Lefer, M.-A. & De Clerck, M. (2021). L'apport des corpus intermodaux en lexicologie contrastive : Étude comparative de la traduction écrite et de l'interprétation simultanée des séquences de noms. In : S. Hanote & R. Nita (eds), *Morphophonologie, lexicologie et langue de spécialité*. Rennes : Presses Universitaires de Rennes, 145-162.

L'APPORT DES CORPUS INTERMODAUX EN LEXICOLOGIE CONTRASTIVE
ÉTUDE COMPARATIVE DE LA TRADUCTION ÉCRITE ET DE L'INTERPRÉTATION SIMULTANÉE DES SÉQUENCES DE NOMS

Marie-Aude LEFER & Marie DE CLERCK
Université catholique de Louvain, Belgique

Résumé

En linguistique contrastive, de nombreuses études reposent sur des corpus parallèles, constitués de textes sources et de leur traduction écrite. À l'heure actuelle, de nouveaux corpus parallèles sont en train de voir le jour. Ces corpus contiennent non seulement des textes écrits (sources et cibles), mais aussi des transcriptions de discours oraux et de leur interprétation simultanée. Dans cet article, nous illustrons l'utilisation de ce type de ressource en lexicologie contrastive, en revisitant une dissymétrie bien connue entre l'anglais et le français, à savoir les difficultés liées à la traduction en français des séquences de noms de l'anglais. Notre étude montre que les données issues de l'interprétation apportent un éclairage nouveau sur ces difficultés de traduction.

Abstract

Numerous contrastive studies rely on parallel corpora, i.e. corpora made up of original texts in Language A and their translation into Language B. Today new parallel corpora are being compiled. In addition to written source and target texts, they contain transcripts of source speeches and their simultaneous interpretation. This paper sets out to show how this new type of corpus can be used in the field of contrastive lexicology. To do so, we revisit a well-known English-French contrast, namely the translation challenges caused by English sequences of nouns. Our study shows that interpreting data can shed new light on the translation difficulties triggered by these units.

Introduction

En linguistique contrastive, de nombreuses études reposent sur des corpus parallèles (également appelés corpus de traduction), typiquement constitués de textes sources et de leur traduction écrite. À l'heure actuelle, de nouveaux corpus parallèles, appelés corpus intermodaux, sont en train de voir le jour. Ces corpus contiennent non seulement des textes écrits (sources et cibles), mais aussi des transcriptions de discours oraux et de leur interprétation simultanée. Cette approche, mise au point par Shlesinger (1998) et Shlesinger & Ordan (2012) dans la foulée des travaux pionniers de Baker (1993, 1995), a amené quelques chercheurs à rassembler des corpus parallèles intermodaux (Bernardini *et al*., 2016 ; Kajzer-Wietrzny, 2012). En traductologie et en interprétologie, ces ressources sont utilisées pour étudier les spécificités de l'interprétation simultanée et de la traduction écrite, étant entendu qu'il s'agit là de deux formes de médiation

interlinguistique. Cette approche novatrice peut s'avérer bénéfique pour la linguistique contrastive, dans laquelle il est désormais possible d'explorer des données orales interprétées, en plus des données traduites principalement utilisées jusqu'à présent.

Notre article, situé dans le domaine de la lexicologie contrastive anglais-français, porte sur l'interprétation simultanée et la traduction écrite des séquences de noms (par exemple, *tuna imports* > *importations de thon*). Nous revisitons ainsi un contraste bien connu entre l'anglais et le français, sur lequel Michel Paillard s'est penché dans plusieurs de ses travaux (voir Paillard 2000, 2011), à travers le prisme de l'approche intermodale, qui, à notre connaissance, n'a pas encore été utilisée en linguistique contrastive.

Dans la Section 1, nous proposons un bref aperçu des principaux contrastes entre les noms composés de l'anglais et du français. La Section 2 est ensuite consacrée à la description (i) du corpus utilisé (le *European Parliament Translation and Interpreting Corpus*, EPTIC), (ii) des séquences de noms extraites d'EPTIC et (iii) de la taxonomie des procédés de traduction que nous avons utilisée. Dans la Section 3, nous présentons les résultats de notre étude de cas. Enfin, nous concluons en soulignant les nouvelles perspectives qu'ouvre l'approche intermodale en lexicologie contrastive basée sur corpus.

1. Les noms composés et les séquences de noms : contrastes anglais-français

Notre étude part d'une dissymétrie lexicale bien connue entre l'anglais et le français : les noms composés anglais du type N+N. Comme l'a souligné Paillard (2000 : 49-51), le schéma[1] N+N est bien plus productif en anglais qu'en français (cf. aussi Van Hoof, 1989 : 18-19). En effet, « le processus est exploité avec la plus grande souplesse en anglais où il aboutit à la juxtaposition de composants liés par des relations syntactico-sémantiques très variées » (*sunshine* : « the sun shines » [sujet-verbe], *handlebar* : « the bar is a handle » [identification], *searchlight* : « X searches with the light » [instrumental], *call box* : « X calls from the box » [locatif]) (Chuquet & Paillard, 1987 : 186). En français, par contre, le schéma N+N correspond presque exclusivement à une relation de localisation (*rayon hommes*, *coin cuisine*) (*ibid*.). Le schéma N+N est donc également attesté en français, mais sa productivité est nettement moins marquée (voir à ce propos Arnaud, 2003 ; Arnaud & Renner, 2014 ; Paillard, 2011 : 918).

En français, on a davantage tendance à recourir à des suffixés (*ice cube* > *glaçon*) ou des composés N+prép+N (*handbag* > *sac à main*) et N+A (*city council* > *conseil municipal*) pour traduire les composés N+N anglais (Paillard, 2000 : 49-51 ; Paillard, 2011 : 915-916 ; Arnaud & Renner, 2014 : 8 ; Astington, 1983 : 95). À propos de la correspondance entre les schémas N+N de l'anglais et N+prép+N du français, Chuquet & Paillard (1987 : 185) soulignent que « le type roman, avec préposition, explicite dans une certaine mesure la relation entre les termes du composé alors que celle-ci est totalement effacée dans le type germanique » (*tasse à thé* : *teacup*).[2] Cependant, ils ajoutent qu'il s'agit là d'une différence relativement superficielle car en français, la préposition tend à être grammaticalisée (*lunettes de soleil*, *propositions de paix*) (*ibid*.). De plus, Paillard (2000 : 51) précise que « la récursivité du processus, qui donne des "surcomposés" (*rack railway* : *chemin de fer à crémaillère*) joue en principe de part et d'autre, même si l'accumulation des prépositions peut en limiter l'application en français ».

[1] Le terme « schéma » est utilisé pour faire référence aux configurations syntaxiques des noms composés et des séquences de noms.

[2] Il est également intéressant de noter qu'il n'y a pas de consensus sur le statut (lexical *vs*. syntaxique) des unités N+prép+N en français (Paillard, 2000 : 46 ; Paillard, 2011 : 915).

La traduction de l'anglais vers le français pose deux types de problèmes. Le premier est lié à l'interprétation des séquences de noms, potentiellement ambiguës, comme *modern history section* ([*modern history*] *section* vs. *modern* [*history section*]) (Chuquet & Paillard, 1987 : 186 ; voir également Garnier, 2012). Le second problème, « c'est le caractère elliptique de la composition par juxtaposition qui fait difficulté vis-à-vis du français » (Chuquet & Paillard, 1987 : 187). En effet, le français requiert l'introduction de liens prépositionnels entre les éléments de la séquence (*first-year linguistics course* > *cours de linguistique de première année*), ce qui mène à des modifications syntaxiques au-delà de trois éléments (*Industrial Development Act Report* > *Rapport concernant la loi sur le développement industriel*) (*ibid*.).

En interprétation, Gile (1995) identifie les noms propres composés (les séquences de deux ou plusieurs substantifs et adjectifs reliés entre eux par des mots grammaticaux) comme des « déclencheurs de difficultés » ou « déclencheurs de problèmes », c'est-à-dire des « éléments et caractéristiques du discours original qui engendrent des problèmes de saturation et de déficit individuel ». Ainsi, selon Gile (*ibid*., 107) :

> « Ces noms propres présentent deux éléments de difficulté. D'une part, ils sont denses, car les mots pleins qui les composent ne sont séparés que par un petit nombre d'éléments de liaison à faible densité informationnelle (…). D'autre part, dans la plupart des cas, pour des raisons syntaxiques, leur restitution en langue d'arrivée demande un réagencement de ces éléments : par exemple, d'anglais en français :

1	2	3	4
« International	Association	of Conference	Interpreters »
« Association	internationale	des interprètes	de conférence »
2	1	4	3

> Il en résulte une activité de mémoire à court terme fort intense. »

En bref, ces différences lexico-syntaxiques « obligent l'interprète à une gymnastique mentale qui modifie l'ordre des informations dans la langue d'arrivée » (*ibid*., 205) (voir également Cappelli, 2015).

Dans cette étude, nous élargissons la thématique des noms composés pour nous pencher sur les séquences de noms, notion qui couvre à la fois les noms composés et les syntagmes libres (ou séquences fortuites) (voir Section 2.2).

2. Données et méthodologie

2.1. Le corpus EPTIC

Notre étude est basée sur le corpus EPTIC (*European Parliament Translation and Interpreting Corpus*), dont la collecte est coordonnée par l'Université de Bologne (Bernardini *et al*. 2016).[3] EPTIC comporte quatre parties : (1) les transcriptions des sessions plénières du Parlement européen (*st-in*), où chaque orateur (député européen ou invité) est libre de s'exprimer dans la langue de son choix (souvent, sa langue maternelle ou l'anglais), (2) les transcriptions des interprétations simultanées de ces sessions plénières (*tt-in*), (3) les comptes rendus *in extenso* des sessions plénières, tels qu'ils sont publiés sur le site du Parlement européen (*st-tr*), et (4)

[3] Nous tenons à remercier Émilie Degueldre, Sophie Steil, Fiona Thewissen, Florent Thirion et Coraline Zizi, qui, tout comme la seconde auteure de cet article, ont collaboré à la collecte du sous-corpus anglais><français d'EPTIC à la *Louvain School of Translation and Interpreting* (Université catholique de Louvain, Belgique).

les traductions de ces comptes rendus (*tt-tr*).[4] À ce jour, EPTIC comprend cinq langues (l'anglais, le français, l'italien, le polonais et le slovène).

Comme le montre le Tableau 1, lequel indique le nombre total de mots dans EPTIC (version 2016), il s'agit d'un corpus de taille modeste, que l'on pourrait qualifier de « nanocorpus » (Collard & Defrancq, 2016). Néanmoins, l'intérêt principal du corpus réside dans le fait que les textes sources de l'interprétation (les discours) et de la traduction (les comptes rendus de ces discours) sont quasi-identiques, ce qui permet d'établir une comparaison systématique entre l'interprétation simultanée et la traduction écrite. Comme l'illustrent les quelques extraits présentés dans le Tableau 2, les différences entre les discours (oraux) et les comptes rendus (écrits) se limitent fondamentalement à la suppression, à l'écrit, des disfluences (faux-départs, hésitations, répétitions, pauses), typiques de la langue orale[5].

Tableau 1 : Le corpus EPTIC (janvier 2016).

	Sous-corpus	Nombre de textes	Nombre total de mots
anglais	**st-in-en**	**64**	**21 106**
	st-tr-en	**64**	**20 091**
	tt-in-en_from-fr	65	20 836
	tt-tr-en_from-fr	65	21 641
	tt-in-en_from-it	68	16 122
	tt-tr-en_from-it	68	17 775
français	st-in-fr	65	23 943
	st-tr-fr	65	23 159
	tt-in-fr_from-en	**63**	**20 558**
	tt-tr-fr_from-en	**64**	**22 940**
	tt-in-fr_from-it	68	18 095
	tt-tr-fr_from-it	68	19 883
italien	st-in-it	68	16 586
	st-tr-it	68	16 521
	tt-in-it_from-en	64	16 977
	tt-tr-it_from-en	64	19 665
	tt-in-it_from-fr	65	17 946
	tt-tr-it_from-fr	65	20 103
	TOTAL	**1181**	**353 947**

Tableau 2 : Extraits de textes sources (*st-in* et *st-tr*).

Transcriptions des discours (*st-in*)	**Comptes rendus *in extenso* (*st-tr*)**
*Today, it's Martin Luther King Day and Martin Luther King said: "A time comes when silence becomes betrayal". Commissioner Füle, that time has long **bec- beca- ehm** arrived.*	*Today is Martin Luther King Day and it was Martin Luther King who said that a time comes when silence becomes betrayal. Commissioner Füle, that time has arrived.*
*But the **p-...** practicalities of an agreement with a State the size of Congo create **all sorts of ... all sorts of** daunting problems.*	*Yet the practicalities of an agreement with a State the size of Congo create all sorts of daunting problems.*

Dans notre étude, nous avons utilisé les sous-corpus anglais (langue de départ) > français (langue d'arrivée) d'EPTIC (en gras dans le Tableau 1), suivant une approche à la fois *parallèle*

[4] Dans cet article, nous utilisons les abréviations suivantes : st = *source text* (texte source), tt = *target text* (texte cible), in = *interpreting* (interprétation), tr = *translation* (traduction).

[5] Les disfluences peuvent être définies comme suit : « phenomena that interrupt the flow of speech and do not add propositional content to an utterance » (Fox Tree, 1995 : 709).

(textes sources anglais *vs*. textes cibles français ; flèches noires dans la Figure 1) et *intermodale* (interprétation simultanée *vs*. traduction écrite ; flèche blanche dans la même figure).

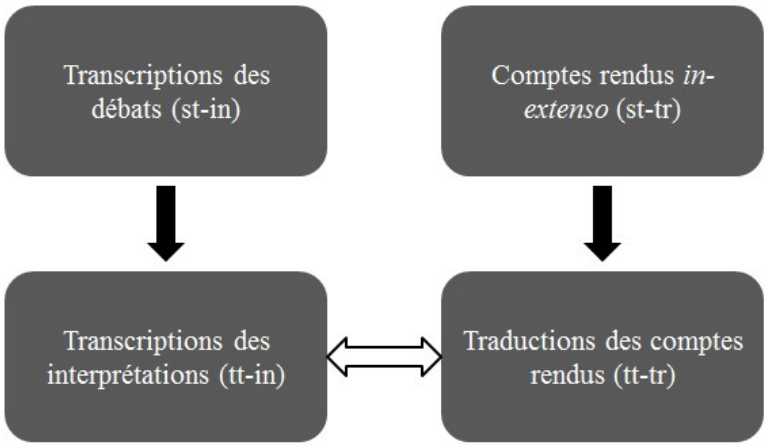

Figure 1 : Approche parallèle et intermodale.

2.2. Séquences de noms : définition et extraction

Nous avons opté pour une définition relativement large des séquences de noms (notamment au vu de la taille limitée des sous-corpus EPTIC utilisés). Toutes les séquences d'au moins deux substantifs ont été prises en compte (*sugar industry*, *labour market*, *disaster risk reduction*), y compris celles qui sont précédées d'un ou plusieurs adjectif(s) (*sustainable forest management*) et/ou suivies d'un syntagme prépositionnel (*dioxin contamination of animal feed*). De plus, nous avons analysé à la fois les unités lexicales composées (*adoption agency*, *war criminal*) et les syntagmes libres (appelés également syntagmes fortuits ; *British egg producer*, *ruling family elite*[6]). Comme le souligne Paillard (2000 : 45), « on a coutume d'opposer les unités lexicales composées (*arc-en-ciel*), totalement figées, aux syntagmes libres (*voiture en panne*), mais entre les deux apparaissent des cas intermédiaires, mémorisés comme des unités mais au statut discutable (*départs en vacances*) ». Nous n'avons pas distingué ces différents cas de figure dans nos données. Néanmoins, il faut souligner que les séquences fortuites relèvent de la syntaxe plutôt que du lexique.

Dans un souci d'exhaustivité, les séquences de noms anglaises et leurs équivalents français ont été extraits manuellement. Notre base de données finale ne reprend que les séquences de noms qui figurent à la fois dans les discours des orateurs (*st-in*) et dans les comptes rendus (*st-tr*), ce qui rend possible la comparaison directe entre l'interprétation simultanée et la traduction écrite des mêmes séquences sources.

Au total, 233 séquences de noms anglaises ont été extraites d'EPTIC. Comme on le voit dans le Tableau 3, le schéma N+N est de loin le plus fréquent (50,6 %). Les autres schémas récurrents, à savoir A+[N+N], [A+N]+N et [N+N]+N, représentent quant à eux de 11 à 14% des cas examinés. Les schémas non récurrents ont été rassemblés dans la catégorie « autres schémas » (27/233).

Tableau 3 : Séquences de noms en anglais original (*st-in* et *st-tr*).

Schémas	**Exemples**	**N**	**%**
N+N	*riot police, market barriers, pork products, enlargement process*	118	50,6

[6] Ces séquences n'apparaissent qu'une seule fois dans le corpus Europarl7 (53 millions de mots).

A+[N+N]	*global fish stocks, immediate health risk, small market share*	34	14,6
[A+N]+N	*organized crime network, intellectual property rights, low income problem*	28	12,0
[N+N]+N	*food safety chain, climate change issue, trade defence instruments*	26	11,2
Autres schémas	*micro-regional development projects of the EU, EU fish processing and canning industry, sizeable Serbian public procurement market*	27	11,6
Total		**233**	**100**

2.3. Typologie des procédés de traduction[7]

Les équivalents français des séquences sources ont été analysés en détail et classés selon la typologie des procédés de traduction de Wadensjö (1998 : 106-108) (voir Tableau 4). Largement utilisée en interprétologie, cette taxonomie a été retenue afin de refléter les spécificités des données interprétées, jusqu'ici encore peu (voire pas) exploitées en linguistique contrastive. Cette classification s'articule autour de la notion centrale de transfert de sens. Ici, nous nous sommes concentrées sur les aspects dénotatifs, dans la mesure où nous étudions des séquences au contenu référentiel/informationnel fort, dans un registre formel.

Tableau 4 : Procédés de traduction (taxonomie basée sur Wadensjö, 1998).

Procédé de traduction	**Terme anglais**	**Définition**	**Exemple tiré d'EPTIC**
Traduction proche	*Close rendition*	La séquence source est rendue par une séquence équivalente	*Member State authorities > autorités des États membres*
Traduction réduite	*Reduced rendition*	La séquence source est rendue par un équivalent moins explicite (ou, pour le dire autrement, plus implicite ou plus simple)	*Christmas Day service > cérémonie*
Traduction étendue	*Expanded rendition*	La séquence source est rendue par un équivalent plus explicite	*December contamination > contamination <u>révélée</u> en décembre*
Traduction divergente	*Divergent rendition* (voir Amato & Mack, 2011)	La séquence source est rendue par une séquence/un mot qui n'est pas équivalent. Il peut s'agir d'un glissement de sens, d'un contre-sens, etc. Dans tous les cas, le sens de la séquence source n'est pas rendu en langue d'arrivée	*young university graduate > étudiant* (vs. *jeune diplômé universitaire*)
Omission	*Zero rendition*	La séquence source n'est pas interprétée ou traduite	

La traduction réduite, la traduction étendue et l'omission s'apparentent à quelques-unes des stratégies et tactiques d'interprétation présentées par Gile (1995 : 129-135). Ainsi, à la *traduction réduite* correspondraient (i) la restitution à un niveau d'abstraction plus élevé (*Monsieur Katzantzakis a déclaré > le délégué grec a dit*) et (ii) la simplification (forme simplifiée, qui ne rend pas tous les éléments) ; à la *traduction étendue*, (iii) l'explication ou la paraphrase ; et, à l'*omission*, (iv) l'omission tactique (« omission consciente de l'information

[7] Nous utilisons le terme « procédé de traduction » pour faire référence aux deux modalités étudiées ici, à savoir l'interprétation simultanée et la traduction écrite.

véhiculée par un segment donné si l'interprète ne l'a pas comprise, s'il l'a oubliée ou s'il a du mal à la restituer en langue d'arrivée » ; *ibid*., 132).

3. Résultats et discussion

L'analyse des 233 séquences de noms révèle que les traducteurs utilisent, à une écrasante majorité, des équivalents proches en français (« traduction proche »), alors que les interprètes recourent à des procédés plus variés, tels que la simplification (« traduction réduite »), l'omission ou la modification de sens (« traduction divergente »).

3.1. La traduction écrite des séquences de noms dans EPTIC

Dans 94 % des cas (218/233), les traducteurs ont recours à des équivalents proches en français pour rendre les séquences de noms anglaises. Dans cette section, nous nous penchons sur ces cas de traduction proche.

Pour les séquences anglaises N+N (109/218), deux schémas français ressortent très clairement : N+prép+N (63 cas) et N+A (41 cas), ce qui reflète les contrastes interlinguistiques soulignés notamment par Paillard (2000) (voir 1 et 2).

> (1) [N+N] > [N+prép+N] : *adoption process* > *procédure d'adoption*, *labour market* > *marché du travail*, *minority rights* > *droits des minorités*, *development aid* > *aide au développement*

> (2) [N+N] > [N+A] : *pork products* > *produits porcins*, *food industry* > *industrie alimentaire*, *birth country* > *pays natal*, *trade agreement* > *accord commercial*

On remarque des tendances assez similaires dans le rendu des séquences N+N qui sont pré-modifiées par un adjectif (A+[N+N] ; voir 3 et 4). Il est intéressant de noter que lorsque la partie N+N de ces séquences est traduite par N+prép+N en français, la place de l'adjectif n'est pas stable : soit l'adjectif est inséré entre le premier substantif (N1) et la préposition, soit il suit la séquence N+prép+N. Il s'agit là d'une piste intéressante pour des travaux ultérieurs.

> (3) A+[N+N] > [N+prép+N]+A : *open market economy* > *économie de marché ouverte*, *national control systems* > *systèmes de contrôle nationaux*, *small market share* > *part de marché restreinte*
> A+[N+N] > N+A+prép+N : *sustainable forest management* > *gestion durable des forêts*, *global fish stocks* > *stocks mondiaux de poisson*, *excessive price volatility* > *volatilité excessive des prix*

> (4) A+[N+N] > [N+A]+A : *national forest authorities* > *autorités forestières nationales*, *erratic weather conditions* > *conditions météorologiques irrégulières*, *existing trade preferences* > *préférences commerciales existantes*

Les données traduites montrent également que le schéma français N+prép+[N+A] est souvent utilisé pour rendre les séquences anglaises [A+N]+N (*national unity government* > *gouvernement d'unité nationale*, *low income problem* > *question des bas revenus*) et [N+N]+N (*food traceability system* > *système de traçabilité alimentaire*, *food stocks issue* > *question des stocks alimentaires*, *trade defence instruments* > *instruments de défense commerciale*). De plus, on constate que les séquences anglaises [N+N]+N donnent lieu au schéma récursif N+prép+[N+prép+N] en français (*legality assurance system* > *système de vérification de la légalité*, *tuna processing facilities* > *installations de transformation du thon*, *disaster risk reduction* > *réduction des risques de catastrophes*). Ces exemples révèlent que le français fait également preuve de récursivité dans son utilisation du schéma N+prép+N (du moins le français

traduit de l'anglais). Cependant, comme le soulignent Chuquet & Paillard (1987 : 187), « le composé de taille plus raisonnable *Industrial Development Act Report* passe mal dans sa traduction directe en français : *Rapport sur la loi sur le développement industriel* ». On est donc en droit de se demander si les séquences N+prép+[N+prép+N] reprises ci-dessus sont élégantes d'un point de vue stylistique.

On peut supposer, au vu de l'ambiguïté potentielle des séquences de noms anglaises et des réorganisations syntaxiques qu'elles imposent en français, que certaines de ces séquences se sont révélées difficiles à traduire. Malheureusement, les données de corpus ne permettent pas de les identifier. Pour ce faire, il faudrait adopter une approche connaissant un vif essor en traductologie empirique, à savoir la triangulation des données (*eye-tracking*, enregistrement de frappe, raisonnement à voix haute, etc.) (cf. par exemple Alves, 2003). Cette triangulation nous permettrait d'examiner de plus près le *processus* traductionnel, en plus du *produit*, et ainsi de mettre au jour les séquences qui requièrent un effort cognitif particulier. À cet égard, les données interprétées offrent un éclairage plus riche.

3.2. L'interprétation simultanée des séquences de noms dans EPTIC

Comme l'illustre le Tableau 5, les interprètes représentés dans EPTIC ont recours à une large gamme de procédés de traduction pour rendre les séquences de noms, contrairement aux traducteurs (traduction proche dans 94 % des cas ; voir Section 3.1). Dans les données interprétées, les traductions proches représentent seulement la moitié des cas examinés (45,5 %). Elles sont suivies des cas de simplification (traductions réduites ; 24,5 %), des omissions (13,7 %) et des cas de traductions divergentes (10,7 %). Les traductions étendues (explicitation) sont moins fréquentes et ne seront pas examinées ici. Ces différences marquées entre la traduction et l'interprétation reflètent clairement les contraintes temporelles fortes et la lourde charge cognitive qui caractérisent l'interprétation simultanée. La distribution de ces procédés de traduction ne semble pas varier d'un schéma à l'autre (cf. Tableau 6), même lorsque les schémas composés de trois mots sont examinés conjointement.

Tableau 5 : Procédés de traduction (interprétation simultanée).

Procédés de traduction	**Nombre d'occurrences**	**%**
Traduction proche	106	45,5
Traduction réduite	57	24,5
Omission	32	13,7
Traduction divergente	25	10,7
Traduction étendue	13	5,6
Total	**233**	**100**

Tableau 6 : Procédés de traduction (interprétation simultanée) en fonction des schémas les plus fréquents.

Schémas	**Trad. proche**	**Trad. réduite**	**Omission**	**Trad. divergente**	**Trad. étendue**	**Total**
N+N	60	23	18	14	3	118
A+[N+N]	15	8	6	2	3	34
[A+N]+N	11	6	3	4	4	28
[N+N]+N	13	6	3	3	1	26

3.2.1. Traductions proches

Comme le montre le Tableau 5, il y a 106 cas de **traduction proche** dans les données interprétées. En fait, dans 65 de ces cas, les interprètes et les traducteurs utilisent exactement le même équivalent en français. Une analyse qualitative des séquences sources de ces équivalents identiques révèle qu'il s'agit en partie de séquences lexicalisées (noms composés, relevant de la langue générale, et termes complexes, relevant de la langue de spécialité[8]) qui sont relativement fréquentes dans les débats parlementaires européens (voir Tableau 7). Ces composés et termes complexes, sans doute mémorisés par les interprètes de manière holistique (comme des unités lexicales à part entière ; voir Wray, 2002 ; Schmid, 2010a), ne semblent donc pas poser de difficulté particulière lors de l'interprétation simultanée.[9]

Tableau 7 : Traductions proches (équivalents identiques dans *tt-in* et *tt-tr*).

Anglais (*st-in*, *st-tr*)	**Français (*tt-in*, *tt-tr*)**
Composés	
war criminal	*criminal de guerre*
road map	*feuille de route*
Christmas day	*jour de Noël*
climate change	*changement climatique*
Termes complexes (fréquents dans Europarl7 et repris dans IATE)	
candidate status	*statut de candidat*
enlargement process	*processus d'élargissement*
European Neighbourhood Policy	*Politique européenne de voisinage*
food security	*sécurité alimentaire*
free trade zone	*zone de libre-échange*
integration process	*processus d'intégration*
Member States	*États membres*
State aid	*aides d'État*

3.2.2. Traductions réduites

Dans environ un quart des cas, les interprètes ont recours à une **traduction réduite**, c'est-à-dire un équivalent simplifié (voir 5 à 7). La simplification constitue un sujet central dans les études de la « langue médiée » (*mediated language*) ou de la « langue contrainte » (*constrained language*). Par exemple, sur la base des sous-corpus anglais et italien d'EPTIC et d'indicateurs linguistiques comme la longueur moyenne des phrases et la densité lexicale, Bernardini *et al*. (2016) indiquent que la langue de l'interprétation est typiquement plus simple que la langue de la traduction, et que, de manière plus globale, la langue médiée (qu'elle soit interprétée ou traduite) est moins complexe que la langue non médiée (la langue originale). De même, dans une approche monolingue, Kruger (2017) démontre que les textes révisés sont plus simples, d'un point de vue à la fois lexical et syntaxique, que les textes non révisés. Il ressort clairement de nos résultats que cette tendance à la simplification des séquences de noms est bien plus marquée en interprétation qu'en traduction, et ce quelle que soit la configuration syntaxique des séquences. Il nous semble dès lors que les séquences de noms, qui participent de près à la densité

[8] La distinction entre noms composés (langue générale) et termes complexes (langue de spécialité) n'est pas aisée. Dans cette étude, sont considérées comme noms composés les séquences reprises dans les dictionnaires de langue générale, et comme termes les séquences reprises dans IATE. Quelques séquences relèvent des deux catégories (par exemple *labour market*).

[9] Il est également probable que certaines des séquences sources soient particulièrement saillantes du fait qu'elles ont été utilisées au préalable dans le discours de l'orateur (ou, plus généralement, lors de la session plénière) (voir à ce propos la distinction entre « saillance ontologique » et « saillance cognitive » ; Schmid, 2010b : 119-120). Toutefois, à ce stade, nous n'avons pas encore examiné ce phénomène de près.

informationnelle des discours et des textes, pourraient être envisagées comme un nouvel indicateur pour l'étude de la simplification (jusqu'à présent, en traductologie et interprétologie de corpus, la plupart des études se sont concentrées sur les paramètres proposés par Laviosa, 1998).

(5) st-in : *As I speak, the **security situation** remains precarious. Looting and violence are still reported*
tt-in : *À cette heure ehm où je vous parle ehm, **la sécuri- la sécurité** est précaire, il y a encore des piages et des violences*
tt-tr : *Au moment où je vous parle, **la situation** reste précaire **en ce qui concerne la sécurité**. Des cas de pillage et de violence continuent d'être signalés*

(6) st-in : *The so-called police of the Turkish occupation regime entered the church and forced the pli- let the priest to terminate the **Christmas Day service** and forced the Greek Cypriots enclaved attending the mass to evacuate the church*
tt-in : *L'église /ééglise/ dans un petit village, la /laa/ ... police est intervenue, est entrée ehm dans l'église et a sommé /ssommé/ le prêtre de terminer la **cérémonie** dans les 10 minutes qui suivent*
tt-tr : *La soi-disant police du régime occupant turc a pénétré dans l'église, forcé les prêtres à interrompre l'**office de Noël** et a ensuite obligé les Chypriotes Grecs enclavés assistant à la messe à évacuer l'église*

(7) st-in : *In the meantime, my group calls for maximum restraint by the security forces and for the arrest and trial of the **ancient regime's presidential guard leadership** responsible for the recent shooting, in the last few days, of innocent bystanders, in a futile attempt to de- destabilise the country*
tt-in : *Entre-temps, il nous faut ehm faire appel ehm lancer un appel modération ehm peu de violence. ehm Il faut qu'il y ait ehm un traitement juste de **ceux** qui sont coupables et il ne faut pas ehm qu'il y ait des tentatives de déstabilisation /déspabilisation/ de /dee/ ehm du pays*
tt-tr : *Simultanément, mon groupe appelle à une retenue maximale de la part des forces de sécurité et à l'arrestation et au procès du **chef de la garde présidentielle de l'ancien régime** responsable du meurtre, ces derniers jours, de spectateurs innocents, dans une tentative futile de déstabiliser le pays*

3.2.3. Omissions et traductions divergentes

On remarque également que les **omissions** et les **modifications de sens** (traductions divergentes) représentent ensemble près d'un quart des données. Comme l'illustrent les exemples 8 à 11, les omissions et les traductions divergentes impliquent des pertes d'information ou des déformations de l'information (plus ou moins importantes) pour les députés qui écoutent le discours en langue cible (voir Gile, 1995 : 16, 84). Cependant, il est utile de rappeler qu'en interprétation simultanée, « certaines omissions sont stratégiques et peuvent être considérées comme 'légitimes' » (*ibid*., 82). Nous ne nous prononcerons pas sur la légitimité de ces pertes ou déformations, l'objectif de notre étude n'étant pas d'évaluer la qualité des performances des interprètes.

(8) st-in : *We've had problems with mozzarella from Italy, **pork products** in Ireland and **cattle products** in Northern Ireland*
tt-in : *Il y a eu des problèmes avec la mozzarella en Italie, en Irlande, en Irlande du Nord*
tt-tr : *Nous avons eu des problèmes avec la mozzarella en Italie, avec les **produits porcins** en Irlande et avec les **produits carnés** en Irlande du Nord*

(9) st-in : *By the time the alarm was raised, contaminated products were all over Germany... and other parts of the EU, even in UK quiches, which were removed as a precautionary... measure from* ***supermarket shelves****. This is not the first dioxin scare*
tt-in : *Mais quand on tire la son- la sonnette d'alarme, souvent il est trop tard /tare/. On avait de la dioxine partout en Allemagne, y compris au Royaume-Uni dans les quiches. Ça n'est pas le premier scandale à la dioxine que l'on connait*
tt-tr : *Lorsque l'alarme a été donnée, les produits contaminés avaient été distribués dans toute l'Allemagne et dans d'autres pays de l'Union européenne. Ils ont même été retrouvés dans des quiches britanniques qui ont été retirées des* ***supermarchés*** *par mesure de précaution. Ce n'est pas la première alerte à la dioxine*

(10) st-in : *On the 17 December, a* ***young university graduate*** *set himself on fire out of sheer desperation (…)*
tt-in : *Un* ***étudiant*** *s'est immolé ... par désespoir (…)*
tt-tr : *Le 17 décembre 2010, un* ***jeune diplômé universitaire*** *s'est immolé par le feu par pur désespoir (…)*

(11) st-in : *The VPAs are good for the planet, good for our* ***partner countries*** *and good for the EU*
tt-in : *Ce sera positif pour les ehm* ***pays d'origine*** *et pour l'Union européenne*
tt-tr : *Les APV (accords de partenariat volontaires) sont bons pour la planète, pour nos* ***pays partenaires*** *et pour l'Union européenne*

Pour conclure, on peut dire que les données issues des transcriptions des interprétations simultanées indiquent clairement que les séquences de noms anglaises représentent de véritables pierres d'achoppement pour les interprètes professionnels qui travaillent vers le français (cf. Gile, 1995 : 107-108). La section suivante est consacrée aux disfluences, relativement fréquentes dans l'environnement direct des séquences de noms.

3.2.4. Les séquences de noms et les disfluences en interprétation simultanée

Comme le souligne Gile (1995 : 215) à propos du processus d'interprétation, « les opérations de mémoire à court terme ne sont pas directement observables, et les opérations d'énonciation ne laissent apparaître que le résultat, une fois les décisions prises. Les incidents de parcours tels que les faux départs, les hésitations, les fautes et les maladresses donnent des indices pour l'analyse du processus ». Nous avons abordé les omissions et les traductions divergentes ci-dessus. Dans cette section, nous nous penchons sur les disfluences, telles que les faux départs et les hésitations, qui se trouvent dans l'environnement direct des équivalents français des séquences de noms (nous n'analyserons pas les disfluences dans les discours sources, bien qu'elles constituent indubitablement une piste de recherche intéressante).

Les disfluences, qui reflètent dans une certaine mesure la charge cognitive à laquelle l'interprète est confronté, ont été fréquemment étudiées en interprétologie (de corpus). Plevoets & Defrancq (2016), par exemple, ont montré que les pauses remplies sont plus fréquentes en interprétation lorsque le débit de parole de l'orateur est élevé (côté *input*) et lorsque l'interprète produit un texte dense sur le plan lexical (côté *output*) (voir également Plevoets & Defrancq, 2015 ; Bendazzoli *et al*., 2011).

Dans nos données, on remarque que des disfluences apparaissent assez souvent dans l'environnement direct des équivalents des séquences de noms (33/233, soit environ un cas sur sept). On retrouve en grande majorité des pauses remplies (retranscrites dans EPTIC à l'aide de « *ehm* », quelle que soit la longueur de la pause), ainsi que quelques cas de faux départs (mots

tronqués et auto-réparations), pauses silencieuses et marqueurs de discours (*alors*, *en fait*), souvent combinés aux pauses remplies (voir 12 à 14). Ces disfluences montrent qu'au moins une partie des séquences de noms anglaises entraînent une lourde charge cognitive pour l'interprète. On observe d'ailleurs que les marques de disfluence sont majoritairement présentes lorsque l'interprète a recours à des traductions réduites ou divergentes (20 disfluences/combinaisons de disfluences sur 33 apparaissent dans de tels contextes).

(12) st-in : *The report that Ms Sârbu put in front of us underlines the issue of* ***excessive price volatility****, which is closely linked to* ***food security***
tt-in : *Le rapport présenté par madame Sarbu /Starbu/ insiste sur ehm le problème que constitue la* ***volatilité*** */volab_tilité/* ***des ehm prix*** *qui a un lien direct avec* ***ehm la sécurité alimentaire*** */alémentaire/*

(13) st-in : *Not an easy subject. We had experiences in the past of* ***commodity boards****. They have proved to not be successful*
tt-in : *Ce n'est pas facile, bon nous avons une expérience* ***en la matière ehm ce qui concerne certaines matières premières*** *et les résultats n'étaient pas au rendez-vous*

(14) st-in : *I'm sure that SAA will bring* ***concrete economic and trade benefits*** *for the country in many areas like environment, energy, transport and many others*
tt-in : *L'accord de stabilisation et d'association permettra* ***... à ... la Serbie de tirer son épingle du jeu dans le domaine économique****, dans le domaine du transport ou de l'énergie, entre autres*

La moitié des pauses remplies sont en fait insérées au milieu du segment équivalent en langue cible (*European Forest Institute > Institut ehm de la forêt européen*, *Illegal Timber Regulation > règlement ehm sur le déboisement illégal*, *market barriers > obstacles ehm qui soient mis en place sur le marché*) (un phénomène similaire a été observé chez les interprètes qui travaillent vers le néerlandais ; Plevoets & Defrancq, 2015). L'autre moitié des pauses remplies précède l'équivalent de la séquence de noms (*food security > ehm la sécurité alimentaire*, *EU fishing fleets > ehm nos flottes de pêche*).

Une analyse détaillée des séquences sources qui induisent ces disfluences permet d'identifier trois cas de figure : (i) des séquences fortuites (*British egg producer*, *December contamination*), parfois elles-mêmes composées de séquences lexicalisées (*concrete economic and trade benefits*, *excessive price volatility*), (ii) des séquences lexicalisées (typiquement reprises dans IATE), mais peu fréquentes dans les débats parlementaires européens (*commodity board*, *food processing industry*, *European Forest Institute*), et de manière plus générale, (iii) des séquences de trois mots ou plus (*EC-Pacific Interim Economic Partnership Agreement*, *high-level food and feed safety system*, *dioxin contamination of animal feed*).

Si nous disposions de données triangulées pour la traduction, nous pourrions sans doute observer des phénomènes similaires à l'écrit, comme les pauses, les corrections et révisions, l'utilisation de dictionnaires et bases de données terminologiques, etc.

4. Conclusion

Notre étude intermodale a permis de revisiter un contraste bien connu entre l'anglais et le français, à savoir les difficultés inhérentes à la traduction, de l'anglais vers le français, des séquences de noms du type N+N, A+[N+N], [A+N]+N et [N+N]+N. Alors que les données traduites d'EPTIC illustrent cette dissymétrie en faisant clairement apparaître les patrons correspondants en français (le français ayant souvent recours aux schémas N+prép+N et N+A), les données interprétées indiquent quant à elles que les séquences de noms de l'anglais

représentent de véritables « déclencheurs de difficultés » (Gile, 1995) pour les interprètes, qui dans plus de la moitié des cas, en simplifient le contenu référentiel (traductions réduites), les omettent, ou en modifient le sens (traductions divergentes). En interprétation, on observe par ailleurs des cas de disfluence dans l'environnement direct des équivalents de ces séquences sources, en particulier lorsque celles-ci contiennent trois mots pleins ou plus, sont des séquences fortuites (syntagmes libres) ou des séquences lexicalisées (composés, termes) peu fréquentes.

De manière plus générale, dans cette étude, nous avons tenté d'illustrer l'utilité des corpus intermodaux en linguistique contrastive, dans laquelle on exploite à ce jour presque exclusivement des corpus parallèles de traductions écrites (y compris les sous-titres). Nous sommes convaincues que l'utilisation de données issues d'autres types de médiation interlinguistique, comme l'interprétation simultanée, peuvent apporter un éclairage nouveau sur les contrastes interlinguistiques. En outre, l'utilisation de ce type de ressource permet de mettre en place une nouvelle forme de dialogue entre la linguistique contrastive, la traductologie et l'interprétologie, qui, selon nous, ont intérêt à mutualiser leurs ressources et leurs efforts (cf. Granger, 2003). L'ouvrage de Michel Paillard dédié à la lexicologie contrastive anglais-français (Paillard, 2000), ainsi que son autre monographie co-écrite avec Hélène Chuquet (Chuquet & Paillard, 1987), regorgent de descriptions contrastives fouillées que nous pouvons aujourd'hui ré-explorer à l'aide de nouveaux types de données empiriques, dans une perspective multidisciplinaire profondément tournée vers l'avenir. Notre dette envers Michel Paillard, et sa collègue Hélène Chuquet, est immense.

Remerciements

Nous tenons à remercier Maïté Dupont, Sylviane Granger, John Humbley et Geneviève Maubille pour leur relecture attentive et leurs commentaires avisés.

Bibliographie

ALVES, Fabio, 2003, *Triangulating Translation*, Amsterdam, John Benjamins.

AMATO, Amalia & MACK, Gabriele, 2011, « Interpreting the Oscar Night on Italian TV: An interpreters' nightmare? », *The Interpreters' Newsletter*, 16, 37-60.

ARNAUD, Pierre J.L., 2003, *Les composés* timbre-poste, Lyon, Presses universitaires de Lyon.

ARNAUD, Pierre J.L. & RENNER, Vincent, 2014, « English and French $[NN]_N$ lexical units: A categorical, morphological and semantic comparison », *Word Structure*, 7-1, 1-28.

ASTINGTON, Eric, 1983*, Equivalences: Translation difficulties and devices, French-English, English-French*, Cambridge, Cambridge University Press.

BAKER, Mona, 1995, « Corpora in translation studies: An overview and some suggestions for future research », *Target*, 7-2, 223-243.

BAKER, Mona, 1993, « Corpus linguistics and translation studies: Implications and applications », *in* M. Baker, G. Francis & E. Tognini-Bonelli (éd.), *Text and Technology: In Honour of John Sinclair*, Amsterdam, John Benjamins, 233-250.

BENDAZZOLI, Claudio, SANDRELLI, Annalisa & RUSSO, Mariachiara, 2011, « Disfluencies in simultaneous interpreting: A corpus-based analysis », *in* A. Kruger, K. Wallmach & J. Munday (éd.), *Corpus-Based Translation Studies*, *Research and Applications*, London/New York, Bloomsbury, 282-306.

BERNARDINI, Silvia, FERRARESI, Adriano & MILIČEVIĆ, Maja, 2016, « From EPIC to EPTIC - Exploring simplification in interpreting and translation from an intermodal perspective », *Target*, 28-1, 61-86.

CAPPELLI, Rita, 2015, « *Lost in interpreting or … maybe not*. Strings of nouns as a challenge for simultaneous interpreters working from Polish into Italian », *Corpus-Based Interpreting Studies: The State of the Art, First Forlì International Workshop*, Université de Bologne, Forlì, 7-8 mai 2015.

CHUQUET, Hélène & PAILLARD, Michel, 1987, *Approche linguistique des problèmes de traduction anglais-français*, Paris/Gap, Ophrys.

COLLARD, Camille & DEFRANCQ, Bart, 2016, « How to use a nanocorpus. Enriching corpora of interpreting », *Doing corpus linguistics with large and small corpora: keeping both the corpus and the linguistics components meaningful*, University of Sussex, 27 février 2016.

FOX TREE, Jean E., 1995, « The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech », *The Journal of Memory and Language*, 34, 709-738.

GARNIER, Marie, 2012, « Correcting erroneous N+N structures in the productions of French users of English », *The EuroCALL Review*, 20-1, 59-63.

GILE, Daniel, 1995, *Regards sur la recherche en interprétation de conférence*, Lille, Presses du Septentrion.

GRANGER, Sylviane, 2003, « The corpus approach: A common way forward for contrastive linguistics and translation studies? », *in* S. Granger, J. Lerot & S. Petch-Tyson (éd.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*, Amsterdam, Rodopi, 17-29.

KAJZER-WIETRZNY, Marta, 2012, *Interpreting Universals and Interpreting Style*, Thèse de doctorat, Adam Mickiewicz University, Poznan.

KRUGER, Haidee, 2017, « The effects of editorial intervention: implications for studies of the features of translated language », *in* G. De Sutter, M.-A. Lefer & I. Delaere (éd.), *Empirical Translation Studies: New methodological and theoretical traditions*, Trends in Linguistics, Studies and Monographs, Berlin, De Gruyter.

LAVIOSA, Sara, 1998, « Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose », *Meta*, 43-4, 557-570.

PAILLARD, Michel, 2011, « English-French contrasts in word-formation. Morphological patterns and stylistic effects », *Poznan Studies in Contemporary Linguistics*, 47-4, 913-927.

PAILLARD, Michel, 2000, *Lexicologie contrastive anglais-français : Formation des mots et construction du sens*, Paris/Gap, Ophrys.

PLEVOETS, Koen & DEFRANCQ, Bart, 2016, « The effect of informational load on disfluencies in interpreting: A corpus-based regression analysis », *Translation and Interpreting Studies*, 11-2, 202-224.

PLEVOETS, Koen & DEFRANCQ, Bart, 2015, « *Over-uh-load*. The occurrence of *uh(m)* between elements of compounds during interpreting », *Corpus-Based Interpreting Studies: The State of the Art, First Forlì International Workshop*, Université de Bologne, Forlì, 7-8 mai 2015.

SHLESINGER, Miriam, 1998, « Corpus-based interpreting studies as an offshoot of corpus-based translation studies », *Meta*, 43-4, 486-493.

SHLESINGER, Miriam & ORDAN, Noam, 2012, « More spoken or more translated? Exploring a known unknown of simultaneous interpreting », *Target*, 24-1, 43-60.

SCHMID, Hans-Jörg, 2010a, « Does frequency in text really instantiate entrenchment in the cognitive system? », *in* D. Glynn & K. Fischer (éd.), *Quantitative Methods in Cognitive Semantics*, *Corpus-driven Approaches*, Berlin, De Gruyter, 101-133.

SCHMID, Hans-Jörg, 2010b, « Entrenchment, salience, and basic levels », in D. Geeraerts & H. Cuyckens (éd.), *Oxford Handbook of Cognitive Linguistics*, Oxford, Oxford University Press, 117-138.

VAN HOOF, Henri, 1989, *Traduire l'anglais : Théorie et pratique*, Paris/Louvain-la-Neuve, Duculot.

WADENSJÖ, Cecilia, 1998, *Interpreting as Interaction*, London/New York, Longman.

WRAY, Alison, 2002, *Formulaic Language and the Lexicon*, Cambridge, Cambridge University Press.