduction à appliquer en cliquant sur le bouton qui correspond à **la paire de langue choisie**, (AID-f, 6-7)
- 2) Choose then the direction of translation to be applied while clicking to the button which corresponds to **the language pair chosen**, (AID\_BF, 6-7)
- 2) Please , choose then the translation direction to be applied by clicking the button which corresponds to **the pair of chosen language**, (AID\_REV, 6-7)

<sup>84</sup> Grevisse 1993 : 526.

<sup>85</sup> Grevisse 1993 : 527.

Dans ce dernier exemple, il s'agit d'un pseudo-complément<sup>86</sup>. *Une paire de langue* (ou de préférence *languages*) se traduit sans problème par *a language pair* ou *a pair of languages*, mais quand ce groupe est modifié par un adjectif postposé, comme ici par *choisi*, la phrase est plus difficile à analyser. L'adjectif modifie la tête du GN, dans cette phrase *paire*. BF a laissé l'adjectif après le GN, en choisissant un ordre de mots qui n'est pas idiomatique en anglais. Reverso a attribué l'adjectif *choisi* au mauvais nom. La bonne traduction serait *the chosen language pair* ou *the chosen pair of languages*.

### 2.5.2.2 Autres compléments

L'ordre des mots habituel d'une phrase affirmative en français est SVO (sujet – verbe – objet / attribut du sujet). Il y a pourtant des exceptions à cette règle, notamment l'ordre SOV quand l'objet (direct ou indirect) est remplacé par un pronom.

... il est venu voir le joli feu d'artifice **qui lui est offert** par Monsieur le Député-Maire.  
(TEL\_f, 10-12)

... it came to see the pretty fireworks **which is offered to him** by Mister the Deputy-Mayor.  
(TEL\_BF, 10-12)

... it(he) came to see the beautiful fireworks **which is offered to him(her)** by Sir Representative - mayor. (TEL\_REV, 10-12)

Cet ordre de mots n'est apparemment pas inconnu des deux systèmes et il a été bien analysé dans cette subordonnée relative au passif. *Qui* est le sujet et *lui* le COI.

A l'Egypte et à la Jordanie, **les deux seuls pays** du Proche-Orient avec lesquels Israël a signé des traités de paix, **il a demandé** qu'ils aident « à rétablir la paix et le calme dans toute la région ». (SHA\_f, 57-60)

In Egypt and Jordan, **the two only countries** of the Middle East with which Israel signed peace treaties, **it required that they help** " to restore peace and calm it in all the area ".  
(SHA\_BF, 53-55)

In Egypt and in Jordan, **both only countries** of the Near East with which Israel signed peace treaties, **he asked that they help** " to restore the peace and the peace in all the region ".  
(SHA\_REV, 60-63)

Le COI du verbe *demandé* (*à l'Egypte et à la Jordanie*) est placé au début de la phrase, probablement pour faciliter l'insertion d'une longue antéposition (*les deux seuls pays du Proche-Orient avec lesquels Israël a signé des traités de paix*) qui modifie ces deux noms. Sans cette antéposition il serait plus naturel d'écrire *Il a*

---

<sup>86</sup> Grevisse 1993 : 526.

*demandé à l’Egypte et à la Jordanie qu’ils aident...* L’ordre des mots peu habituel de cette phrase a fait que les groupes *à l’Egypte et à la Jordanie* ont été analysés en tant que compléments de lieu (bien qu’avec la plupart des noms des pays soit utilisée la préposition *en* : *en Egypte, en Jordanie*). De plus, Reverso a mal analysé *les deux seuls pays* en traduisant *les deux* par *both* ‘chacun des deux’.

### 2.5.2.3 D’autres cas

**Arrive** enfin **la surprise** tant attendue. (BOY\_f, 69)  
**Arrive** finally the so much awaited **surprise**. (BOY\_BF, 68)  
**The surprise** so waited **arrives** finally. (BOY\_REV, 69)

Dans cette phrase, pourtant assertative, le verbe vient avant le sujet, pour des raisons de style. Aucune des deux traductions n’est parfaitement satisfaisante. Il semble que BF a analysé le mot *arrive* en tant qu’impératif s’adressant à *la surprise*, ce qui expliquerait la forme anglaise *arrive* et sa position en début de phrase. Reverso a bien vu que *la surprise* est le sujet du verbe *arriver*, mais la place du groupe adjectival *so waited* ‘tant attendu’ par rapport au nom *surprise* qu’il définit n’est pas très idiomatique. Une meilleure traduction pourrait être *The so-awaited surprise arrives finally*.

Ariel Sharon a appelé **hier** les Palestiniens à « cesser la violence immédiatement et à revenir à la table des négociations » (SHA\_f, 3-4)  
 Ariel Sharon invited the Palestinians **yesterday** " to cease violence immediately and to return to the table of the negotiations " (SHA\_BF, 3-4)  
 Ariel Sharon called the Palestinians " to stop **yesterday** the violence at once and to return to the table of the negotiations " (SHA\_REV, 3-4)

Bien que la place de *hier* dans la phrase française soit toute à fait grammaticale, elle a apparemment perturbé Reverso dans l’analyse de la phrase. Le verbe *appeler* signifie ici ‘demander, inviter’ et il demande deux compléments, un COD (*les Palestiniens*) et un infinitif (*à cesser la violence*). Reverso a analysé *appeler* comme un simple verbe transitif, d’où le verbe *call* ‘téléphoner à’ ou ‘nommer’. Rien ne justifie pourtant le nouvel emplacement de *hier, to stop yesterday the violence* ‘à arrêter **hier** les violences’. La citation est devenue quelque peu absurde, vu qu’elle contenait déjà une autre expression de temps, *immédiatement*, qui est incompatible avec *hier*.



## 2.6. Les temps

### 2.6.1 Les temps de l'anglais

#### 2.6.1.1 Remarques préliminaires

Face à un texte à traduire de l'anglais en français, un traducteur humain hésite souvent entre les temps du passé (voir ch.1.3). Une machine qui n'a pas de connaissance du monde affronte encore plus de difficultés.

#### 2.6.1.2 La traduction du prétérit anglais

Le prétérit anglais est un temps particulièrement problématique, car il peut correspondre à trois différents temps du français, c'est-à-dire l'imparfait, le passé composé et le passé simple.

For a long time, my co-workers **thought I lived** near the beach because I frequently **showed up** with a milky-white patch on my shoulder (SPI\_a, 14-16)

Pendant longtemps, mes collègues **ont pensé** que **j'ai vécu** près de la plage parce que **j'ai fréquemment révélé** avec une connexion laiteux-blanche sur mon épaule (SPI\_BF, 14-16)

Pendant une longue période de temps, mes collaborateurs **ont pensé** que **j'ai vécu** près de la plage parce que je **démasquais** fréquemment avec une pièce laiteuse-blanche sur mon épaule... (SPI\_REV, 15-17)

Ici on utiliserait de préférence l'imparfait à la place du passé composé *j'ai fréquemment révélé*<sup>87</sup>, car il s'agit d'une action répétée (l'adverbe *frequently* en donne un indice), aussi bien qu'à la place de *j'ai vécu* car il fait partie d'un discours indirect au passé (cf. *Ils pensent que je vis, ils pensaient (ou ont pensé) que je vivais*). L'imparfait serait également envisageable pour remplacer *ont pensé* car les limites de la période durant laquelle l'action s'est déroulée ne sont pas clairement indiquées. Reverso a probablement su profiter de l'information offerte par l'adverbe *frequently* 'fréquemment', qui lui a permis de choisir l'imparfait plutôt que le passé composé pour traduire *I showed up* 'je me présentais'. Pour les autres verbes, il a pourtant préféré le passé composé.

---

<sup>87</sup> Bien que *show up*, s'il est suivi d'un objet, ait la signification de *révéler (les fautes d'une personne), démasquer*, une meilleure traduction dans cette phrase intransitive serait *venir, se présenter*, Murray et al. 1989 : 360-361.

Nous avons remarqué que le verbe *said* peut être difficile à analyser. BF traduit la forme du passé *said* tantôt par le passé composé *a dit*, tantôt par le participe passé *dit*, voire par l'adjectif *ledit*. Le choix entre ces formes se fait selon ce qui suit immédiatement le verbe dans la phrase d'origine. Si *said* est suivi d'un groupe nominal ou d'une phrase subordonnée qui ne peut être que le COD, il se traduit par le passé composé.

He **said that** he had given the money... (TOR\_a, 34)  
 Il **a dit qu'**il avait donné l'argent... (TOR\_BF, 36)  
 Il **a dit qu'**il avait donné l'argent... (TOR\_REV, 36)

A Conservative spokesman **said**: "We are very grateful... (TOR\_a, 43)  
 Un porte-parole conservateur **a dit**: " nous sommes très reconnaissants... (TOR\_BF, 45-46)  
 Un porte-parole Conservateur **a dit**: "nous sommes très reconnaissants à la générosité... (TOR\_REV, 45-46)

Si cette forme est suivie d'un groupe nominal dont la fonction dans la phrase est peu claire, BF la traduit plus facilement par l'adjectif *ledit*.

He **said last night**: "I think William Hague would... (TOR\_a, 29)  
 Il **ladite nuit passée**: " je pense que William Hague ferait... (TOR\_BF, 30)  
 Il **a dit hier dans la nuit**: "je pense que William Hague ferait... (TOR\_REV, 30)

Dans ce cas, *ledit* est associé au groupe nominal qui le suit, *ladite nuit*<sup>88</sup>. Reverso a bien analysé cette phrase.

Dans certains cas où la principale suit la citation, et le verbe *said* est à la fin de la phrase, BF le traduit par le participe passé.

Arevalo, previously a Navy officer based in the Galapagos, could face a jail term of four to five years if charged and convicted, local police **said**. (OIL\_a, 52-54)  
 Arevalo, précédemment un officier de marine basé dans Galapagos, a pu faire face à une limite de prison de quatre à cinq ans si chargé et police condamnée et locale **dite**. (OIL\_BF, 58-61)  
 Arevalo, précédemment un officier Naval basé dans les Galapagos, pourrait faire face à un terme de prison de quatre à cinq ans si chargé et reconnu coupable, la police locale **a dit**. (OIL\_REV, 43-45)

... some animals are going to die," Dr. Fernando Espinosa of the Charles Darwin Foundation ecology center **said**. (OIL\_a, 104-106)  
 ... quelques animaux vont mourir, " Dr. Fernando Espinosa du centre d'écologie de base de Charles Darwin **dit**. (OIL\_BF, 118-119)  
 ... quelques animaux vont mourir," docteur Fernando Espinosa de Charles Darwin le centre d'écologie de Base(fondation) **a dit**. (OIL\_REV, 88-90)

---

<sup>88</sup> *Ledit* se place devant le nom qu'il définit, tandis que les autres participes passés sont postposés, Grevisse 1993 : 501.

Nous ne trouvons d'autre explication pour ce phénomène que le fait que *say* soit un verbe transitif. Comme BF ne trouve pas de COD possible dans la phrase, il en déduit que *said* n'est pas un verbe à l'imparfait mais autre chose, dans ce cas participe passé utilisé comme adjectif qualificatif.

### 2.6.1.3 Les temps composés anglais

Nous avons repéré des cas intéressants où BF a traduit le *present perfect* avec l'aspect progressif par un plus-que-parfait. Reverso, par contre, semble savoir affronter cette forme verbale sans problème (l'échec de la traduction de *with* est dû aux facteurs techniques expliqués dans le chapitre 1.2).

With the advent of molecular biology, researchers in recent years **have been working** with known SNPs... (GEN\_a, 60-61)

With l'arrivée de la biologie moléculaire, chercheurs ces dernières années **avaient fonctionné** avec SNPs connu... (GEN\_BF, 80-81)

Avec l'apparition de biologie moléculaire, les chercheurs **ont** ces dernières années **travaillé** avec connu SNPs (GEN\_REV, 71-72)

La bonne traduction de *researchers ... have been working* serait *des chercheurs ... ont travaillé (continuellement)*. BF a traduit *working*, un participe présent, par *fonctionnant*. Il est vrai que *work* peut vouloir dire aussi bien *travailler* (quand on parle de personnes) que *fonctionner* (quand on parle de machines ou de systèmes)<sup>89</sup>. Il s'agirait donc en quelque sorte de *chercheurs qui fonctionnent*. BF a probablement analysé à part *have been* 'ont été'. Les deux fragments (*ont été* + *fonctionnant* en tant qu'adjectif qualificatif qui définit *chercheurs*) mis ensemble pourraient, avec beaucoup d'imagination, se traduire par *avaient fonctionné*.

BF semble suivre le même procédé quand il rencontre la structure *have been* + participe présent, comme on peut le constater également dans l'exemple suivant.

Local fishermen, who **have been scooping** up the oil by hand with small plastic buckets protected by flimsy masks, **fear** their livelihoods and tourism on the island **are** in peril (OIL\_a, 107-111)

les pêcheurs locaux de , qui **avaient écopé** vers le haut de l'huile à la main avec de petites positions en plastique se sont protégés par les masques fragiles, **craignent** que leurs vies et tourisme sur l'île **soient** dans le péril (OIL\_BF, 120-121)

---

<sup>89</sup> Pilard *et al.* 2000 : 1108.

Des pêcheurs locaux, qui **ont été scooping** en haut l'huile(pétrole) à la main avec des petits seaux en plastique protégés par des masques fragiles, **la crainte** leurs gagne-pain et tourisme dans l'île **sont** dans le péril (OIL\_REV, 91-93)

Rien dans la phrase ne donne de raison pour mettre le premier verbe au plus-que-parfait, au contraire le fait que le reste de la phrase soit au présent devrait aider à choisir un temps lié au présent, dans ce cas le passé composé *ont ramassé (avec une pelle)*. Quant à la solution de Reverso, la traduction du *present perfect* avec l'aspect progressif a échoué, car il n'a pas connu le mot *scooping*<sup>90</sup>. De plus, le verbe *fear* 'craindre' a été analysé en tant que substantif, probablement à cause de la distance entre le sujet, *local fishermen*, et le verbe.

L'exemple suivant pourra justifier notre hypothèse sur la progression de la traduction de BF.

"Everyone wants to do these studies, but so far we've **been doing** them in a very incomplete way (GEN\_a, 27-28)

"Everyone veut pour faire ces étude, mais jusqu'ici nous **ont été faisant** ils dans très inachevé voie (GEN\_BF, 36)

"Chacun veut faire ces études, mais jusqu'ici nous les **avons faits** d'une façon très incomplète (GEN\_REV, 31-32)

*Have been* (ici sous la forme '*ve been*) a donné *ont été* [probablement *them* 'les (forme à l'accusatif de *ils*)' a été analysé comme le sujet] et *doing* s'est traduit par *faisant*. Nous ne savons pas pourquoi ici BF n'a pas fait la même transformation que dans les autres cas où il a rencontré *have been* + participe présent. Reverso a bien analysé le *present perfect* mais l'accord du participe passé n'est que partiellement juste. Il faudrait dire *nous les avons faites*, car *étude* est un nom féminin.

Il semble que c'est l'aspect progressif qui gêne BF car *have been* + participe passé se traduit correctement, malgré la distance entre le sujet (*places*) et le verbe.

... **places** in our DNA where those differences occur **have been cataloged** (GEN\_a, 22-23)

... **endroits** en notre ADN où ces différences se produisent **ont été catalogués** (GEN\_BF, 30-31)

... **places** dans notre ADN où ces différences arrivent **ont été cataloguées** (GEN\_REV, 25-26)

D'autre part, BF traduit également le plus-que-parfait anglais par un plus-que-parfait français, ce qui est tout à fait compréhensible, vu que les emplois de ce temps se

---

<sup>90</sup> *To scoop* 'écoper', *a scoop* 'une pelle', Pilard *et al.* 2000 : 813.

correspondent dans les deux langues. Ici le plus-que-parfait est utilisé dans un discours indirect.

Galapagos National Park Director Eliecer Cruz said oil **had contaminated** a 500-yard stretch of beach (OIL\_a, 87-89)

Directeur national Eliecer Cruz de parc de Galapagos a dit que l'huile **avait souillé** un bout droit de 500-yards de plage (OIL\_BF, 99-102)

les Galapagos le Directeur de Parc national Eliecer Cruz a dit de l'huile(pétrole) **avait contaminé** une période(bout) de 500 yards de plage (OIL\_REV, 74-75)

## 2.6.2 Les temps du français

### 2.6.2.1 Les temps du passé

Le système des temps du passé du français est bien différent de celui de l'anglais. Le passé composé ou le passé simple du français se traduit par un *present perfect* ou le prétérit anglais. D'une autre part, l'imparfait français peut correspondre à ce même prétérit ou encore au prétérit avec l'aspect progressif.

Elle **regardait** son imbécile. Lequel **réussit ce soir-là** l'exploit de communiquer ainsi pendant une heure et demie. (TEL\_f, 95-98)

It **looked at** its imbecile. Which **succeeds in that evening** the exploit to communicate thus during an hour and half. (TEL\_BF, 96-99)

She(it) **looked** imbecile at sound. Who(which) **manages this evening** of the exploit to communicate so during one hour and a half. (TEL\_REV, 99-103)

Dans cette phrase *réussit*, un passé simple, a été analysé en tant que présent. Comme ces deux formes de ce verbe ont la même graphie, il est normal que les machines se trompent. De plus, comme la phrase subordonnée relative dont le verbe fait partie (*lequel réussit...*) est coupée de sa principale par un point final au lieu d'une virgule, l'analyse en devient encore plus difficile.

En aval de Lensk, un "bouchon" de glace de 50 kilomètres de long, avec des pics atteignant par endroits 17 mètres de haut, **s'était constitué** en quelques jours (SIB\_f, 50-54)

Downstream from Lensk, a "stopper" of 50 kilometers length ice, with peaks reaching by places 17 meters top, **had been constituted** in a few days (SIB\_BF, 49-52)

Downstream to Lensk, a "cork" of ice(mirror) 50 kilometres long, with peaks achieving by places 17 metres high, **had established(constituted)** in some days (SIB\_REV, 46-48)

Dans cet exemple la combinaison plus-que-parfait et verbe pronominal a été mal analysée. La traduction de BF pourrait être considérée juste dans un autre contexte, car il est vrai que le pronom réfléchi *se* peut s'utiliser pour exprimer le passif (*s'utiliser = être utilisé*). Ici, pourtant, ce n'est pas le cas. Nous ne pouvons pas

considérer que quelqu'un ait constitué le « *bouchon* » *de glace*, mais il s'est constitué de lui-même. La traduction de Reverso omet l'objet du verbe *constituer*, qui est un verbe transitif tout comme *establish* et *constitute*, les deux traductions qu'il propose. Il serait donc mieux de dire *a cork of ice... had established itself* 'un bouchon de glace s'était constitué (lui-même)'. Une solution qu'il faut aller chercher un peu plus loin serait *a cork of ice... had formed*.

Iakoutsk, capitale de la République Sakha, un territoire grand comme six fois la France au cœur de la Sibérie orientale, **tente depuis** trois jours d'échapper à la furie (SIB\_f, 8-10)  
 Iakoutsk, capital of the Republic Sakha, a large territory like six times France with the cœur of Eastern Siberia, **has tried for** three days to escape from fury (SIB\_BF, 7-9)  
 Yakutsk, capital of the Republic Sakha, a big territory as six times France in the cœur of the Eastern Siberia, **tries for** three days to escape the fury (SIB\_REV, 8-10)

La construction PRESENT + *depuis* est utilisée pour exprimer une action qui a commencé dans le passé (souvent récent) et qui continue encore dans le moment de l'énonciation. Quand *depuis* introduit la durée de l'action, la construction se traduit par exemple comme BF l'a traduit ici. Une autre traduction possible serait *has been trying* 'a été en train d'essayer', pour traduire l'aspect duratif de l'action. La solution de Reverso, la traduction par un présent anglais, n'est pas idiomatique ni compréhensible.

Ces deux pays **n'ont plus** d'ambassadeur à Tel-Aviv **depuis** le début de l'intifada. (SHA\_f, 60-61)  
 These two countries **do not have more** ambassador with Such-Aviv **since** the beginning of the intifada. (SHA\_BF, 55-57)  
 These two countries **have no more** ambassador in Tel Aviv **since** the beginning of the intifada. (SHA\_REV, 63-64)

Dans le cas où *depuis* introduit le moment dans le temps depuis lequel l'action à lieu l'utilisation du présent dans la phrase anglaise est plus acceptable. BF s'est trompé dans l'analyse de *ne ... plus*, qui doit se traduire par *no more* ou *not... anymore*. De plus, la traduction de *Tel-Aviv* par *Such-Aviv* (*such* 'tel') ne contribue pas à la compréhension de la phrase. Il serait également possible d'utiliser le *present perfect* mais sans l'aspect duratif (*these two countries have not had an ambassador in Tel Aviv any more since...*). Les deux machines ont fait la différence entre ces deux utilisations de *depuis* dans la traduction de la préposition elle-même (*for three days* 'depuis trois jours' vs. *since the beginning of the intifada* 'depuis le début de l'intifada'), mais, faute d'autres exemples, il est impossible de dire si cette différence a influencé la traduction des formes verbales.

### 2.6.2.2 Impératif

L'impératif français est facile à repérer dans le texte, du moins celui qui utilise la forme de politesse, car une forme verbale qui n'est pas précédée d'un sujet et qui finit en *-ez* ne peut être qu'un impératif.

**Cliquez** sur l'onglet 'Texte' :

- 1) **Tapez** le texte source dans la zone de gauche, ou **importez** du texte par simple copier-coller ou glisser-déposer depuis votre application,
- 2) **Choisissez** ensuite la direction de traduction à appliquer en cliquant sur le bouton qui correspond à la paire de langue choisie, (AID\_f, 3-7)

**Cliquez** on the mitre ' Texte':

- 1) **Type** the source text in the zone of left, or **import** text by simple copy-sticking or slip-depositing since your application,
- 2) **Choose** then the direction of translation to be applied while cliquant to the button which corresponds to the language pair chosen, (AID\_BF, 3-7)

**Please** , **click** the tab 'Text':

- 1) **Please** , **type** the source text in the left zone, or **matter** of the text by simple to copy to stick or drag and drop since your application,
- 2) **Please** , **choose** then the translation direction to be applied by clicking the button which corresponds to the pair of chosen language,  
(AID\_REV, 1-7)

Toutes les deux machines s'en sortent plutôt bien, si l'on excepte le fait que BF n'ait pas reconnu le mot *cliquer* et que Reverso ait choisi une mauvaise traduction parmi les différents sens du verbe *importer*. Reverso a un souci particulier de la politesse, puisqu'il ajoute *please* 's'il vous plaît' devant chaque impératif. Cela peut parfois être la bonne solution, mais seulement si utilisée avec modération. Ici, comme il s'agit d'un mode d'emploi, les *please* sont excessifs.

« Ohé, Ducon-Malin ! c'est pour toi, **attrape** ! » (TEL\_f, 154-156)

" Ohé, Ducon-Malignant ! it is for you, **catches** ! " (TEL\_BF, 152-154)

" Ohé, Ducon-Malin ! It is for you, **catches** ! " (TEL\_REV, 162-164)

Dans le cas d'un impératif plus familier, l'analyse est plus difficile. Les deux machines ont analysé la forme *attrape* comme une forme de la III<sup>e</sup> personne du singulier, à laquelle sa graphie est identique, et l'ont traduite conséquemment.

## 2.7. Noms propres

### 2.7.1 Remarques préliminaires

Un outil de traduction ne sait pas toujours faire la différence entre un mot à traduire et un nom propre, l'utilisateur lui-même devant lui indiquer les mots à laisser tels quels. Quand on utilise BF, il est possible de le faire soit en tapant un point immédiatement devant le nom, par ex. *.Babel.Fish* (sans les points, le nom deviendrait « des poissons de Babel »), soit en l'entourant avec des *x* (*xBabelx xFishx*). Ce petit travail de préédition ne devrait pas prendre trop de temps, surtout s'il s'agit d'un texte relativement court. Cependant, il est impossible d'éditer les pages web, qui sont très fréquemment traduites<sup>91</sup>. Bien sûr ce traitement n'est pas nécessaire quand il s'agit de noms propres qui ne sont pas homographes des mots de la langue source. Comme ce sont des mots inconnus pour le système, ils ne sont pas traduits (p.ex. *Hosni Mubarak*). Reverso ne propose aucun moyen de marquer les noms propres dans le texte.

### 2.7.2 Fautes dues aux noms propres dans les textes en anglais

Nous avons repéré un exemple intéressant qui illustre ce problème. BF a traduit au sein du même texte le nom *Cook* deux fois par *cuisinier* et trois fois par *Cook*. De plus, *Robin Cook* est devenu mystérieusement *le cuisinier de groupe de travail*. Reverso, de sa part, a traduit ce nom une fois par *le cuisinier* et ailleurs par *Cook*.

Israeli soldiers killed as **Cook** continues mission (TIM\_a, 3)

Les soldats israéliens détruits comme **cuisinier** continue la mission (TIM\_BF, 3)

Des soldats israéliens ont tué comme le **Cuisinier** continue la mission (TIM\_REV, 3)

Meanwhile, **Robin Cook**, the Foreign Secretary, flew to Egypt today... (TIM\_a, 20-21)

En attendant, le **cuisinier de groupe de travail**, le ministre des affaires étrangères, a volé en

Egypte aujourd'hui... (TIM\_BF, 23-24)

En attendant, **Robin Cook**, le Ministre des Affaires Etrangères, a volé en Egypte aujourd'hui (TIM\_REV, 22-23)

**Mr Cook** met President Hosni Mubarak of Egypt... (TIM\_a, 24)

**M. Cook** a rencontré le Président Hosni Mubarak de l'**Egypte**... (TIM\_BF, 28)

**M. Cook** a rencontré le Président Hosni Mubarak de l'**Egypte**... (TIM\_REV, 26)

---

<sup>91</sup> En 1998, 40% des demandes de traduction étaient des pages web et 60% des documents tapés par les utilisateurs, Yang – Lange 1998 : 4.



Dans les cas où *Cook* est interprété comme nom propre, il est précédé de *Mr.* Il semble donc que la présence de cet élément permette à ces deux machines de considérer le nom comme nom propre. Il est à noter que certains noms géographiques figurent dans le dictionnaire de BF (*Egypt – Egypte, Middle East – Moyen-Orient*).

Dans plusieurs dictionnaires d'anglais<sup>92</sup>, les seules significations données à *robin* sont 'rouge-gorge' et 'rossignol' (deux oiseaux de la famille des turdidés). La consultation des dictionnaires généraux ne permet donc pas d'expliquer pourquoi BF traduit le mot *robin*, même pris isolément, toujours par *groupe de travail*. Il est donc probable que *robin* soit un terme scientifique. Dans le langage sportif de l'anglais d'Amérique, *round robin* signifie un type de compétition où chaque joueur ou équipe joue contre chacun des autres joueurs ou équipes<sup>93</sup>. Il est imaginable que la notion d'*équipe* ou de *groupe de travail* dérive de cette signification. D'autre part, en informatique on appelle *Round-Robin* un protocole de file d'attente où le travail qui arrive le premier dans la file est exécuté le premier<sup>94</sup>. Un protocole de file d'attente est nécessaire quand plusieurs personnes (= un groupe de travail ?) utilisent les mêmes ressources (par exemple quand elles partagent une imprimante), mais aussi quand une seule personne utilise un ordinateur multi-tâches. Quoi qu'il en soit, cette explication semble moins satisfaisante que celle relative au langage sportif.

### 2.7.3 Fautes dues aux noms propres dans les textes en français

Parfois le fait que le dictionnaire contienne des termes spécifiques peut avoir des conséquences surprenantes.

Clarisse Fabre et Michel Noblecourt (PRE\_f, 101)

Clarisse Fabre and Michel Noblecourt (PRE\_BF, 104)

Nun of the order of St Clare Fabre and Michel Noblecourt (PRE\_REV, 113)

Une *clarisse* est effectivement une religieuse appartenant à l'ordre de sainte Claire, fondé au XIII<sup>e</sup> siècle<sup>95</sup>. Dans cette phrase, pourtant, il s'agit simplement du prénom d'une journaliste.

---

<sup>92</sup> Hanks 1991 : 1261 ; Engineer – Gray 1981 : 1392 ; Cowie 1989 : 1095.

<sup>93</sup> Engineer – Gray 1981 : 1403.

<sup>94</sup> McKinney 1983 : 1256.

<sup>95</sup> Robert 1985 : 643.

## 2.8. Différences quantitatives entre le français et l'anglais

### 2.8.1 Remarques préliminaires

Comme nous l'avons déjà constaté, la traduction ne peut pas être du simple décodage. Le nombre de mots exigés pour exprimer une idée varie d'une langue à l'autre. Parfois, en traduisant, il faut ajouter des éléments, parfois il faut en omettre.

### 2.8.2 Omissions en anglais

En anglais il est possible d'omettre des éléments syntaxiques, tels que le pronom relatif objet *that* 'que'.

They report several million places where people's genomes can differ from those of others, **changes researchers refer to** as "SNPs." (GEN\_a, 37-38)

l'état de They plusieurs million d'endroits où les génomes des personnes peuvent différer de ceux de d'autres, **chercheurs de changements se rapportent** en tant que " SNPs. (GEN\_f, 50-51)

Ils annoncent plusieurs millions de places où des génomes populaires peuvent différer de ceux d'autres, **changent les chercheurs se réfèrent à** comme "SNPs". (GEN\_REV, 44-45)

En fait, il s'agit ici d'un complément d'objet indirect (*to which*, 'auquel') qui est divisé et dont une partie est omise. La grammaire anglaise accepte l'omission du pronom relatif restrictif quand celui n'est pas le sujet de la relative (*the man I love* = *the man that I love*, 'l'homme **que** j'aime')<sup>96</sup>. Il est également possible d'exprimer exactement la même chose d'une manière moins ambiguë en écrivant *changes to which researchers refer*. Une traduction correcte de cette phrase serait 'changements auxquels les chercheurs se réfèrent'. Les deux machines ont chacun produit une traduction fautive. Celle de Reverso présente un mystère, car le mot *changes* 'changements' a été traduit par *changent*, troisième personne du pluriel du verbe *changer*. Or, la forme anglaise *changes* ne peut en aucun cas être analysé en tant que tel. BF, de sa part, a analysé *changes researchers* en tant que groupe nominal, bien que le premier de ces mots soit l'antécédent de la relative et le deuxième le sujet de la relative.

---

<sup>96</sup> Greenbaum – Quirk 1990 : 369.

*That* ‘que’ en tant que conjonction peut également être omis s’il est le complément d’objet direct, un autre complément ou s’il est extraposé<sup>97</sup>.

And, what's especially frightening about that, is a parent is never tagged while wearing a ratty T-shirt... (SPI\_a, 25-26)

Et, ce qui est particulièrement effrayant à ce sujet, est un parent n'est jamais étiqueté tout en portant un t-shirt ratty... (SPI\_f, 27-29)

Et, ce qui est particulièrement effrayant de cela, est un parent n'est jamais étiqueté en portant un T-shirt grincheux(miteux)... (SPI\_REV, 27-29)

Cette omission est fréquente en anglais mais il est tout aussi acceptable d’écrire *what’s especially frightening [--] is that a parent...* ‘ce qui est particulièrement effrayant [--] c’est qu’un parent...’ En français il faut utiliser l’introducteur *c’est* pour introduire cette proposition conjonctive essentielle<sup>98</sup>. Évidemment, si le texte est destiné à un traitement automatique, il est conseillé de ne pas omettre ces éléments.

Especially sickening is I'm right there begging for handouts. (HAL\_a, 17)

Rendre malade particulièrement est moi suis bon prier là pour des aumônes. (HAL\_BF, 18)

Particulièrement pour dégoûter est je prie juste là pour des aumônes(prospectus). (HAL\_REV, 18)

Les constructions de cette phrase ressemble à celles de l’exemple précédent. Elle peut également être exprimée d’une manière plus explicite : *It is especially sickening that I'm right there...* ‘Il est particulièrement révoltant que je sois (juste) là...’ Si la phrase précédente s’est traduite d’une manière compréhensible, ici l’omission de *that* combinée à l’omission de *it is* ‘il est’ a mené à une traduction ayant peu de valeur communicatif.

### 2.8.3 Expressions françaises contenant moins de mots que celles anglaises

Il y a les belles bleues. Les belles rouges. (TEL\_f, 14)

There are the beautiful blue ones. Beautiful reds (TEL\_BF, 14-15)

There are the beautiful blue. The beautiful Red (TEL\_REV, 14-15)

Le texte parle de feux d’artifice. En français il est possible d’exprimer les qualités d’un objet sans devoir en répéter le nom. En anglais il serait préférable d’utiliser la formule *the + adj. + one(s)*, ce qu’a fait BF pour la première phrase. Il est vrai que

<sup>97</sup> L’extraposition est la substitution d’un élément par une autre forme et le déplacement du premier vers la fin la phrase, Greenbaum – Quirk 1990:306, 417.

dans certains cas il suffit d'ajouter *the* devant l'adjectif singulier pour former le pluriel<sup>99</sup>, mais ici cette solution semble peu idiomatique.

#### 2.8.4 La négation en français

La négation en français demande deux éléments, dont le premier est la particule *ne*, le deuxième pouvant varier entre *pas*, *plus*, *aucun(e)* et *rien*. En anglais cette négation dite « double » est considérée fautive. La phrase *je n'ai rien dit* ne doit pas se traduire par *(\*I didn't say nothing* mais par *I said nothing* ou par *I didn't say anything*.

si un mot est mal écrit, il y a de grandes chances que Reverso Online **ne le traduise pas** correctement. (AID\_f, 84-85)

if a word is badly written, there are great chances that Reverso Online **does not translate it** correctly. (AID\_BF, 82-83)

if a word is badly written, there are big chances that Reverso Online **does not translate him(it)** correctly. (AID\_REV, 86-87)

Dans les cas simples les deux machines traduisent la négation correctement.

Quarante pour cent de nos gènes environ peuvent être supprimés (individuellement) **sans que se manifeste aucune conséquence perceptible**. (BIO\_f, 95-98)

Forty percent of our genes approximately can be removed (individually) **without appearing no perceptible consequence**. (BIO\_BF, 92-93)

Forty percent of our genes about can be killed(abolished) (individually) **but that demonstrates itself no perceptible consequence**. (BIO\_REV, 86-88)

La traduction de cette phrase demande quelques transformations de structure. Une possibilité serait : *Approximately forty percent of our genes can be (individually) removed without the appearance of any perceptible consequence*. La solution de BF est clairement manquée, car la construction *without appearing* demande un sujet explicite<sup>100</sup>, qui manque à cette phrase au passif. De même, *without... no consequence* n'est pas acceptable dans l'anglais écrit, mais il faudrait écrire *without... any consequence*. La solution de Reverso est très intéressante et presque créative, mais également fautive. Traduire *sans que se manifeste aucune conséquence perceptible* par *but that demonstrates itself no consequence* 'mais cela

<sup>98</sup> Une proposition conjonctive essentielle est une proposition qui remplit une fonction nominale essentielle (sujet, complément etc.) dans la phrase, Grevisse 1993 : 1596, 1602-1603.

<sup>99</sup> Cf. chapitre 2.5.1.2.

<sup>100</sup> Une utilisation acceptable de cette construction serait par ex. *She told the joke without laughing* (= *She told the joke and she did not laugh*) 'Elle a raconté la blague sans rire (Elle a raconté la blague et elle n'a pas ri)'

ne manifeste **à soi-même** aucune conséquence' dévoile une erreur dans l'analyse de *se*. Il ne s'agit pas de « manifester des conséquences à quelqu'un » mais de « conséquences qui se manifestent » (ou plutôt ne se manifestent pas dans ce cas).

De plus, le remplacement par *de* de l'article indéfini de l'objet direct dans une phrase négative cause des problèmes.

Trois obus de mortier tirés par des Palestiniens à partir de la bande de Gaza se sont abattus en territoire israélien et un quatrième sur la colonie de Rafah Yam, dans le sud de la bande de Gaza, **sans faire de victime**. (SHA\_f, 75-78)

Three mortar shells fired by Palestinians starting from the Gaza Strip fell down in Israeli territory and a fourth on the colony of Rafah Yam, in the south of the Gaza Strip, **without making victim**. (SHA\_BF, 69-71)

Three mortar shell pulled(fired) by Palestinians from Gaza Strip beat(came) down on Israeli territory and a fourth on Rafah Yam's colony, in the South of Gaza Strip, **without making(doing) of victim**. (SHA\_REV, 77-81)

« Si les Palestiniens n'entendent pas un mot clair sur la fin de l'occupation et des colonies, ils **n'ont pas d'autre choix que de continuer** la résistance au moyen de l'intifada », a-t-il dit. (SHA\_f, 47-50)

" If the Palestinians do not hear a clear word on the end of the occupation and the colonies, **they do not have an other choice to continue** resistance by means of the intifada ", has it says. (SHA\_BF, 43-46)

" If the Palestinians do not hear(understand) a clear word at the end of the activity(occupation) and the colonies, they **have not of the other choice than to continue** the resistance by means of the intifada ", he said. (SHA\_REV, 50-53)

Apparemment, Reverso ne supprime pas *de*, mais il le traduit dûment par *of*, bien que la grammaire anglaise ne le veuille pas. Dans la première phrase, BF a bien traduit la négation, mais dans la deuxième, il a omis la conjonction *que*, ce qui a inversé de quelque sort le sens à la phrase par rapport à l'originale : *they do not have an other choice to continue* 'ils n'ont pas d'autre choix de continuer'. La traduction de Reverso: *They have not of the other choice than to continue* 'ils n'ont pas de l'autre choix que de continuer', bien que contenant le mot *than* 'que', est tout aussi trompeuse.

### 3. Conclusion

Nous avons examiné des traductions générées par deux systèmes de traduction automatique (Babel Fish et Reverso, tous les deux en accès gratuit sur Internet), d'une part des traductions vers le français de textes en anglais et d'autre part des traductions vers l'anglais de textes en français. Le but de cette étude était d'étudier le genre de problèmes que la traduction automatique doit résoudre et d'établir la façon dont les outils de traduction automatique le font. Nous avons repéré des difficultés dans l'analyse des ambiguïtés (les homographes et les pronoms) et des groupes nominaux (les groupes *nom + nom* de l'anglais et les groupes *nom + de + nom* du français), ainsi que dans la traduction d'autres groupes de mots (les expressions) et des formes verbales (l'impératif et les temps du passé).

Notre première hypothèse était que les mots et les structures ambiguës soient une importante source d'erreurs dans l'analyse et la traduction. Effectivement, nous avons rencontré un bon nombre de traductions incompréhensibles ou du moins contraires aux textes originaux suite à une mauvaise analyse des homographes. Notamment les impératifs de l'anglais ont été analysés en majorité comme s'il s'agissait des noms. Les pronoms français ont également été mal analysés. Par contre, les mauvaises traductions dues aux structures ambiguës étaient relativement rares, ce qui implique que les textes du corpus sont bien rédigés de ce point de vue.

Les structures qui n'ont rien d'ambigu pour un lecteur humain peuvent pourtant l'être pour une machine. Notamment la modification des noms en anglais provoque des traductions parfois amusantes (cf. *les tortues célèbres de géant de Galapagos* vs. *Les célèbres tortues géantes des Galapagos*). Bien que les deux machines réussissent à faire les transformations nécessaires dans les cas simples, comme la transformation de la prémodification anglaise en postmodification française et l'inverse, dès que les phrases deviennent plus compliquées ou que l'on utilise des constructions peu courantes, la qualité de la traduction baisse. En bref, la TA ne sait pas s'y prendre avec la créativité du langage.

Cette étude nous a permis de nous rendre compte de l'abondance des expressions difficiles à traduire dans les textes français. Des structures aussi anodines que *les deux seuls pays* ou *il en va de même* ont été traduites avec un succès varié. Par contre un bon nombre de locutions de nature proverbial ont été reconnus par ces systèmes.

La traduction des impératifs réussit rarement parce que les impératifs sont souvent des homographes d'une autre forme verbale (par ex. en français *attrape* peut être un impératif ou le présent de l'indicatif de la III<sup>e</sup> personne du singulier) ou même d'un nom (en anglais *copy* peut être un impératif, un présent de l'indicatif ou un nom). Le passage du prétérit anglais aux temps correspondants français s'est également avéré particulièrement problématique.

La plus grande difficulté que nous avons rencontrée pendant l'étude était la question du classement des problèmes. Certaines phrases ont pu servir comme exemple sous plusieurs titres, par ex. les problèmes de traduction des lexies figées étaient souvent liés à une mauvaise analyse des homographes qu'elles contenaient.

Le but de cette étude n'était pas de comparer la qualité des traductions des deux machines, mais simplement d'illustrer certains problèmes que tous les systèmes de traduction et d'autres types de traitement automatique de langues devront affronter. Il serait néanmoins intéressant d'adopter une approche plus pragmatique à ce sujet et de pouvoir établir des critères pour l'évaluation de la qualité et de l'utilisabilité de la traduction automatique.

## 4. Bibliographie

### Corpus :

**AID** = Anon. *Aide*. <http://www.reverso.net/help.asp> (le 30 mai 2001).

**BIO** = Ball, Ph. 'Le génome décrypté donne les clés d'une biologie nouvelle',  
[http://www.lemonde.fr/imprimer\\_article/0,6063,148536,00.html](http://www.lemonde.fr/imprimer_article/0,6063,148536,00.html) (le 13 juin 2001).

**BOY** = Lazuly, P. 'Quand je pense aux boys band'.  
<http://www.menteur.com/cahiers/stars.shtml> (le 31 juillet 2001).

**FAQ** = Anon. *Babel Fish Frequently Asked Questions*.  
[http://live.altavista.com/s?spage=help/babel\\_faq.htm](http://live.altavista.com/s?spage=help/babel_faq.htm) (le 27 septembre 2000).

**GEN** = Mestel, R. 'Tiny Gene Disparities Go a Long Way',  
[http://www.latimes.com/news/science/science/lat\\_gene010212.htm](http://www.latimes.com/news/science/science/lat_gene010212.htm) (le 12 janvier 2001).

**HAL** = Swarner, K. 'Halloween candy overload'.  
<http://www.amused.com/columns/familyman/index.ph> (le 24 octobre 2000).

**OIL** = Gardner, S. '« My Fault, » Says Galapagos Oil Spill Captain',  
<http://www.latimes.com/wires/winternat/20010126/tCB00a1710.html> (le 26 janvier 2001).

**PRE** = Fabre, C. – Noblecourt, M. '2002 : le RPR et le PS préparent la campagne de leurs candidats', [http://www.lemonde.fr/imprimer\\_article/0,6063,203922,00.html](http://www.lemonde.fr/imprimer_article/0,6063,203922,00.html) (le 14 août 2001).

**SHA** = Anon. 'Sharon appelle les Palestiniens à faire le premier pas',  
[http://www.dna.fr/dna/jour/internat/rdmonde/7744\\_0.html](http://www.dna.fr/dna/jour/internat/rdmonde/7744_0.html) (le 23 mai 2001).

**SIB** = Bonnet, F. 'Des inondations en Sibérie provoquent l'évacuation de dizaines de milliers de personnes',  
[http://www.lemonde.fr/imprimer\\_article/0,6063,186772,00.html](http://www.lemonde.fr/imprimer_article/0,6063,186772,00.html) (le 22 mai 2001).

**SPI** = Swarner, K. 'Spit-Up',  
<http://www.amused.com/columns/familyman/index.php> (le 18 janvier 2001).

**TIM** = Anon. 'Israeli soldiers killed as Cook continues mission.'  
<http://www.thetimes.co.uk> (le 12 octobre 2000).

**TEL** = Lazuly, P. 'Heureux qui communique',  
<http://www.menteur.com/cahiers/mobiles.shtml> (le 21 juin 2001).



**TOR=** Webster, P. 'Donor stakes £5 million on Tories',  
<http://www.thetimes.co.uk/article/0,2-70091,00.html> (le 18 janvier 2001).

### Ouvrages consultés :

**Adams 1973** = Adams, V. *An Introduction to Modern English Word-formation*. London.

**Adams 1982** = Adams, D. *Le guide galactique*. Tr. de l'anglais par Jean Bonnefoy. Paris.

**Anis 1994** = Anis, J. 'Ordinateurs et traduction : survol d'un demi-siècle', *Langages* 116 : 111-122.

**Anon. 2000a** = 'Past and Present', <http://www.systransoft.com/PastPresent.html> (le 27 septembre 2000).

**Anon. 2000b** = 'Machine Translation', [http://www.systransoft.com/mt\\_arch.html](http://www.systransoft.com/mt_arch.html) (le 24 octobre 2000).

**Anon 2000c** = 'PROMT's machine translation systems museum', <http://www.e-prompt.com/company/museum.shtml> (le 3 septembre 2001).

**Anon. 2001a** = 'Livre blanc traduction',  
<http://www.softissimo.com/products/infotrad.htm> (le 22 mai 2001).

**Arnold et al. 1994** = Arnold, D. – Balkan, L. – Meijer, S. – Humphreys, R.L. – Sadler, L. *Machine Translation. An Introductory Guide*. Londres.

**Ball 1999** = Ball, C. *Machine Translation Links*.  
[http://www.gerogetown.edu/cball/ling361\\_mt\\_links.html](http://www.gerogetown.edu/cball/ling361_mt_links.html) (le 3 septembre 2001).

**Boitet – Blanchon 1994** = Boitet, C. – Blanchon, H. 'Promesses et problèmes de la « TAO pour tous » après LIDIA-1, une première maquette', *Langages* 116 : 20-47.

**Cowie 1989** = Cowie, A. éd. *Oxford Advanced Learner's Dictionary of Current English*. Aylesbury.

**Dubois et al. 1973** = Dubois, J. – Giacomo, M. – Guespin, L. – Marcellesi, Ch. – Marcellesi, J-B. – Mevel, J-P. éd. *Dictionnaire de linguistique*. Paris.

**Engineer – Gray 1981** = Engineer, S. – Gray, H. éd. *Longman Dictionary of the English Language*. Hong Kong.

**Green 1987** = Green, J. éd. *Dictionary of Jargon*. London / New York.


- Greenbaum – Quirk 1990** = Greenbaum, S. – Quirk, R. *A Student's Grammar of the English Language*. Hong Kong.
- Greenfield-Cardey 1996** = Greenfield-Cardey, S. 'Grammaires électroniques et ambiguïtés', *LINX (Linguistique Institut Nanterre Paris X)* 34-35 : 315-325.
- Grevisse 1993** = Grevisse, M. – Goosse, A. *Le Bon Usage*<sup>13</sup>. Grammaire française refondue par A. Goosse. Paris – Louvain-la-Neuve.
- Hanks 1991** = Hanks, P. éd. *Collins Dictionary of the English Language*. Glasgow.
- Ladmiral 1994** = Ladmiral, J-R. 'Le traducteur et l'ordinateur', *Langages* 116 : 5-19.
- Laporte 1996** = Laporte, É. 'L'évaluation des résultats de la levée d'ambiguïtés lexicales', *LINX (Linguistique Institut Nanterre Paris X)* 34-35 : 291-305.
- Loffler-Laurian 1994** = Loffler-Laurian, A-M. 'La traduction automatique : Son utilisation par le « grand public »', *Langages* 116 : 87-94.
- McKinney 1983** = McKinney, J.M. 'Queuing theory' 1255-1257, in Ralston, A. éd. *Encyclopedia of Computer Science and Engineering*. New York.
- Murray et al. 1989** = Murray, J. – Bradley, H. – Craigie, W. – Onions, C. éd. *The Oxford English Dictionary*. Oxford.
- Pilard et al. 2000** = Pilard, G. – Cronin, G. – Larroche, L. – Stevenson, A. éd. *Harrap's Shorter Dictionary English – French / French – English*<sup>6</sup>. Edinburgh.
- Robert 1985** = Robert, P. *Le Grand Robert de la langue française*. Paris.
- Sinclair 1987** = Sinclair, J. éd. *Collins Cobuild English Language Dictionary*. Glasgow.
- Yang – Lange 2000** = Yang, J. – Lange, E. 'SYSTRAN on AltaVista. A user study on Real-time Machine Translation on the Internet', [http://www.systransoft.com/Papers/ppr\\_alta.html](http://www.systransoft.com/Papers/ppr_alta.html) (le 24 octobre 2000).

# Appendice 1 : l'interface de Babel Fish



[Search](#) | [Comparison Shop](#) | [Channels](#) | [Rewards](#) | [Email & Tools](#)

## World: translation

<a href="#">World Home</a>	<a href="#">News</a>	<a href="#">Regions</a>	<a href="#">Message Boards</a>
<b>BABEL FISH</b> 	<b>Translate Anything</b>		
Compare the translation with the original by clicking the "View Original Language" link on a translated web page.	<input checked="" type="radio"/> <b>Text</b> Enter text for translation:		
Find out about the <a href="#">Systran technology</a> , the brains of the Babel Fish.	<div style="border: 1px solid black; height: 80px; width: 100%;"></div>		
Questions? Check out our <a href="#">FAQs</a> .	<input type="radio"/> <b>Website</b> Or enter the Web Address of the page you wish to translate:		
Like this new version? Send us <a href="#">feedback</a>	<div style="border: 1px solid black; padding: 2px;">http://</div>		
	Translate from <a href="#">English to French</a> ▼	<input type="button" value="Translate"/>	<a href="#">How to use Babel Fish?</a>




# Appendice 2 : l'interface de Reverso

Reverso translate online

wysiwyg://principal.15/http://www.reverso.net/textonly/default\_net.asp

 un texte court ou une page web  
**à traduire ?**  
www.reverso.net

Softissimo's  
site in  


**MON REVERSO**      **MY REVERSO**

Texte    Page WEB    Comment.    (?) Aide

① Enter your text    Entrez votre texte  
Fügen Sie den Text ein    Escriba su texto

② The translation is here    Ici la traduction  
Die Übersetzung ist da    Aquí está la traducción

③

EN-FR  
DE-FR  
FR-EN  
FR-DE  
DE-EN  
EN-DE  
EN-SP  
SP-EN  
FR-RU  
EN-RU

Special characters    Caractères spéciaux  
Sonderzeichen    Caracteres especiales

à    é    î    ô    ù    Autre

Effacer - Clear - Löschen - Borrar

**E-MAIL**

[Conditions d'utilisation](#) - [Conditions of use](#)

© Softissimo, PROject MT Ltd. Tous droits réservés.