

# Les constructions à verbe support dans un dictionnaire bilingue français- néerlandais

Les possibilités et applications d'une analyse de corpus bilingue pour  
la traductologie

Mémoire de fin d'études  
Master Vertalen  
Tilly van Paassen (3114325)  
Août 2010

Sous la direction de Prof. Dr. Henriette de Swart et Dr. Bert Le Bruyn  
Université d'Utrecht, Faculté de Lettres  
Département de langue et culture françaises

# Table des matières

<b>Remerciements</b>	4
<b>Chapitre 1: Introduction</b>	
1.1 L'analyse de corpus, un domaine partagé de la linguistique contrastive et de la traductologie	5
1.2. Le schéma de Holmes (1988 [1972])	6
1.3. Les constructions à verbe support	7
1.4 La délimitation des constructions à verbe support pour le français	8
1.5 L'encadrement de l'analyse des constructions à verbe support dans la traductologie	9
<b>Chapitre 2: La linguistique de corpus et la traductologie</b>	
2.1 L'avènement de la linguistique de corpus	12
2.2. La terminologie	16
2.3. Les applications de l'analyse de corpus dans des domaines divers	17
2.4 Les critères pour l'analyse de corpus dans le domaine de la traductologie	18
<b>Chapitre 3 : Méthodologie, Van Dale et les corpus</b>	
3.1. La notion de 'dictionnaire pré-corpus'	20
3.2 La recherche qualitative et quantitative	
3.2.1. L'analyse de corpus appliquée à la traductologie	22
3.2.2. Hanks et al. (2006, Hanks (2004) et Hanks&Pustejovsky (2005)	23
3.2.3. Schroten (2000, 2002)	32
3.3 Méthodologie, les corpus et Van Dale	
3.3.1. Méthodologie	36
3.3.2. Méthodes de recherche dans les corpus	36
3.4. La composition et fonction des corpus	
3.4.1. L'accessibilité et le choix des corpus	38
3.4.2 Le statut des textes dans le Corpus EU	40
3.5 Analyse de l'entrée lexicale d'un Vsupp	
3.5.1. La composition de l'entrée lexicale	40
3.5.2. L'emploi léger du verbe commettre	41
3.5.3 Les occurrences dans Van Dale F-N et N-F	43
<b>Chapitre 4 : Résultats des corpus monolingues</b>	
4.1. Les résultats issus de Frantext	48
4.1.1. Résultats rapatriés	52
4.1.2. Une comparaison entre Van Dale et le corpus Ft	52
4.2. Résultats issus de la partie française du corpus EU et de NexisLexis	52
4.3. Le corpus Ft, la partie française du corpus EU et le corpus LN : une comparaison	56
4.4. Les corpus monolingues vs. Van Dale	60
<b>Chapitre 5 : Résultats du corpus bilingue EU</b>	
5.1. Résultats de la recherche parallèle dans le corpus EU bilingue	62

5.2. Une proposition de catégorisation par Npréd	67
5.3. Une proposition de catégorisation par verbe	72
5.4. La nouvelle entrée lexicale	76
5.5. Conclusion	77

### **Annexes**

1.a l'entrée lexicale <i>commettre</i> dans Van Dale	82
1.b Données de l'option de recherche <i>binnen artikelen</i>	83
2. Occurrences, traductions et statuts des Npréd	90
3. Une catégorisation par Vsupp	94

## Remerciements

Le présent mémoire est le résultat d'une tentative de relier la traductologie à la linguistique de corpus. En raison de la littérature riche trouvée au niveau de la recherche empirique à l'aide des corpus de textes, je me suis permis d'insister sur les développements les plus récents dans ces deux filières. Bien que l'encadrement de ce sujet de recherche se soit finalement établi par l'application lexicographique, j'ai ainsi cru mieux encadrer le sujet dans la filière de la traductologie. Le développement d'une méthodologie apte à la recherche exécutée et de conception d'une nouvelle entrée lexicale d'un verbe support ont été des étapes logiques qui ont résulté de la recherche de corpus initiale.

Je suis gratifiante pour l'approbation donné par Ton Naaijkens à ce projet et pour son enthousiasme porté à l'insertion d'une approche linguistique dans le courant de la traductologie à l'Université d'Utrecht.

Mes remerciements les plus profonds sont destinés à Bert le Bruyn et Henriëtte de Swart, pour leurs idées qui ont mené à la conception finale de l'analyse de corpus bilingue avec toutes les pistes de réflexion, pour leurs questions critiques au sujet de l'application des résultats dans l'application lexicographique et pour leurs multiples suggestions pratiques qui ont toutes contribué au résultat final.

Les nombreuses discussions à trois qui ont eu lieu régulièrement m'ont toujours apporté un sentiment de privilège et presque d'une double direction de ce projet. Aussi sont-elles significatives des très bonnes conditions, du temps investi et finalement de la chance unique que j'ai eue avec ce projet encore inhabituel au sein du master Vertalen.

Je remercie Bert le Bruyn et particulier de ses efforts de la conception d'un corpus français-néerlandais. Sans lui, cette recherche ne serait pas fournie d'un corpus bilingue apte à cette recherche et nécessaire pour rendre la recherche caractérisée par son 'corpus-based approach'.

Je remercie mes membres de ma famille et mes amis de leur support, du 'room service', de leur compréhension et de leur enthousiasme.

# Chapitre 1: Introduction

## 1.1 L'analyse de corpus, un domaine partagé de la linguistique contrastive et de la traductologie

Depuis l'antiquité, la traduction se fait sujet de réflexion. L'intérêt pour l'aspect scientifique du phénomène de la traduction est né au milieu du XXe siècle. L'intérêt à la traduction comme objet de recherche a commencé dans les années 1950 avec les linguistes Vinay et Darbelnet (1958), Catford (1965), Jakobson (1959) et Mounin (1963). La traductologue D. Gile<sup>1</sup> met en avant que les années 1950 marquent aussi le début de l'interdisciplinarité de la réflexion sur la traduction (traductologie) avec d'autres champs scientifiques. La traductologie comme domaine de recherche n'a eu son avènement que dans les années 1970, avec le pionnier J. Holmes. G. Toury, qui a adapté le schéma en 1995. Leur encadrement a contribué au développement d'une branche scientifique empirique et indépendante.

Bien que le domaine de la traductologie soit déterminé indépendamment, il connaît un assez grand nombre de champs d'intérêt en commun avec d'autres domaines de recherche. La linguistique a toujours occupé une place prédominante dans le champ d'études de *translation studies* et a contribué à la formalisation du domaine de recherche de la traductologie. L'objectif de l'analyse théorique dans ce mémoire est de faire un inventaire des points communs entre la traductologie et la linguistique contrastive et d'esquisser le développement de la recherche interdisciplinaire, en prenant comme objet d'études le domaine partagé de la traductologie et de la linguistique contrastive: la linguistique de corpus. Cette branche de la linguistique empirique connaît un spectre divers d'applications dans la linguistique contrastive et la traductologie. La linguistique contrastive constitue une comparaison systématique de au moins deux langues, ayant l'objectif de décrire leurs similarités et leurs différences. On remarque qu'à présent, la comparaison de langues sert aux branches appliquées ainsi que théorique, sur le niveau d'une seule langue ainsi que la langue en général. (Johansson et Hofland 1994 : 25)<sup>2</sup> L'analyse d'un corpus dans la linguistique contrastive constitue une recherche après un phénomène langagier dans les textes insérées et sélectionnées selon les objectifs spécifiques. En général, la source exacte des données n'est pas d'importance, à condition que les textes reflètent l'emploi langagier originel. Par conséquent, on essaie de réduire les influences éventuelles de traductions individuelles par l'utilisation de grands corpus afin d'obtenir des données d'équivalence crédibles.

Quand l'objectif est d'inventorier les effets de traduction, la sélection de textes pour le corpus est plus importante. Dans le cadre de la traductologie, la recherche après le processus de traduction et le produit (*product of translation*) est le plus souvent exécutée. L'objectif est de généraliser les résultats au niveau de simplification, explicitation ou normalisation d'éléments textuels. Ainsi, l'analyse de corpus mène à une recherche après les traits universelles de la traduction, aussi appelées *tendances universelles*. Dans les années 1990, l'analyse de corpus est également incorporée dans la traductologie et dans les deux décennies qui ont suivi, la méthodologie de recherche a subi d'adaptations importantes afin de mieux servir la recherche empirique dans le cadre de la traductologie. Après une analyse théorique de l'utilisation de l'analyse de corpus, je veux inventorier aussi bien la méthodologie de l'analyse de corpus en fonction de l'exécution d'une analyse de corpus que l'application des résultats aux fins traductologiques. Je voudrais montrer que l'analyse de corpus peut être utilisée dans plusieurs domaines de la traductologie. La première question centrale est la suivante :

**De quelle manière, l'analyse de corpus peut contribuer à la recherche empirique dans le domaine de la traductologie ?**

---

<sup>1</sup> Gile D., 'Présentation traductologie' sur <http://www.cirinandgile.com/ESIT%20home.htm> (2009)

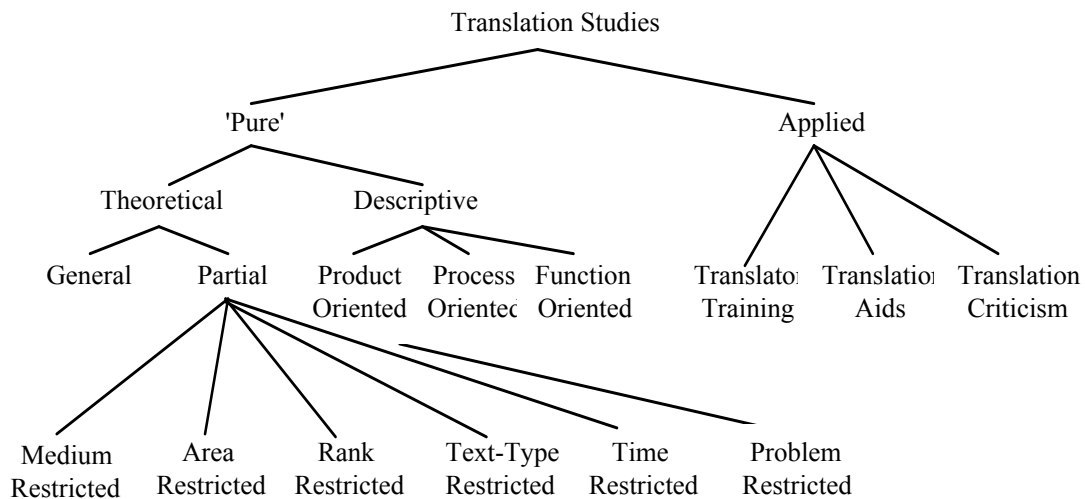
<sup>2</sup> Johansson, S, and Hofland K., Towards an English-Norwegian parallel corpus. In: U. Fries, Editor, *Creating and Using English Language Corpora*, Rodopi, Amsterdam (1994), pp. 25–37.

Afin de mieux classifier l'analyse de corpus que je vais exécuter dans les domaines de linguistique et de traductologie, j'analyserai d'abord l'insertion de l'analyse de corpus dans la traductologie.

## 1.2. Le schéma de Holmes (1988 [1972])

A l'intérieur de la traductologie, le schéma de Holmes fournit une classification des domaines de la traductologie. Voici le schéma:<sup>3</sup>

Figure 1 : La taxonomie de *translation studies*



La branche théorique constitue avec la traduction descriptive l'ensemble d'études de traductologie pure et descriptive. Holmes la définit comme la branche qui vise une conception de traduction ayant l'objectif d'arriver à une définition d'une théorie générale de la traduction. Avec l'introduction du schéma répartitif de Holmes, la branche DTS, la traductologie théorique descriptive a vu jour. La DTS est la branche empirique dans laquelle le processus, le produit (*product of translation*) et la fonction de traduction et de textes traduits trouvent sa place. La dernière catégorie traite la réception des textes dans leur environnement de langue et culture cible. La branche appliquée rassemble la didactique, les outils et la politique de traduction. Comme il est difficile d'attribuer la notion de traductologie, la notion de la traduction appliquée est plus souvent utilisée.

La branche DTS est critiqué et adapté en 1995 par G. Toury remet en cause la valeur scientifique d'analyses dites 'individuelles'.<sup>4</sup> En outre, il privilège la branche de DTS comme l'unique branche scientifique de la traductologie. Dans son adaptation, il insiste sur la conception de la traductologie appliquée comme partie 'extensionnelle', c.à.d. constituant les 'extensions' de DTS, est élabore donc une relation indirecte entre DTS et la branche appliquée. DTS essaye a mieux indiquer des régularités probabilistes dans le 'comportement traductionnelle'. Bien que Holmes catégorise l'analyse contrastive dans la branche théorique, Toury met l'analyse contrastive à l'intérieur de DTS. Finalement, il met l'accent dans ses adaptations sur l'objectif de formuler une théorie générale qui appuie sur les théories partielles. La recherche empirique dans la DTS est censé de toujours servir une théorie générale sur la traduction (Laviosa : 17).

<sup>3</sup> Holmes, J. 'the name and nature of translation studies', *Translated! Papers on literary translation studies and translation studies*. Amsterdam: Rodopi, (1988 [1972]) Les notions de français ont été empruntés de D. Gile: 'initiation à la traductologie', [www.cirinandgile.com/2Initttraductol.ppt](http://www.cirinandgile.com/2Initttraductol.ppt), 2009.

<sup>4</sup> Toury, G. *Descriptive translation studies and beyond*. Amsterdam/Philadelphia: John Benjamins, 1995

La deuxième partie de cette section est consacré à l'élaboration de trois catégorisations possibles pour l'analyse de corpus.

Premièrement, l'analyse de corpus aux fins purement traductologiques, après les traits universels de la traduction, peut être défini comme partie de la traductologie descriptive. (Oholan : 24)<sup>5</sup> Laviosa met en avant un argument contre cette catégorisation : dans la linguistique du corpus, les théories partielles sont considérées comme un reflet elle-même de l'hétérogénéité et la dynamique de l'emploi langagier dans les textes traduits. Par conséquent, l'analyse de corpus a moins tendance à formuler une théorie générale. (Laviosa : 17)

Deuxièmement, l'analyse de corpus pourrait se définir également comme une méthode de recherche dans les sous-théories de la traductologie théorique. Une analyse de corpus bilingue est déjà délimitée par le domaine de textes (selon le corpus utilisé), par la période d'édition des textes ainsi qu'elle peut être restreint à un phénomène linguistique en particulier. Le domaine de *problem restricted theoretical translation studies* est la catégorisation la plus spécifique possible.

Je propose ici un élargissement des applications pour l'analyse de corpus aux fins lexicographiques. Une troisième possibilité de catégorisation de l'analyse de corpus pourrait ainsi être dans la branche de la traductologie appliquée. Une analyse de corpus peut ainsi trouver sa place sous le développement des outils à la traduction. Cette catégorisation est soutenue par les suggestions de Holmes (1988) : il plaide dans sa taxonomie que les moyens lexicographiques doivent être adaptés de manière plus adéquate à la traduction. Il met en avant que les grammaires sont plutôt visées sur l'apprenant, donc à l'acquisition de la langue, et qu'avec la lexicologie et la linguistique contrastive, les moyens de support doivent être adaptés aux besoins du traducteur. Avec cette troisième possibilité, les résultats de recherche peuvent être appliqués afin d'améliorer les stratégies et également les outils de traduction.

Il est donc la possibilité d'appliquer l'analyse de corpus dans des domaines divers de la traductologie. Dans le chapitre 1.4, j'élaborai la définition finale la de catégorisation de ma recherche dans la taxonomie de Holmes.

### 1.3. Les constructions à verbe support

Dans mon analyse de corpus j'ai choisi de sélectionner un phénomène linguistique qui surgit fréquemment dans la traduction. Tout comme dans le domaine de *natural language processing*, les *multiword expressions* posent problèmes au traitement automatique des langues<sup>6</sup>, les collocations aux verbes supports posent des problèmes au traducteur aux niveaux sémantique et compositionnelle. Les verbes supports constitueront le sujet de recherche pour mon analyse de corpus. Les verbes (*light verbs* en anglais) sont marqués par un emploi et signification spécifique dans les constructions avec un nom prédicatif, des constructions à verbe support (*support verb constructions* ou SVCs). Les verbes comme *donner* et *faire* ne prédominent pas la collocation : les verbes sont déficients au niveau sémantique :

1. Elle a fait une présentation.
2. Jean prend une décision.

Si dans la collocation *faire une présentation*, l'objet direct *présentation* est supprimé, le verbe ne convoque plus de signification de '*présenter*'. De la même manière, la collocation du verbe support avec l'objet direct dans l'exemple 2 peut être remplacée par un seul verbe entier (*main verb*), *décider*. Pour ce type de collocations, la signification de base des verbes ne suffit pas. C'est à partir des noms prédicatifs comme *décision* que ces Vsupp prennent un sens en collocation. La construction Vsupp ne porte explicitement pas de signification idiomatique parce que les constructions Vsupp se caractérisent par leur flexibilité syntaxique, et leur productivité sémantique spécifique, ce qui sera traité dans le chapitre 1.4.

---

<sup>5</sup> Oholan, M. *Introducing corpora in translation studies*, Oxfordshire: Routledge, 2004.

<sup>6</sup> Sag et al. 'Multiword expressions: a pain in the neck for NLP', 2002

Bien que les verbes support se combinent également avec les compléments, dans ce mémoire les constructions à verbe support se définissent par une construction d'un verbe support (abrégé par Vsupp) avec un objet direct (abrégé par Npréd). Le Npréd déterminera le contenu sémantique du Vsupp et par conséquent, sa signification est spécifique au Npréd. J'emploie dans ce mémoire également les notions construction à verbe support (abrégé par construction Vsupp) et verbe léger (opposé à verbe entier).

Quand on passe d'une analyse monolingue à une analyse bilingue, il s'avère que la spécificité est également un élément spécifique à une langue :

3. a. She gives a lecture.
- b. Sie hält eine Volesung.

Une comparaison du nom *présentation /lecture/Vorlesung* relève qu'en anglais et en allemand, ce nom ne se combine pas avec un équivalent de 'faire' : *lecture* prend le verbe *to give* au lieu de *to make* et *Vorlesung* prend le Vsupp *halten* au lieu de *machen*. Il se peut donc que la traduction de la construction Vsupp dans une autre langue n'a qu'une traduction partielle.

Dans la linguistique, les recherches se sont consacrés à développer des modèles syntaxiques et sémantiques afin de mieux définir les constructions à verbe support. Je veux élaborer la recherche dans un contexte bilingue et analyser l'occurrence des constructions Vsupp en français et en néerlandais. Par conséquent, notre recherche tirerait un grand profit d'une méthode de recherche de la linguistique contrastive, au lieu d'utiliser une méthode de recherche traductologique. Les corpus utilisées dans l'analyse contrastive prétendent être un reflet de l'emploi langagier et par conséquent, les effets de traduction y sont effacés afin d'obtenir des données crédibles. La recherche dans ce mémoire ne vise pas de représentation sémantique ou syntaxique, mais a l'objectif d'une application lexicographique des données tirées d'une analyse de corpus. La recherche de corpus permettra une analyse linguistique qui pourrait contribuer à l'amélioration de l'entrée lexicale des Vsupp. L'objectif final est concevoir une représentation lexicographique où les occurrences seront fournis d'informations sémantiques en syntaxiques. Le deuxième volet de ce mémoire est constitué par la question:

**Comment les constructions à verbe support peuvent-elles être mieux affichées dans un dictionnaire bilingue ?**

Avant d'encadrer la recherche de corpus que j'ai défini dans cette section, il est d'abord nécessaire de mieux délimiter ce qui est désigné par la notion de construction Vsupp.

#### **1.4. La délimitation des constructions à verbe support pour le français**

Suite à la définition de constructions à verbe support par Mel'cuk (1996, 2007)<sup>7</sup>, S. Langer (2004) fournit un ensemble de tests pour les constructions à verbe support dans l'article '*A linguistic test battery for support verb constructions*',<sup>8</sup> La délimitation du phénomène des constructions Vsupp pose plusieurs problèmes en termes de la description du degré de la légèreté sémantique des Vsupp en question. Dans l'exemple 5, le verbe est plus léger que dans l'exemple 4 :

4. porter une robe
5. porter un jugement

---

<sup>7</sup> Mel'cuk, I. 'Verbes supports sans peine', *Linguisticae Investigationis*, 27 :2, 2004, p. 203-228.

Mel'cuk, I. *Lexical Functions : A tool for the description of lexical relations in a lexicon*. Dans Wanner, L.(ed) *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1996.

<sup>8</sup> Langer, S. 'A linguistic test battery for delimiting support verb constructions', *Linguistica Investigationis*, 27 :2 John Benjamins Publishing Company (2004) p. 171-184



Formellement, Langer adopte les caractéristiques du grammaire-lexique de Mel'cuk (1996) et la sémantique des cadres (*frame semantics*)<sup>9</sup>. Dans ces programmes, beaucoup d'attention est porté à la structure argumentale du nom : tandis que le verbe réalise ses arguments syntaxiques, le nom doit réaliser ses 'actants sémantiques'. Ainsi, les deux constituants dans une construction Vsupp sont complémentaires. Langer détermine les constructions Vsupp comme 'semi-compositionnelles'. (Langer: 171-172)

Ma recherche après la distribution lexicale des constructions Vsupp nécessite une définition précise de la notion de construction Vsupp afin d'exclure les collocations qui ne remplissent pas les conditions posées dans le chapitre 1.3. La délimitation de Langer permet de supprimer les emplois idiomatiques et l'emploi du verbe comme verbe principal.<sup>10</sup> Dans la première partie, l'ensemble de tests est analysé. Ensuite, les tests sont vérifiés sur leur pertinence à la recherche bilingue dans la deuxième partie de cette section.

L'objectif général de Langer est de distinguer les constructions Vsupp prototypiques des 'cas marginaux' afin de décrire les propriétés qui valent pour l'ensemble de constructions Vsupp. (Langer : 174). Bien qu'il utilise des exemples dans trois langues, j'ai vérifié la pertinence des tests pour notre recherche du français avec le Vsupp *prendre*. Il s'agit de deux types de tests les tests de référentialité des Npréd, et les tests pour contrôler la fonction sémantique réduite du Vsupp et la sensibilité à la collocation en construction Vsupp. Pour vérifier les tests, trois phrases sont mis en avant avec le verbe *prendre* qui sont marquées par un stade graduel de légèreté.

6. a. Jean prend le sac à dos dans la main.
- b. Jean prend un café.
- c. Jean prend une photo.

Dans le sens de *prendre* illustré dans l'exemple a, le Npréd de l'objet direct se remplace facilement par d'autres noms d'objets, avec comme seule condition pragmatique que les N dans cette catégorie soient portable dans la main. Ce type est entièrement productif au niveau sémantique et n'est pas délimité par l'emploi particulier en collocation. Dans l'exemple b, *café* peut être remplacé par une catégorie de Npréd qui est délimitée par la signification 'consommer quelque chose' : *un verre, un coca, une pression* font partie de cette variation ainsi que *un croque monsieur* ou *ses médicaments*. Dans l'exemple c, *prendre une photo*, le Npréd *photo* est le seul type de Npréd qui est équivalent à 'photographier'. Les catégories b et c sont envisagées en particulier dans notre recherche après les constructions Vsupp.

Les tests de référentialité des Npréd

Le test de pronominalisation s'effectue de deux manières : externe à la construction, par une phrase avec *qu* ou par la pronominalisation interne. Les constructions Vsupp selon Langer, permettent la pronominalisation.

7. a. Il fait une présentation que personne ne comprend. (pronominalisation externe)
- b. Il regrettait la décision qu'il avait prise. (pronominalisation interne) (Langer : 176)

Les idiomes sont ainsi exclus puisque dans les constructions idiomatiques, le Npréd dans l'exemple d. n'est pas référentiel.

8. a. Jean prend le sac à dos qui est lourd dans la main.
- b. Jean regrettait le café qu'il avait pris.
- c. Jean prend une photo qui lui servira d'aide mémoire plus tard.

La substitution par interrogatifs afin de vérifier la référentialité s'avère moins applicable aux constructions Vsupp:

9. Qu'est-ce que Jean prenait?

---

9 Fillmore, C., Johnson, C.R., Petruck, M.. Background to *framenet*. International journal of lexicology, 16(3), 2003.

<sup>10</sup> Afin d'opposer *support verb meaning* à *full verb meaning*, la notion de emploi principal n'est pas tout à fait satisfaisante mais une notion choisie par défaut.

- a. Le sac.
- b. Un café.
- c. ?Une photo

Dans l'exemple c, je doute de la réponse quand il s'agit de l'action de photographier. Par contre, s'il s'agit de l'action de prendre quelque chose dans la main, une photo est une bonne réponse.

La flexibilité syntaxique du Npréd est dénotée par Langer comme 'la variabilité du Npréd' (176). Comme le Npréd dans les constructions à verbe support est syntaxiquement plus flexible que dans les constructions idiomatiques, les constructions Vsupp ne sont pas retraintes au nombre et au déterminant, et se produisent avec différents types de négation et avec des pronoms possessifs.

- 10. a. Jean prend son sac à dos dans la main.
- b. Jean ne prend jamais de café.
- c. Jean prenait un dizaine de photos.

Selon le test de variabilité, le Npréd permet également la passivation :

- 11. a. Le sac est pris dans la main
- b. Dans un café français, le café est pris debout.
- c. Les photos ont été prises dans le jardin.

Ce test a un caractère syntaxique au lieu de sémantique, et par conséquent, je ne l'utiliserai pas pour la sélection des constructions Vsupp.

Les tests pour contrôler la fonction sémantique réduite du Vsupp et la sensibilité à la collocation en construction Vsupp.

En français, le verbe en construction Vsupp ni la construction Vsupp ne peut pas être nominalisé:

- 12. a. \*La prise du sac.
- b. \*La prise d'un café.
- c. \*La prise d'une photo.

Deuxièmement, le test Zeugma consiste en l'insertion d'une coordination à l'intérieur de la construction Vsupp et une 'construction non-Vsupp'. Selon Langer, les constructions Vsupp ne permettraient pas la coordination :

- 13. a. Jean prend son sac à dos dans la main et un parapluie.
- b. \*Jean prend un café et un parapluie.
- c. \*Jean prend une photo et un parapluie.
- d. Jan prend un café et une bière.

Ce test n'est pas univoque, comme montre l'exemple d : *prendre un café* et *prendre une bière* sont tous les deux une construction Vsupp. Apparemment, si une des constructions avec prendre n'est pas une construction Vsupp, la coordination n'est pas compatible : *prendre (Vsupp) une photo* et *prendre(V entier) un parapluie* ne se combinent pas.

Langer met en avant que les constructions à verbe support peuvent être substituées par un seul verbe principal. Il propose d'abord le test de compositionnalité en remplaçant la construction Vsupp par un seul verbe principal. En deuxième lieu, la substitution du Vsupp seul constitue également un test valable :

- 14. a. Jean prend/met son sac à dos dans la main (substitution du verbe seulement)
- b. Jean prend/boit un café
- c. Jean prend une photo ->Jean prend [quelque chose] en photo -> Jean photographie [quelque chose].

Les exemples a et b ne permettent pas cette substitution. Langer fait observer que les constructions Vsupp portent souvent un Npréd déverbal qui permet la substitution. Langer : 180)

L'omission du verbe support est également proposée comme un test valable, emprunté à J. Giry-Schneider (1987). Elle constate que la construction Vsupp ne perd pas d'information quand le Vsupp est supprimé. Cette caractéristique s'applique bien au verbe *faire* :

15. Jean était content qu'il ait fait une bonne présentation.

La substitution d'une phrase subordonnée avec le Npréd est plus difficile avec le verbe *prendre* :

16. a. ?Anne est contente du sac de Jean dans sa main

b. \*Anne est contente du café de Jean.

c. ?Anne est contente de la photo par Jean.

Les Npréd seuls ne réfèrent pas (a et b) ou peu (exemple c) à l'activité de *porter un sac, boire un café* ou de *photographier*. Toutefois, ce test pourrait bien donner de bons résultats avec le verbe *commettre* dans le paragraphe 3.5.2.

Langer adopte un deuxième critère de sélection de Giry-Schneider (1987) qui consiste dans l'interchangeabilité d'adverbe et l'épithète parce qu'ils prennent la même valeur sémantique dans une construction Vsupp. Il s'avère qu'avec le paire fréquent-fréquemment, seulement dans l'exemple a, on remarque une différence :

17. Marie fait fréquemment des faux pas / Marie fait des faux pas fréquents (Langer : 185)

18. a. Jean prend fréquemment le sac à dos dans la main./

\*Jean prend le sac à dos fréquent à la main

b. Jean prend fréquemment un café. / \*Jean prend le café fréquent/ ?Jean prend les cafés fréquents.

c. Jean prend fréquemment une photo / ?Jean prend des photos fréquentes.

Il se peut que le verbe *faire* se prête mieux à une action qui se répète.

Dans l'analyse des critères de sélection de Langer, il s'avère que l'application des tests dépend souvent de la langue choisie. Pour le français, les tests de la pronominalisation, et le test de substitution du Vsupp par un autre Vsupp sont pertinents pour notre recherche. Les tests de substitution de l'Npréd par des interrogatifs, la nominalisation de la construction Vsupp, la substitution de la construction Vsupp par le verbe principal, l'omission du verbe support et le test sur le contenu sémantique des constructions Vsupp posent plus de problèmes aux constructions Vsupp avec le verbe *prendre*. Le test Zeugma n'était pas univoque dans les résultats. Contraire au test de Langer, notre intérêt réside non seulement dans le français, mais aussi dans la traduction des constructions Vsupp en néerlandais. Pour la sélection de constructions Vsupp dans le chapitre 4, il faudra distinguer les constructions Vsupp qui ne sont pas 'légères' mais qui poseront également des problèmes pour la traduction, notamment aux utilisateurs allophones. Il se peut qu'il faut dévier des critères posés dans cette section et par conséquent analyser la pertinence de constructions comme l'exemple a, 'Jean prend le sac à dos dans la main', dans une application bilingue.

### 1.5 L'encadrement de notre analyse des constructions à verbe support dans la traductologie

L'insertion de notre analyse contrastive dans l'encadrement théorique constitue la dernière étape de définition. Il s'avère des analyses de Holmes, Toury et Laviosa (voir le chapitre 1.2), qu'une analyse contrastive est de nature différente qu'une analyse de corpus des traits universels de traduction. On peut constater du chapitre 1.3. qu'une analyse de constructions Vsupp dans plusieurs langues demande une méthode de recherche d'analyse contrastive, ce qui n'est pas compatible avec l'encadrement par *problem restricted theoretical translation studies*. Ma recherche visée partage également plus avec la linguistique contrastive et par conséquent, avec la traduction théorique. Pourtant, notre recherche ne peut pas être classifiée dans la filière de la traductologie descriptive en raison de ses objectifs : elle ne vise pas les traits universels de la traduction, ni après les textes traduits à part mais vise un phénomène linguistique pour lequel il est important de réduire les influences des traductions dans les textes des corpus. Par conséquent, j'ai une préférence pour un corpus de données de textes qui n'ont pas de statut de langue-cible ou de langue-source. Un tel corpus est fourni par le corpus Europarl, où les actes de conférence de la commission européenne ont été intégrés dans les 11 langues européennes. Il y a donc la possibilité de concevoir un corpus bilingue où les textes ont le même statut 'neutre' de *versions*. On retournera au statut des textes dans le chapitre 3.4.

La problématique des constructions Vsupp ne peut pas être catégorisée non plus sous la traductologie théorique. L'analyse de corpus visé ne se concentre pas sur les caractéristiques textuelles des textes traduits mais sur les caractéristiques spécifiques aux langues. Bien que la sous-théorie autour un problème (*problem restricted theoretical translation studies*) semble une bonne catégorisation, l'objectif n'est pas de soutenir une théorie générale de traduction mais une recherche qui vise l'application lexicographique des résultats. Ce changement est spécifiquement lié à la composition du corpus, comprenant des textes 'neutres', et la méthode de recherche de l'analyse contrastive. En particulier, les résultats seront insérés dans une discussion autour le traitement de constructions Vsupp dans un dictionnaire bilingue.

Les types de corpus utilisés dans la traductologie sont analysés dans le chapitre 2.1. Ensuite, les points communs avec d'autres domaines de recherches sont définis (2.2.), les constructions à verbe support seront délimitées de point de vue linguistique (2.3) et méthodologique (2.4). Dans le chapitre 3, la problématique des constructions à verbe support de point de vue lexicographique est inventoriée. Nous déterminons une méthodologie apte à notre recherche (3.3), les fonctions des corpus (3.4.) et nous aborderons l'analyse du dictionnaire Van Dale qui servira comme point de départ dans l'analyse lexicographique (3.5). Dans le chapitre 4, les données issues des corpus monolingues seront mises à une analyse quantitative. Avec la présentation des résultats du corpus bilingue dans le chapitre 5, je nous présenterons les résultats qualitatifs ainsi qu'une application lexicographique des résultats sous forme d'une conception d'une nouvelle entrée lexicale pour un dictionnaire bilingue français-néerlandais.

## Chapitre 2: La linguistique de corpus et la traductologie

Suite à la présentation, la définition et la délimitation du sujet de recherche, l'analyse de corpus se concentre sur le phénomène des verbes supports. Le deuxième chapitre fournira la terminologie nécessaire pour le développement d'une méthodologie spécifique à notre recherche. Nous insisterons sur les origines (2.1), la terminologie et la méthodologie (2.2), de la linguistique de corpus et l'application de l'analyse de corpus dans des domaines divers (2.3) et les critères de recherche (2.4) afin de faciliter la concrétisation de la délimitation des corpus dans le chapitre 3.

### 2.1. L'avènement de la linguistique de corpus

L'analyse de corpus prend ses origines dans de différentes branches en recherche et sous de différents noms. Je prendrai en compte aussi bien le développement de l'analyse contrastive que la conception des premiers corpus.

L'analyse contrastive est née dans des années 1960 comme un sous-domaine de la linguistique générale. La branche de l'analyse contrastive (plus tard devenu la linguistique contrastive) était considérée comme une branche appliquée de la linguistique ayant l'objectif de développer des outils de l'apprentissage de langue étrangère. Les recherches servaient comme support à la création de grammaires contrastives, souvent dans un domaine spécifique (p.e. la phonétique) et dans une ou plusieurs langues spécifiques. Dans le domaine de recherche après l'acquisition de langue secondaire, l'intérêt aux facteurs internes dans le système langagier, diminuait quand on se réalisait que la problématique du FLE ne se limitait pas à la linguistique. Le premier corpus de textes conçu aux objectifs linguistiques, le Brown Corpus, a été créé par Francis et Kucera et a vu le jour en 1967. Le corpus, qui constituait de textes en anglais américain, a servi de modèle pour le Lancaster-Oslo-Bergen Corpus, et était créé de samples écrits de l'anglais britannique, compilé par Leech, Johansson et Hofland en 1978. (Laviosa 2002 :5) Au lieu de textes entiers, les corpus peuvent également insérer des samples : des fragments de textes qui ont été sélectionnés selon des critères fixes ou de manière arbitraire. (Laviosa : 33) Le Brown Corpus contient ainsi des samples de 2000 mots environ.

Dans les années 1970, il existe deux types de corpus, des corpus monolingues qui servaient comme '*a collection of texts assumed to be representative of a given language, dialect or other subset of a language, to be used for linguistic analysis.*' (Sinclair: 1991 cité dans Laviosa : 6). Il s'agit de l'emploi original de la langue, en raison de laquelle les textes traduits expulsés de grands corpus comme le BNC dans les années 1990. Les textes traduits étaient en général susceptibles de ne pas représenter l'emploi langagier original dans le contexte de la langue-cible. (Oholan 2004 :13) Au début, les corpus bilingues étaient plutôt constitués de textes originaux de plusieurs langues, mais qui se ressemblent par leurs caractéristiques comme période d'édition, registre, etc.

En 1972, J. Holmes lance son taxonomie de la traductologie. L'intérêt à la linguistique contrastive et l'analyse de corpus par la nouvelle branche scientifique réside dans leur nature empirique. Avec la lexicologie, la traductologie prétend partager ses intérêts au développement d'outils d'apprentissage. Holmes décrit la tendance que la traduction est en grandes lignes considérée comme semblable à la recherche dans la linguistique comparée et contrastive. Par conséquent, la traductologie pourrait se servir de la même terminologie (Denken over vertalen : 153).<sup>11</sup> Toutefois, les traductions étaient considérées comme le reflet des possibilités dans une autre langue à l'intérieur de la linguistique contrastive tandis que les deux branches principales de la recherche empirique pour la traductologie

---

<sup>11</sup> Holmes, J.S. 'Wat is vertaalwetenschap?' Nederlandse vertalingen en uitbreiding in: Denken over vertalen, Amsterdam: Van Tilt, 2004. Origineel: 'The name and nature of translation studies' in Translated! Papers on translation and translation studies. Amsterdam: Rodopi, 1988 [1972].

pure, l'intérêt pour le processus et l'activité de la traduction, ne recevaient pas d'attention. (Laviosa : 6)

Le domaine partagé qui était prévu par Holmes se réalisait dans les années 1980, avec l'introduction des premiers corpus bilingues contenant des textes traduits. Selon Holmes, il s'agissait d'une branche. Comme il s'est avéré des chapitres 1.1. et 1.2., la traductologie est présentée comme une branche scientifique indépendante. Toutefois, cette 'deuxième génération' de corpus, conçu dans les années 1980 et 1990, est encore marquée par la linguistique de corpus faisant partie de la linguistique générale. (Laviosa : 6) Dans les années 1980, avec l'intérêt croissant pour la communication intralingue, l'émergence et le succès de la linguistique de corpus a fait revivre le domaine de la linguistique contrastive avec la conception de corpus monolingues à plus grande échelle.

Le domaine de l'analyse de corpus monolingue avait été élargi par l'aspect multilingue : la création du premier corpus parallèle, le corpus Hansard<sup>12</sup>, a marqué le début d'une nouvelle étape de la recherche. Les **corpus parallèles** contiennent les textes sources et leurs traductions dans une ou plusieurs langues, et sont donc bilingues ou multilingues. Dans le domaine exclusivement de la traductologie, il s'agit souvent d'une comparaison entre plusieurs traductions de la même texte-source, ou d'une analyse de plusieurs textes traduits par le même auteur. Le corpus Hansard par contre, est constitué des actes de conférence du parlement Canadien en français et en anglais.

Pourtant, l'analyse envisagée par Holmes dans la branche DTS, la recherche après le produit, le processus et la fonction de la traduction, n'est abordé que dans les années 1990. En 1980 déjà, G. Toury plaide en faveur d'une orientation vers les textes et langues cibles et proposait une analyse de régularités probabilistes de traduction<sup>13</sup>. Finalement, les premières recherches sur les **corpus de textes traduits** ont été exécutées par Gellerstam (1986)<sup>14</sup> et Lindquist (1989)<sup>15</sup>. Gellerstam a étudié le 'translationese', et analyse comment les influences de la traduction sur l'emploi de la langue dans les textes-cibles peuvent interférer avec la langue-source. Par conséquent, l'analyse de corpus est intégrée dans le domaine de la traductologie depuis les années 1990. Une autre recherche prototypique dans la branche DTS est la recherche après les traits généraux de la traduction, par exemple la recherche de Johansson (1998c) sur la distribution lexicale des notions 'love' et 'hate' dans les traductions en norvégien et en anglais.<sup>16</sup>

En 1993, M. Baker (1993, 1995, 1996) donne le coup d'envoi pour le développement d'une méthodologie apte à la traductologie.<sup>17</sup> Dans l'article '*Corpora in translation studies: an overview and suggestions for future research*' (1995), elle plaide en faveur d'une méthodologie à part de la linguistique contrastive aux fins traductologiques. Baker redéfinit le corpus comme

*'a collection of texts held in machine-readable form and capable of being analyzed automatically or semi-automatically in a variety of ways; [...] a corpus may include a large number of texts from a variety of sources, by many writers and speakers and on a multitude of topics. [...] it is put together for a particular purpose and according to explicit design criteria in order to ensure that it is representative of the given area or sample of language it aims to account for.'* (Baker: 225)

---

<sup>12</sup> Hansard Corpus: <http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC95T20>

<sup>13</sup> Toury, G. *Descriptive translation studies and beyond*. Amsterdam/Philadelphia: John Benjamins, 1995

<sup>14</sup> Gellerstam, Martin. "translationese in Swedish novels translated from English", in Lars Wollin & Hans Lindquist, *Translation Studies in Scandinavia*, CWK Gleerup, Lund, 1986, pp. 88-95.

<sup>15</sup> Lindquist, H. *English adverbials in translation: A corpus study of Swedish Renderings*. Lund studies in English 80. Lund: Lund University Press, 1989.

<sup>16</sup> Johansson, S. *Loving and hating in English and Norwegian: A corpus-based contrastive study*. Dans: Albrechtsen et al. (eds) *Perspectives on foreign and second language pedagogy. Essays presented to Kirsten Haastруп on the occasion of her sixtieth birthday*, 93-103. Odense: Odense University Press.

<sup>17</sup> Baker, M. 'Corpora in translation studies: an overview and suggestions for future research', *Target* 7, 1995, John Benjamins Publishing Company: Amsterdam/Philadelphia p. 223-243

En écrivant son article, Baker a encore peu d'exemples de recherches de corpus dans la traductologie à sa disposition : ainsi, à part de sa définition de base de trois types de corpus, elle donne des suggestions pour l'alignement de corpus parallèles. Elle est la première à introduire et délimiter la notion de corpus comparables, comme un ensemble de textes en partie traduits et en partie non-traduits<sup>18</sup> Les conditions de composition du corpus sont plus strictes que pour les corpus multilingues: '*Both corpora should cover a similar domain, variety of language and time span, and be of comparable length. The translation corpus [la partie du corpus comparable qui comprend les textes traduits] should be representative in terms of the range of original author and of translators*' (Baker : 234) Elle met en avant que les résultats de l'analyse de corpus comparables peuvent contribuer aux analyses de corpus parallèles mettant l'accent sur les différences pertinentes entre les textes traduits et les textes non-traduits. Ses deux suggestions au niveau de différences structurelle sont l'affichage d'un ratio type/token et la densité lexicale d'un texte comparable, une recherche après les « schémas » (*patterns*) dans l'emploi langagier de la traduction, par rapport aux textes originaux ou non-traduits (Baker : 236) Elle suggère déjà une tendance universelle responsable pour le changement de ratio type/token dans les comparaisons de textes traduits et non-traduits : la *simplification lexicale* qui se produit '*in a variety of mediated communicative activities, including translation*', cite-elle de Blum-Kulka and Levenston. (Baker : 236) Avec la simplification, le processus de la simplification par rapport au texte-source est désigné et résulte dans un texte qui porte des traits de ce changement. Néanmoins, elle exprime également sa crainte que les résultats soient peu applicables directement 'en classe' e qui aurait pour conséquence des difficultés pour la collecte de fonds financiers afin de créer de plus grands corpus comparables.

Dans les derniers 15 ans, les suggestions de Baker ont mené à l'investigation d'un nombre de traits universels de traduction probables. Depuis les années 1990, le nombre de corpus multilingues dans les langues de l'UE augmente, ainsi que leur ampleur. Deux exemples de la conception de corpus multilingues sont le NERC, Network of European textual Reference Corpus, comprenant neuf langues européennes, et le projet Parole qui vise la composition de corpus comparables, un projet subventionnée par l'Union Européenne. (Laviosa : 10) Dans les années 2000, la composition de nouveaux corpus comparables s'est développée en partie indépendamment de la recherche dans les branches théoriques et appliquées de la traductologie. Les corpus comparables sont étroitement liés à la branche de la DTS et les deux objectifs principaux de Toury. Selon Laviosa, les deux objectifs de recherche sont 'la découverte des régularités dans la traduction dans les tendances universelles', et ils doivent en même temps mener à la définition des normes qui restreignent ce « comportement traductionnelle ». Finalement les régularités observées doivent mener à montrer les normes opérationnelles (notion de Toury, 1995) qui limitent le 'comportement traductionnel' (*translational behaviour*) (Laviosa : 20)

A l'heure actuelle, il s'avère que les grands corpus conçus et les corpus en préparation sont en majorité parallèles ou créés pour les recherches de nature diverses, comme le *English-Norwegian Parallel Corpus*. Ce corpus peut être utilisé comme corpus parallèle ainsi que comme corpus comparable. Les quatre sous-corpus, les ensembles de textes anglais originaux, des textes anglais traduits du norvégien, des textes norvégiens originaux et de textes norvégiens traduits de l'anglais, peuvent être sélectionnées pour des objectifs de recherche différents. (Johansson, p. 39) Les objectifs de recherche récurrents dans les recherches parallèles ainsi que les corpus comparables constituent le renforcement des théories, l'amélioration de la méthodologie et l'expansion du corpus initial avec de nouvelles données.

---

<sup>18</sup> Baker ajoute dans une de ses remarques que la terminologie autour les types de corpus tend en 1993 déjà à la confusion. (p. 240)

## 2.2. La terminologie

Le point de départ pour la détermination de la terminologie pour notre recherche est l'analyse méthodologique de Granger (2003)<sup>19</sup>, auxquels quelques remarques dans d'autres méthodologies de Baker, Laviosa et Johansson seront ajoutées.

Comme on l'a déjà vu dans l'article de Baker(1995), la terminologie de la définition des corpus est loin d'être claire. J'adopte la tripartition de Granger (2003) qui est expliquée par de nombreux autres articles théoriques. Granger, d'abord, remarque que la linguistique contrastive a une autre terminologie afin de distinguer les types de corpus. En outre, elle propose une multitude de combinaisons de caractéristiques des corpus.<sup>20</sup> La terminologie de la linguistique contrastive n'est pas adoptée pour notre recherche, à l'exception d'un seul type de corpus, constituant un alternatif au choix du corpus bilingue dans le chapitre 3.4 : un corpus appelé également 'multilingue', qui contient des textes originaux dans une ou plusieurs langues. Pourtant, les textes insérés ne sont pas marqués par des traces de traduction, comme les textes en langue-source et langue-cible dans les corpus de traduction. La problématique réside dans l'équivalence des textes : dans un tel corpus, il est difficile d'établir des textes équivalents en termes du contenu sémantique. Ce type de corpus ne semble donc pas apte à notre recherche, ni à l'intérieur de la branche de recherche de la traductologie.

Dans la traductologie, on peut distinguer trois types généraux de corpus: la notion de **corpus comparable** est définie comme deux ensembles de textes dans la même langue, l'un de textes originaux, appelés non-traduits et l'autre de textes traduits de une ou plusieurs langues-sources (Baker : 234) Le corpus comparable est dès maintenant donc considérée comme monolingue. Un corpus comparable peut servir à analyser 'the naturalness of translated texts' ce qui est considéré comme un critère de qualité de traductions. (Johansson, 1998 dans Oholan : 13) Le **corpus de traduction** dénote donc la partie de textes traduits. Un exemple d'un corpus de traduction est le Translational English Corpus (TEC).<sup>21</sup> La partie de textes non-traduits sont souvent issus de grands corpus de référence monolingues, comme le British National Corpus (BNC). Le troisième type, des corpus comparables, a été défini déjà dans le chapitre 2.1. Dans le chapitre 3.4, les corpus utilisés dans notre recherche sont décrits à l'aide de ces distinctions. Les recherches qui sont exécutées à l'aide de ces corpus ont souvent la forme d'une comparaison sur le niveau de la langue, entre les textes traduits et non-traduits. On essaye premièrement à révéler des traits universels qui sont supposés d'être présents dans les textes traduits. Autant plus que la simplification, les textes traduits sont considérés comme plus explicites sur plusieurs niveaux (**explicitation**) ainsi que le langage utilisé est plus conventionnel ou normatif (**normalisation**). (Oholan : 37)

A son tour, la méthodologie de la linguistique de corpus consiste en la création de corpus et le développement des outils informatiques pour l'analyse de corpus. La linguistique de corpus ne se distingue pas seulement de la linguistique générale par sa méthodologie et son objet d'études mais surtout par son approche sur les sciences du langage dans lequel le corpus constitue le point de départ (*corpus-based studies*). Les corpus qui sont utilisées à l'intérieur de la traductologie sont généralement délimités par les critères suivants. Premièrement, les textes doivent obligatoirement être authentiques, c'est-à-dire non-intuitifs : l'introspection est rejetée comme méthode d'analyse de phénomènes linguistiques, et le support par corpus (*corpus evidence*) est considérée comme la seule méthode d'analyse de l'emploi langagier (Laviosa : 9) En outre, une nouvelle condition importante à

---

<sup>19</sup> Granger, S. 'The corpus approach: a common way forward for Contrastive linguistics and translation studies?' dans Granger, S. et al. *Corpus-based approaches to contrastive linguistics and translation studies*. Rodopi, Amsterdam (2003) p.17-30

<sup>20</sup> Son table de 'corpora in cross-linguistic research' se trouve dans l'article 'The corpus approach: a common way forward for Contrastive and Translation Studies?', Granger 2003: 21.

<sup>21</sup> <http://www.monabaker.com/tsresources/TranslationalEnglishCorpus.htm> le TEC corpus n'est malheureusement plus accessible sur internet.



la conception des corpus est le critère que les textes doivent être préparés pour le traitement automatique : ‘... a collection of texts held in machine-readable form and capable of being analysed automatically or semi-automatically’(Baker : 225)

Mamkjaer (1998) a élaboré une comparaison entre les recherches de corpus parallèles, incluant également les recherches exécutées dans le domaine de la linguistique contrastive et computationnelle. Il met en avant deux désavantages à l’analyse de corpus aux fins traductologiques. Premièrement, les lignes de concordance ‘n’offrent pas toujours assez de contexte linguistique pour faire l’inventorie des caractéristiques de textes entiers ou de phénomènes sémantiques. La recherche par ‘keywords in context’ (KWIC) n’accède probablement pas à certains types de comportement traductionnelle puisque les possibilités du concordanceur sont limitées. Deuxièmement, Malmkjaer met en avant la nécessité de l’alignement manuel et l’analyse de plusieurs traductions par rapport à un seul texte-source. Les constatations issues de l’analyse manuelle peuvent être vérifiées dans un plus grand corpus. (Laviosa : 27-28)

### 2.3. Les applications de l’analyse de corpus dans des domaines divers

Comme l’analyse de corpus a pris un grand essor dans les années 1990, les tendances dans le domaine de la traductologie sont marquées par l’optimisme envers le renouvellement de la méthodologie afin d’établir des analyses aux fins pluridisciplinaires :

*‘The systematic creation and investigation of corpora is continually opening up new areas of research in descriptive and applied linguistics. Contrastive linguistic, lexicography, educational linguistics, computational linguistics, terminology, second language acquisition, forensic linguistics, literature studies, and critical linguistics have all experienced significant shifts in their methodological and theoretical approaches, thanks to the influence of corpus linguistics.’*  
(Laviosa 2003: 45)

Dans son article ‘Corpora and translation studies’, Laviosa (2003) applique cette tendance à sa proposition pour une nouvelle branche à l’intérieur de la traductologie, le CTS (*Corpus-based translation studies*) qu’elle considère comme une nouvelle branche qui se distingue explicitement de la DTS. En passant sur la discussion de l’insertion de CTS<sup>22</sup>, elle conclut que la contribution la plus importante de CTS constitue sa méthodologie ‘flexible’ envers les branches diverses de la traductologie et, comme le montre la citation ci-dessous, favorise l’échange de résultats avec les branches de recherche en dehors de la traductologie :

*‘Already a corpus-based descriptive approach is emerging in the related disciplines of contrastive linguistics and bilingual lexicography, which specifically use translational corpora not to evaluate the product of translation or to improve translator performance, but to investigate linguistic phenomena which are of interest to translation scholars and practicing translators. Conversely, the insights provided by descriptive CTS can enhance research in these disciplines, by providing evidence, for example, about language pair- and translation-specific phenomena’*  
(Laviosa 2003: 45)

Comme l’apprentissage de la langue étrangère, les corpus parallèles et comparables peuvent également être intégrés dans la classe, remarque Laviosa. Suite à l’encadrement établi dans le chapitre 1.4., l’intérêt de cette constatation réside dans l’application dans d’autres domaines de recherche que la DTS, en particulier dans les domaines de la linguistique contrastive, de la lexicographie ainsi de la didactique de traduction. A part des applications multiples de l’analyse de corpus à l’intérieur de la traductologie, les résultats de l’analyse de corpus peuvent être également appliqués dans les domaines en dehors de la traductologie.

---

<sup>22</sup> Voir Laviosa, S. ‘Corpora and translation studies’ dans Granger, S. et al. *Corpus-based approaches to contrastive linguistics and translation studies*. Rodopi, Amsterdam (2003)p. 49-52

Dans le domaine de la linguistique contrastive, Johansson(2003) discute les applications de l'analyse de corpus et plaide en faveur de l'application d'une 'nouvelle génération de grammaires et de dictionnaires, fondé sur l'étude de l'emploi langagier', en particulier de nouvelles dictionnaires bilingues.<sup>23</sup> Le fondateur du projet COBUILD, J. Sinclair est un des linguistes pionniers sur le plan de l'insertion de l'analyse de corpus dans le domaine de la lexicographie. En 1987, la première version de *Collins' Cobuild English Language Dictionary* voit le jour, un dictionnaire basé sur un lexique de nature 'corpus-driven' et vise les apprenants de l'anglais langue étrangère et donc le développement d'outils didactiques et descriptives.

L'analyse de corpus est également utilisée afin de développer des outils de recherche électronique (*multilingual software tools*). J'utiliserai un outil d'analyse spécifiquement développée pour l'analyse bilingue et multilingue, ParaConc, voir le chapitre 3.4.

Un autre exemple d'un domaine de recherche qui porte l'intérêt dans le traitement des verbes supports est le traitement automatique de langage (*natural language processing*, NLP). M. Salkoff (1990) souligne l'importance de l'insertion des constructions Vsupp en entier dans les systèmes de traduction automatique.<sup>24</sup> Selon lui, les constructions à verbe support doivent être pris en compte dans l'analyseur syntaxique (*parser*) ainsi que dans le lexique d'un programme de traduction automatique. Il propose que les Npréd qui vont avec le Vsupp doivent être marqués comme combinatoire avec le Vsupp en question. Ainsi, on pourrait faire face à plusieurs problèmes dans la traduction, d'entre autres l'absence d'un Vsupp équivalent dans la langue-cible, ou l'absence de toute la construction Vsupp dans la langue-cible. Par conséquent, on peut s'attendre à des traductions irrégulières des constructions Vsupp.

Par cette analyse, j'ai essayé de soutenir ma proposition pour un élargissement de l'insertion de l'analyse de corpus dans des domaines pluridisciplinaires dont le domaine de notre recherche consiste en le domaine partagé par la linguistique contrastive, la linguistique de corpus et la lexicographie. L'ensemble est encadré dans la traductologie par l'objectif d'améliorer les outils pour la traduction et l'application dans un contexte lexicographe bilingue. Dans le chapitre 3, la pluridisciplinarité est concrétisée dans la conception d'une méthodologie et le point de départ dans un dictionnaire bilingue existant susceptible d'être un dictionnaire pré-corpus.

#### **2.4 Les critères pour l'analyse de corpus dans le domaine de la traductologie**

Le linguiste de corpus J. Aarts (1999)<sup>25</sup> a défini un ensemble de critères de conception pour un modèle descriptif de l'emploi langagier, ayant l'objectif de 'référer au produit linguistique écrit ou parlé et pas au processus de production ou de l'activité linguistique'(Aarts : 6-7) Comme notre recherche ne vise pas l'analyse dans le cadre de la DTS, on peut adopter ses critères de la conception d'un modèle de l'emploi langagier afin de délimiter un corpus 'idéale' par rapport à notre objectif de recherche.

Le facteur principal de l'analyse de Aarts est que les données montrent le véritable emploi langagier: la linguistique contrastive s'est toujours opposée à la grammaire générative à cet égard. Le rejet de l'analyse de corpus dans le courant de la grammaire générative est basé sur la représentation de la langue performante, là où le courant Chomskyenne tend à favoriser les données introspectives.

---

<sup>23</sup> Johansson, S. 'Contrastive Linguistics and corpora' dans Granger, S. et al. *Corpus-based approaches to contrastive linguistics and translation studies*. Rodopi, Amsterdam (2003)p. 41

<sup>24</sup> Salkoff, M. 'Automatic translations of support verb constructions', 1990,p. 243-246.  
[www.aclweb.org/anthology/C/C90/C90-3043.pdf](http://www.aclweb.org/anthology/C/C90/C90-3043.pdf)

<sup>25</sup> Aarts, J. The description of language use. The In H. Hasselgård and S. Oksefjell (eds.) *Out of corpora: Studies in honour of Stig Johansson*. Amsterdam: Rodopi, 1999, p. 3-20

- 1) Le modèle doit permettre une combinaison de la description qualitative et quantitative des données
- 2) Le modèle doit obligatoirement établir une relation entre les facteurs à l'extérieur de système langagier et les phénomènes langagiers internes
- 3) Le modèle doit permettre la description de l'ensemble de variétés langagières
- 4) Le modèle doit permettre la description intégrée de caractéristiques syntaxiques et lexicales ainsi qu'une analyse de discours (Aarts dans Oholan : 15)

Parmi les quatre critères, le troisième a peu de pertinence pour notre recherche : le corpus parallèle aux fins traductologiques est délimité à l'emploi langagier non-spontané et édité. Le critère le plus pertinent au contraire, se trouve dans l'analyse combinée des données qui met ensemble les données qualitatives et les données sur la fréquence des occurrences dans l'emploi langagier. Le deuxième critère est un point de critique sur la linguistique contrastive des années 1970 et 1980. A présent, la linguistique contrastive a évolué elle-même, comme il s'est avéré du chapitre 2.1. Bien que la DTS donne la priorité aux facteurs externes, elle ne passe jamais sur les aspects langagiers spécifiques à la langue-source.

Notre recherche peut être définie comme une '*corpus-driven analysis*' et remplit la deuxième condition en respectant le statut des textes (voir le chapitre 3.4). Notre méthode de recherche est tirée de la linguistique contrastive, puisque les conditions et objectifs de Laviosa(2002) sont spécifiques à la DTS. Néanmoins, il faut que notre analyse rende compte des exigences générales que pose l'analyse de corpus à l'intérieur de la traductologie. La première condition que pose Laviosa pour l'analyse de la production de textes traduits est l'élaboration de critères explicites et des processus de sélection et annotation des textes. Cette condition peut non seulement contribuer à la discussion de la méthodologie de notre analyse mais aussi à l'analyse en général à l'intérieur de la traductologie. L'encadrement de notre recherche est élaboré dans le premier chapitre ; le troisième chapitre sera consacré à la description détaillé et les méthodes de recherche. La deuxième condition que pose Laviosa(2002) et Baker (1995) est la définition précise des traits linguistiques qui sont considérées comme d'expressions des universels de la traduction comme la simplification, l'explicitation, et la normalisation. Du premier chapitre, il s'est avéré que ces traits universels de traduction ne sont pas pertinents dans notre recherche. Toutefois, il faut être attentif aux traces éventuelles de traduction dans les textes sélectionnées dans le corpus bilingue à analyser. Ce 'troisième facteur' éventuel nécessite la vérification d'une analyse manuelle de toutes les données trouvées, même quantitatives. Par conséquent, les données de corpus monolingues serviront comme contrôle à la distribution des occurrences dans le corpus bilingue.<sup>26</sup>

Dans ce chapitre, on a essayé à mieux définir l'avènement de la linguistique de corpus, à délimiter les types de corpus d'intérêt pour notre recherche. En outre, ce chapitre soutient l'hypothèse que les applications de l'analyse de corpus sont diverses.

---

<sup>26</sup> La seule référence au support de corpus monolingues relève de Laviosa (2004 : 27-28) mais cette constatation est également la conclusion de plusieurs discussions lors la conception du corpus de recherche bilingue.

## Chapitre 3 : Méthodologie, les corpus et Van Dale

Comme nous l'avons vu dans les chapitres précédents, l'analyse de corpus offre de multiples possibilités pour contribuer à la fois à la connaissance sur l'emploi et la structure des Vsupp et à la création d'une nouvelle entrée lexicographique. Dans le chapitre 3, nous passons de l'approche théorique des objectifs de l'analyse de corpus à la conception d'une méthodologie apte à nos objectifs de recherche. La première section est consacrée à la discussion sur la notion de *dictionnaire pré-corpus* et l'utilisation de l'analyse de corpus dans la conception d'un dictionnaire monolingue allemand et des dictionnaires bilingues français-anglais et français-néerlandais. Ensuite, nous examinerons l'insertion de l'analyse qualitative et quantitative dans le domaine de la traductologie (3.2.1.), dans une proposition pour un dictionnaire allemande monolingue (3.2.2.) en dans une proposition pour un dictionnaire bilingue espagnol-néerlandais (3.2.3.). Nous développerons une méthodologie apte à notre recherche (3.3.) et nous formulons les contributions attendues des corpus individuels (3.4.). Finalement, le dictionnaire bilingue Van Dale constitue le point de départ pour une analyse après la distribution lexicale du verbe *commettre*.

### 3.1. La notion de *dictionnaire pré-corpus*

D'abord, il faut admettre que toute analyse est en service de notre objectif final de présenter une nouvelle conception de l'entrée lexicale pour les Vsupp. Comme je m'attends à ce que la recherche puisse contribuer à la manière d'affichage des verbes supports dans les dictionnaires existants, il faut d'abord examiner la situation actuelle de la lexicographie des dictionnaires bilingues. Il faut savoir de quelle façon les dictionnaires se servent des méthodes d'analyse de corpus afin d'améliorer la lexicographie existante. J'analyserai d'abord les constatations des linguistes-lexicographes Hanks et al. (2006) Hanks et Pustejovsky (2004, 2005), et Corbin et Gasiglia (2009) sur la situation lexicographique actuelle dans les éditions des dictionnaires les plus nouvelles afin de déterminer le renouvellement lexicographique du dictionnaire sélectionné, Van Dale bilingue français-néerlandais.

Tout comme dans la linguistique contrastive, la lexicographie a connu un développement énorme grâce à la disponibilité de l'Internet et les possibilités pour la composition des corpus. Cependant, la lexicographie connaît une progression lente selon Hanks et al. (2006)<sup>27</sup>. Dans leur article, 'German light verb constructions in corpora and dictionaries', ils constatent que les nouvelles éditions de dictionnaires allemands sont typiquement des versions révisées d'éditions précédentes :

*However, lexicography is a slow-moving profession and tends to be accretive, new dictionaries typically being revised versions of old ones, rather than radical new departures. There are a several reasons for this, not least the high cost of undertaking a completely new analysis and a general feeling (whether justified or not) that existing dictionaries, while not perfect, are 'good enough' and only need an occasional update, not radical restructuring.'*  
(Hanks et al : 439)

Ils soulignent que les investissements nécessaires d'une analyse lexicographique tout à fait nouvelle constituent une prise de considération importante auxquels s'ajoute le point de vue rédactionnelle prédominante que les dictionnaires existants, fonctionnent assez bien et ne demandent qu'une mise à jour occasionnelle. Par conséquent, les dictionnaires ne demanderaient pas de restructuration à grande échelle. Ils ressentent que ce sont ces facteurs qui freinent l'insertion et le développement de nouvelles approches linguistiques pour la lexicographie.

P. Corbin et N. Gasiglia (2009b) ont élaboré une analyse de la lexicographie française et l'ont comparé la situation pour le français à la situation dans le monde anglo-saxonne dans l'article 'Changer les

---

<sup>27</sup> Hanks, P, Urbschat, A. et Gehweiler, E. 'The syntagmatics of metaphor and idiom', *International Journal of Lexicography*, 2006, p. 439-457.

dictionnaires ? Une pluralité d'approches'.<sup>28</sup> Ils présentent l'hypothèse qu'un dictionnaire est souvent plus enraciné dans l'intuition des locuteurs et dans l'emploi normatif de la langue, que fondé sur l'emploi langagier. Ils mettent en avant l'importance d'un dictionnaire reflétant l'emploi langagier. De cette façon, une approche linguistique peut contribuer en particulier à l'amélioration de la lexicographie descriptive. Ils signalent qu'en France, à l'initiative de l'université de Lille 3, plusieurs nouvelles propositions linguistiques, comme des approches informatiques et des données métalexigraphiques sont analysés et élaborés (C&G : 8).

Un des pistes de réflexion s'étend vers le développement récent d'une plus grande édition de dictionnaires électroniques. La plus grande contribution des environnements électroniques à la lexicographie serait la conception d'une organisation plus hiérarchisée par rapport au dictionnaire imprimé. Contraire aux attentes, Corbin et Gasiglia constatent que

*'...la version électronique du Robert Junior n'est ni la première à ne plus correspondre à la version imprimée d'un même dictionnaire, ni le même répertoire à perdurer dans un catalogue sans actualisation significative, ni la première victime de décisions à court terme et à courte vue.'* (C&G : 14)

Dans ce point de vue critique, ils soulignent que les décisions prises 'à court terme et à courte vue' d'édition numérique ne contribuent pas à la restructuration de l'ensemble du contenu du dictionnaire.

En analysant un article de H. Béjoint (2009), Corbin et Gasiglia décrivent la situation dans le monde anglo-saxonne plus 'florissante' que dans la monde francophone : les dictionnaires monolingues sont mises à jour à l'aide de nouvelles innovations linguistiques.<sup>29</sup> Comme points forts, Corbin et Gasiglia soulignent les 'corpus coopératifs' dans le BNC qui soutiennent plusieurs grands dictionnaires, ainsi que le courant de l'innovation linguistique avec en tête le Collins COBUILD (voir le chapitre 2.1) qui incite aux diverses applications des statistiques de cooccurrence de mots. Le public visé dans ces dictionnaires est en majorité l'apprenant de l'anglais langue étrangère. Les résultats de recherche contribuant entre autres à la '*simplification de la description sémantique, [...] à la clarté de son affichage, au contrôle du lexique qui est mis en œuvre et à l'articulation avec les contextes d'emploi des items*' (Béjoint 2009, dans C&G :16) En ce qui concerne l'affichage de collocations, appelées combinatoire lexicale, il est bien développé dans la tradition française mais également critiqué pour leur qualité des descriptions : 'Déficit d'affichage ou manque de systématité dans le traitement des cooccurrences peuvent être au débit de la lexicographie française' (C&G :17-18)

De cette analyse, il s'avère que dans les dictionnaires anglais, la priorité est donnée non seulement à la combinatoire lexicale mais aussi la systématité d'affichage. En outre, la lexicographie anglaise prend une avance sur la situation française par rapport à l'insertion linguistique contribuant à la conception de nouvelles approches pour les dictionnaires pour allophones.

Corbin et Gasiglia sont très nuancés sur l'emploi des données obtenus de l'analyse de corpus dans les dictionnaires unilingues français et anglais. Ils signalent que de plus en plus, les approches linguistiques, comme la sémantique des cadres, sont insérées dans la création d'une nouvelle distribution d'informations dans des dictionnaires français et anglo-américains. Hanks et al. (2006) par contre, considèrent dans leur analyse le *Deutsches Universalwörterbuch* (DUW) comme un **dictionnaire pré-corpus**, en le définissant comme un dictionnaire dans lequel le support de l'analyse de corpus n'est pas ou guère intégré.

Les analyses de l'allemand, du français et de l'anglais mènent à la situation des dictionnaires néerlandais. Leurs procédures d'édition et les méthodes de composition et de sélection, comme pour le dictionnaire bilingue Van Dale, ne sont souvent pas expliquées clairement. L'édition de 2000 de Van Dale français-néerlandais est une réimpression de l'édition de 1985. Le rédacteur P. Bogaards,

---

<sup>28</sup> Corbin, P. Gasiglia, N., « Changer les dictionnaires ? Une pluralité d'approches », *Lexique 19* « Changer les dictionnaires ? », Pierre Corbin & Nathalie Gasiglia eds., 2009 b, p. 7-38.

<sup>29</sup> Béjoint, H. 'Lexicographie et linguistique : quelques réflexions sur l'anglais', *Lexique 19*, p. 117-158

explique dans l'introduction les plus grands changements dans la nouvelle édition, à savoir des investissements pour la langue-cible, le Néerlandais, avec entre autres l'insertion de 'plus de traductions'<sup>30</sup>. Ce dernier facteur aurait eu la priorité pour cette édition, bien qu'un 'omvangrijk nieuw bestand Nederlands' ne fût pas encore disponible pour le travail de rédaction au moment de la troisième édition. La quatrième édition de 2008 est une série de 'Groot Handwoordenboeken' et affiche une entrée lexicale de *commette* qui est identique à la version électronique de 2010. Dans l'introduction, le corpus dont il avait question n'est plus mentionné. Il semble être plutôt un renouvellement d'affichage au niveau de l'emploi du dictionnaire au lieu d'un renouvellement du contenu des entrées.

La question se pose dans quelle mesure le support de corpus est intégré dans le Van Dale électronique 2010. Par manque de plus d'informations sur le renouvellement lexicographique, je définis le statut du dictionnaire Van Dale comme inconnu. L'utilisation de l'analyse de corpus peut être visible de manière indirecte dans le choix des exemples qui reflètent la distribution qualitative de l'emploi d'un verbe support. En outre, la variation sémantique donne une certaine indication, bien que les choix de l'édition puissent s'y mêler : on ne peut pas supposer que la restriction lexicale est liée seul aux ressources de la conception lexicographique.

Notre analyse permettra d'examiner de quelle façon les données reflètent l'évidence éventuelle du corpus. Il se peut que les dictionnaires Van Dale sont toujours marqués de l'analyse d'un corpus, au moins dans une des deux langues dans un dictionnaire bilingue. La première fonction de l'analyse de corpus est d'évaluer la distribution, la quantité et la pertinence des informations dans l'entrée lexicographique de Van Dale 2010.

La question de recherche pour l'analyse de la situation allemande du Deutsches Universalwörterbuch (DUW) était comment l'entrée lexicale dans ce dictionnaire peut être améliorée par l'insertion du support de l'analyse de corpus (Hanks et al : 440). Dans la recherche de ce mémoire, la même constatation est probable pour le Van Dale. Notre dictionnaire est de nature différente : le dictionnaire est bilingue et électronique au lieu de monolingue et imprimé. Par conséquent, une analyse du dictionnaire Van Dale constitue la première phase dans la conception d'un nouveau type d'entrée lexicale afin de mieux définir les contributions et la fonction des corpus qui j'utiliserai. Il s'agit de trois corpus monolingues Frantext, LexisNexis et la partie en français du corpus EU (corpus parallèle Europarl) et un corpus bilingue, un corpus également issu du corpus EU.

Une analyse des entrées lexicales des Vsupp dans Van Dale apportera une bonne compréhension des critères de distribution des constructions Vsupp et les variations lexicales dans l'affichage, et fournit des possibilités de recherche électronique. *Commette* est un Vsupp non seulement relativement restreint sur le plan syntaxique mais aussi d'un verbe qui combine avec un grand nombre de Npréd. Je m'attends donc à une grande variation lexicale.

La section suivante est consacrée d'abord à la méthodologie d'autres projets de conception d'une nouvelle entrée lexicale, et insiste sur les aspects qualitatifs et quantitatifs de l'analyse de corpus.

### **3.2 La recherche qualitative et quantitative**

Dans cette section, la recherche quantitative et qualitative est d'abord introduit dans le domaine de la traductologie, et ensuite fourni d'un encadrement théorique et une analyse de deux conceptions pour une entrée lexicographique.

#### **3.2.1. L'analyse de corpus appliqué à la traductologie**

---

<sup>30</sup> Edition 1985, NF, derde druk 2000. 'Ter inleiding', IX-XI. 'Dans cette édition', P. Bogaards.

J'ai choisi d'effectuer une analyse détaillée de deux propositions pour une entrée pour les Vsupp afin d'intégrer les résultats de cette analyse de corpus dans une proposition pour une entrée lexicale bilingue. L'analyse de Hanks et al (2006) applique des critères de sélection aux constructions Vsupp mais il faut que les conditions de recherche monolingues soient adaptées à une recherche bilingue. En outre, cette analyse contribuera surtout à la méthodologie. Comme deuxième proposition, j'ai choisi le dictionnaire bilingue Spaans-Nederlands par J. Schroten (1989) afin d'évaluer la méthode d'affichage des Vsupp.<sup>31</sup> Dans une réflexion générale sur la problématique sémantique des constructions Vsupp, Schroten propose une nouvelle approche de l'affichage des verbes supports dans les dictionnaires bilingues qui lie l'approche théorique à un dictionnaire bilingue édité. En vue de la langue ainsi que le public visé, le Collins'COBUILD ne sera pas analysé malgré le fait que ce dictionnaire peut être considéré comme le dictionnaire exemplaire de l'insertion de l'analyse de corpus.

### 3.2.2. Hanks et al. (2006 Hanks (2004) et Hanks&Pustejovsky (2005)

Parmi les trois articles, Hanks et al 'German light verb constructions in corpora and dictionnaires' (2006<sup>32</sup> constitue l'élaboration la plus détaillée d'une approche qui combine la linguistique de corpus avec l'affichage des verbes support dans une entrée lexicale. La méthodologie utilisée est celle de la 'théorie des normes et exploitations' (*Theory of Norms and Exploitations*, TNE) présenté dans Hanks (2004) et Hanks&Pustejovsky (2005) dans lesquels le support de l'analyse de corpus est intégrée dans la conception lexicographique.

Dans la théorie des normes et exploitations, on part de l'hypothèse que le lexique 'ne consiste pas de mots isolés mais un ensemble de schémas syntagmatiques (*syntagmatic patterns*) qui sont liées à un mot (Hanks 2004 :246).<sup>33</sup> Dans l'article 'the syntagmatics of metaphor and idiom', la notion de schéma est définie comme une structure syntagmatique ayant des valeurs sémantiques. Les valeurs sémantiques sont spécifiés en relation avec le mot cible, dans le cas des constructions à verbe support le Npréd, souvent en position de d'objet direct. La notion de valeurs n'est pas expliquée de façon claire bien que cela constitue la différence entre l'analyse de noms et l'analyse de verbes. Bien que Hanks (2004) adhèrent au principe de différences individuelles pour les associations de chaque mot, l'emploi langagier (*language use*) est stéréotypé du fait que les emplois non-stéréotypés sont supprimés grâce aux normes d'une communauté langagière. (Hanks et al : 246)

L'article 'A pattern dictionary for Natural Language Processing' de Hanks et Pustejovsky (2005) insiste plus sur la définition des emplois (non)-stéréotypés.<sup>34</sup> L'objectif de la TNE est d'obtenir une catégorisation lexicologique afin d'identifier des emplois prototypiques en collocation, appelés *emplois normaux*. La catégorisation des emplois normaux dans les schémas dépend de la façon dont les schémas (*patterns*) d'emploi de la langue ont été définis. La théorie distingue entre l'emploi normal et l'emploi exploité, qui désigne la variation sur les emplois normaux. Le premier a de l'intérêt lexicographique, tandis que le second constitue entre autres le domaine d'effets rhétoriques. Finalement, la notion d'**emplois non-normaux** couvre les emplois exploités ainsi que les alternances lexicales des emplois normaux. (Hanks et Pustejovsky 2005 : 1-2)

---

<sup>31</sup> Schroten, J. (m.m.v. D. Bloemraad, M. Buursma, N. Gerbrandy, T. Ceelen, J. Hulst, J. Boone, B. van Maris, M. Inklaar). *Spaans-Nederlands woordenboek*. Muiderberg, Coutinho, 1989, xi-807 pags.

<sup>32</sup> Hanks, P. Urbschat, A., Gehweiler, E. 'German light verb constructions in corpora and dictionnaires' *International Journey of Lexicology*, 2006, p 439-457.

<sup>33</sup> Hanks, P. 'The syntagmatics of metaphor and idiom', *International Journal of Lexicography*, 2004, p. 245-272.

<sup>34</sup> Hanks, P. and Pustejovsky, J. 'A pattern dictionary for Natural Language Processing', *Revue Française de Langue appliqué*, (2005), 10 :2

Si on considère la langue conceptuelle comme la base des dictionnaires pré-corpus, le support issu des corpus montre l'emploi d'une langue de manière qualitative ainsi que quantitative. Hanks (2004) constate que les dictionnaires actuels fournissent surtout des syntagmatiques inadéquates, principalement puisque des emplois individuelles sont insérées. Il propose l'**analyse de schémas de corpus** (*corpus pattern analysis*, CPA) comme un appareil pour établir les normes prototypées pour des mots individuelles. Par comparaison avec FrameNet (Fillmore 1975), la sémantique des cadres, la CPA est présentée comme un système plus pratique et appliqué aux mots au lieu de cadres. Il contraste l'emploi des techniques d'introspection à l'emploi des corpus; par conséquent, la priorité est donnée au support du corpus.

Hanks prend l'exemple du verbe *take* pour exemplifier l'emploi les verbes complexes : *to take* fournit plus de 200 normes. A l'aide du verbe moins complexe *to abate*, il explique que les emplois qui ne sont pas des normes, constituent des alternances ou bien des exploitations. Le premier terme n'est pas clairement défini mais seulement exemplifié par une alternance sur le nom [humain]: [groupe humaine], [personne], [partie politique], [association sociale] (Hanks 2004 : 253)

Pour la méthodologie, Hanks adopte la catégorisation suivante :

*'The methodology is to extract a concordance for each target word, scan it to get a general overview of the word's behavior, then select a random sample of between 200 and 1000 concordance lines for detailed analysis. In the course of detailed analysis, concordance lines are sorted into groups that have approximately the same meaning and similar syntactic structures. Semantic values are given for the arguments or valencies of the target word in each group.'* (Hanks 2004 :255)

Les collocations trouvées doivent être classifiées individuellement. Hanks et al. (2006) mettent en avant d'afficher seulement les normes dans la nouvelle entrée lexicographe. L'article et l'entrée lexicale en annexe ne mentionnent pas les informations sur le nombre d'occurrences analysées. En ce qui concerne l'ordre d'affichage dans les entrées de Hanks et al. (2006), la description n'est pas concrète non plus. Il se peut que l'ordre soit hiérarchique par fréquence décroissante, mais les fréquences ne sont pas intégrées dans les schémas ni dans l'explication dans l'article.

Afin d'analyser la proposition finale de Hanks et al (2006), je veux analyser plus en détail les entrées lexicales des verbes allemands *leisten* et *erteilen* :

### **Table 1 : l'entrée lexicale du Vsupp allemand *erteilen***



**erteilen, erteilte, hat erteilt** /Vb./

[verblasst, mit Nomen, die eine bestimmte Art von Information bezeichnen] geben, zuteil werden lassen

[light verb, with nouns that denote a particular kind of information] issue, give

a) jemandem {eine Zustimmung | eine Genehmigung | eine Erlaubnis | eine Vollmacht | eine Zulassung | eine Befugnis |...} erteilen: Die Jerusalemer Stadtverwaltung hatte am Donnerstag **eine Genehmigung** für das Bauprojekt **erteilt** | Nach Fernsehberichten hatte der Kontrollturm...**dem Flugzeug...eine Erlaubnis** zur Landung **erteilt** | Nachdem die...Kartellbehörde **ihre Zustimmung erteilt** habe, werde das neue Unternehmen noch im August den Betrieb aufnehmen | Viele der 34 blinden Heiminsassen hatten **dem Mann...Vollmachten** für ihre Konten und Sparbücher **erteilt**.

a) give someone permission

b) jemandem {eine Anweisung | einen Auftrag | eine Weisung | einen Befehl | eine Order |...} erteilen: Dann **erteilt** der Mann über Sprechfunk **der Mannschaft** wieder **Anweisungen** | Die Kreditanstalt...hat **der...Gesellschaft den Auftrag** zur Erarbeitung eines...Energieversorgungskonzeptes...**erteilt** | fünf Tage nach der Kapitulation Deutschlands **erteilte** die amerikanische Militärregierung **den Befehl** zur Wiederaufnahme der Produktion.

b) give someone an order

c) jemandem {eine Absage | eine Abfuhr |...} erteilen: Der Fußball-Zweitligaverein **erteilte seinem...Wunschkandidaten...eine Absage**, weil er keine Fußballlehrer-Lizenz besitzt | Die US-Pop-Sängerin hat **ihrem glühenden Verehrer, dem britischen Prinzen William, eine Abfuhr erteilt**.

c) turn someone down

(continued)

---

Appendix: Continued

---

- **etwas eine Absage erteilen:** etwas verurteilen: *Die Zeitung hatte mit Aufklebern der Ausländerfeindlichkeit eine Absage erteilt.* • issue a condemnation of something
- d) **jemandem {eine Rüge | einen Verweis |...} erteilen:** *Außerdem erteilte der Presserat den Lübecker Nachrichten eine Rüge wegen Schleichwerbung | Die Lehrerin wollte Veronika einen Verweis erteilen, weil sie zu spät zum Unterricht gekommen war.* d) reprimand someone
- e) **jemandem {Unterricht | eine Lehre | eine Lektion |...} erteilen:** *Die Lehrkräfte erteilen Unterricht nach der Stundentafel und halten Pausenaufsicht | Das auf Aufklärung erpichte iranische Kino will dem Zuschauer gern eine Lehre erteilen | Danach erklärte er mir die wichtigsten Dinge im Leben und erteilte mir eine Lektion über die allerheimlichsten Geheimnisse von Bibliotheken.* e) give instruction to someone; teach
- **jemandem {eine Lektion | eine Lehre} erteilen:** jemanden zurechtweisen: *Dass die kleine Lehrerin Mayawati aus der unteren Kaste den Mächtigen in Neu-Delhi eine Lehre erteilt hat, das freut ihre Anhänger | Bei einer Nachwahl dagegen wollen viele Wähler ihren Protest gegen die Politik deutlich machen und der Regierung eine Lektion erteilen.* • teach someone a lesson
- f) **jemandem {Informationen | Auskunft | einen Rat | Ratschläge | eine Antwort |...} erteilen:** *Weitere Informationen zu den neuen Bestimmungen erteilt das Offenbacher Finanzamt | Laßt schlafende Hunde liegen: Diesen Rat hat Sir Robert Walpole, der erste Premierminister, seinen Nachfolgern erteilt | Es ist einfacher, anderen Ratschläge zu erteilen, als selbst tätig zu werden | Hansa-Kapitän Beinlich erteilte den permanenten Fragen nach der Zukunft des Trainers eine klare Antwort.* f) give someone information
- g) **jemandem das Wort erteilen:** *Er erteilte dem Delegationsleiter Hyde das Wort und entließ danach die Abordnung.* g) hand over to (another speaker)
- h) **jemandem den Segen erteilen:** (von Geistlichen) durch das Kreuzeszeichen segnen: *Ein protestantischer Bischof erteilte zu Beginn den Segen, und ein jüdischer Rabbiner hatte das Schlusswort | In Vilnius erteilten Erzbischof* h) give a blessing

---

(continued)

## Appendix: Continued

*Backis und Alexii II. in einem ökumenischen Gottesdienst mehreren tausend Gläubigen gemeinsam den Segen.*

- {jemandem | etwas} seinen Segen erteilen: Zustimmung geben: *Die Stadt Bad Homburg hat dem geplanten Neubau... ihren Segen erteilt.*

- give one's approval to someone or something

i) jemandem die Absolution erteilen: (durch einen Geistlichen) jemanden von seinen Sünden lossprechen: *Auf dem Sterbebett ließ er sich von dem greisen Erzbischof Bernard von Palermo die Absolution erteilen.*

i) give absolution to someone

- [übertragen] Zustimmung geben: *In seiner Antwort erteilt der Stadtrat dem Bauvorstand völlige Absolution: Er erachte die Beteiligung... sowohl juristisch als auch politisch und sachlich für unbedenklich | der Widerstand hat merklich nachgelassen, seit das Umweltbundesamt dem Konzept die Absolution erteilte.*

- give one's approval

j) (Rechtswesen) jemandem {ein Patent | eine Lizenz | ein Visum | eine Konzession | ...} erteilen: *Das erste Patent auf einen Schreibapparat war... dem Engländer Henry Mill erteilt worden | Die amerikanische Militärregierung erteilte dem... Verleger... eine Lizenz zur Herausgabe einer Tageszeitung | besorgte Bürger wollen von den zuständigen Behörden wissen, wieso L. eine Konzession für die Gaststätte erteilt wurde | Dafür erteilen die Konsulate... ein... Visum, das für sechs Monate gilt und mehrere Einreisen ermöglicht.*

j) grant someone {a patent | a licence | a visa | a concession | ...}

Source : Hanks et al. (2006) p. 454-456

Hanks et al. (2006) mettent en avant que pour un verbe comme *erteilen*, un seul schéma suffit puisque ce schéma couvre plus de 80% de l'emploi de *erteilen* (p. 446-447). En vue de l'objectif de constituer une nouvelle entrée lexicale pour un public cible de traducteurs, il se peut que ces emplois non-normaux puissent également être intéressants pour le traducteur et qu'il faut insérer une sélection d'alternances pour compléter l'emploi.

A la tête de l'entrée, une définition générale, est affichée : il s'agit d'un verbe sémantiquement semblable à *geben* (donner). Les collocations affichées servent d'exemple de l'emploi en contexte, tandis qu'en théorie, Hanks et al. (2006) passent sur les emplois particuliers, les emplois non-normaux, en développant un seul schéma.

## Table 2 : L'entrée lexicale du Vsupp allemand *leisten*

---

APPENDIX: Sample Entries for a *Dictionary of German Light Verbs*

---

**leisten, leistete, hat geleistet** /Vb./

1 [verblasst, normalerweise mit Nomen, die eine Aktivität bezeichnen]

1 [light verb, normally with nouns denoting an activity]

a) [mod] **Arbeit leisten**: eine (bestimmte Art von) Arbeit zustande bringen: *Das alleinige Ziel sollte es sein, gute Arbeit zu leisten und dafür gut bezahlt zu werden* | *Jeden Monat würden in der Bundesrepublik 240 Millionen Stunden ehrenamtlicher Arbeit geleistet, sagte Frau Nolte* | *Wer fruchtbare Friedensarbeit leisten will, muß auf dem Boden der Wirklichkeit bleiben.*

a) do (a particular kind of) work

- **ganze Arbeit leisten**: eine Arbeit bestmöglich zustande bringen: *Die Sondergerichte hätten in Karatschi ganze Arbeit geleistet, die Zahl der Morde und Überfälle habe abgenommen, seit die Gerichte im Dezember eingeführt worden seien.*

• make a good job of something

b) [mod] **Dienste leisten**: eine (bestimmte Art von) Diensten erweisen: *Als Präsident großer Wissenschaftsinstitute ... hat Pallottino der archäologischen Wissenschaft und der Kultur insgesamt unschätzbare Dienste geleistet* | *Bis zu 75 Prozent aller freiwilligen Dienste würden von Frauen geleistet.*

b) perform services (of a particular kind)

---

(continued)

## Appendix: Continued

- **Dienst** [sing.] **leisten**: (normalerweise) im Militär dienen: *Sein Vater hat noch kurz vor dem Krieg bei der finnischen Garnison in Wyborg **Dienst geleistet** | Frauen dürfen in Deutschland auf keinen Fall **Dienst** mit der Waffe **leisten**.*
  - serve in the armed forces
  
- c) [mod] **Hilfe leisten**: eine bestimmte Art von (besonders offizieller) Hilfe geben: *Jetzt liegt die Bitte des UN-Generalsekretärs auf dem Tisch, in Somalia mit 1500 Soldaten **humanitäre Hilfe** zu **leisten** | Das Rio-Abkommen verpflichtet die Industriestaaten dazu, den Entwicklungsländern **finanzielle und technische Hilfe** zu **leisten**.*
  - e) come to someone's aid, especially in an official capacity
  
- d) **einen** [mod] **Beitrag zu etwas leisten**: einen (bestimmten) Beitrag zu etwas dazu geben: *Aus Umweltschutzgründen müssen Autofahrer **einen wesentlich höheren Beitrag zur Schuldentilgung leisten** als Bahnreisende | Im Eurokorps werden deutsche und belgische Soldaten ... **ihren Beitrag zur Sicherung von Frieden und Freiheit leisten**.*
  - d) make a contribution (to something)
  
- e) **jemandem Widerstand leisten**: sich wehren: *Wie die Gendarmerie weiter mitteilte, **leistete** er bei seiner Festnahme **keinen Widerstand** | Ein Teil des Landes sei besetzt, und dort gebe es Leute, die **der Besatzung Widerstand leisteten** | Ein unbeugsamer Streiter **hört nicht auf, den Kritikern Widerstand** zu **leisten**.*
  - e) offer resistance (to someone)
  
- f) **etwas** (einer Aufforderung) **Folge leisten**: einer Aufforderung nachkommen: *Bitte **leisten** Sie **den Anweisungen** von Polizei und Feuerwehr **Folge** | Einige Zuhörer **leisteten der Aufforderung unverzüglich Folge** und verließen den Saal | »Hände hoch, oder wir schießen!« - Total verduzt **leisteten sie dem Befehl** erst einmal **Folge** und drehten sich dann um.*
  - f) obey orders
  
- g) **einen Schwur leisten**: einen Schwur ablegen: *Der Briefempfänger **leistet einen Schwur** zur Geheimhaltung | Dann **leisteten** sie einander **den feierlichen Schwur**, niemandem den Verwahrungsort zu verraten.*
  - g) swear an oath
  
- h) **einen Eid leisten**: einen Eid schwören: *Der Witwer Thomas Jefferson ging in Washington nach dem Frühstück aus seiner Pension zu Fuß zum Kapitol, **leistete seinen Eid** aus seiner Pension zu Fuß zum Kapitol, **leistete seinen Eid** und war zum Mittagessen wieder zurück | **Einen Eid** auf die*
  - h) (bureaucratic and legal) swear an oath

(continued)

Appendix: Continued

*Verfassung muss niemand leisten | Sie verweigern den Eid auf die Queen, den alle Parlamentarier leisten müssen.*

**i) einen Meineid leisten:** einen Meineid schwören: *Er gibt zu, im Fall Monica Lewinsky einen Meineid geleistet zu haben, um seinen Seitensprung zu vertuschen.*

i) commit perjury

**j) etwas (Schlechtem) Vorschub leisten:** etwas Schlechtes fördern: *Daß sie beim Gesetz keinen Schutz finden, leistet dem Mißbrauch Vorschub | Der Fiskus darf der Korruption nicht Vorschub leisten und sie als Kavaliersdelikt erscheinen lassen.*

j) encourage (a bad action)

**k) Zahlungen leisten:** regelmäßig bezahlen: *Die Zahlungen soll Seat geleistet haben, um in Madrid die umstrittene Genehmigung für ein lukratives Immobiliengeschäft zu erhalten | Nomura wird beschuldigt, illegale Zahlungen an japanische Verbrechergruppen geleistet zu haben.*

k) make regular payments

• **eine Anzahlung leisten:** die erste Zahlung einer Reihe regelmäßiger Zahlungen machen: *Darunter fallen Geschäfte, für die ein Vertrag unterzeichnet und eine Anzahlung geleistet ... wurde.*

• make a down payment

**l) jemandem Gesellschaft leisten:** bei jemandem bleiben; jemanden unterhalten: *nur wenn Rudolf mir Gesellschaft leistete, ging sie fort | Sie malt, und er leistet ihr Gesellschaft.*

l) keep (someone) company

**m) jemandem Beistand leisten:** für jemanden da sein; jemanden unterstützen: *Die Spielführerin wollte trotz des Bänderrisses im Knie nicht vorzeitig abreisen, sondern ihren Kolleginnen Beistand leisten | Unter den Gefangenen sind einige Ärzte, die Beistand leisten.*

m) give (someone) support

**n) Verzicht leisten:** verzichten: *Ich weiß, daß viele glauben: Vielleicht rette ich meinen Arbeitsplatz, wenn ich jetzt Verzicht leiste | Haben Sie den Eindruck, daß die Gesellschaft bereit ist, Verzicht im Interesse der Einheit zu leisten?*

n) do without (something)

**o) Ersatz leisten:** Ersatz geben: *Zeigen sich trotz dieser Sorgfalt nach dem Auspacken dennoch Frostschäden ... so leistet der Lieferant gerne Ersatz | Die Regierung ... hat ... die ... Botschaft Israels aufgefordert, finanziellen Ersatz für Gegenstände zu leisten, die während eines ... Besuchs des israelischen Ministerpräsidenten ... aus dem Gästehaus ... verschwunden seien.*

o) supply a replacement or pay compensation

(continued)

## Appendix: Continued

<p><b>p) eine Unterschrift leisten:</b> unterschreiben: <i>Viele Behinderte sind nur mit Mühe in der Lage, im Auto eine Unterschrift zu leisten   Sie leisteten die Unterschrift, um die Beziehung zu retten   Die Unterschriften können bei einer der Kontaktadressen, nicht jedoch im Rathaus geleistet werden.</i></p>	<p>p) write one's signature</p>
<p><b>q) jemandem Gehorsam leisten:</b> jemandes Befehlen folgen: <i>In bester Hirtentradition haben sie ihrem absoluten Monarchen Gehorsam geleistet und sich gleichzeitig selbst aus der Verantwortung gestohlen   20 Rekruten ... erklärten, daß sie der Bundeswehr keinen bedingungslosen Gehorsam leisten wollten.</i></p>	<p>q) obey (someone)</p>
<p><b>r) Abbitte leisten:</b> um Verzeihung bitten: <i>Mittlerweile haben sie Abbitte geleistet und auf ihrer Titelseite um Verzeihung gebeten   Brandt leistet als Kanzler Abbitte für die Verbrechen der Faschisten.</i></p>	<p>r) ask for forgiveness</p>
<p><b>s) eine Bürgschaft leisten:</b> sich bereit erklären für eine fremde Verbindlichkeit die Haftung zu übernehmen: <i>die finnische Staatsbank will eine Bürgschaft für das Projekt leisten   Der Konzern verpflichtete sich ... als Sicherheit eine Bürgschaft von knapp 1, 4 Mill. DM zu leisten.</i></p>	<p>s) stand as guarantor for something</p>
<p><b>t) Garantie leisten:</b> Garantie übernehmen: <i>Für Ihre Waschkombination leisten wir nach unseren Lieferbedingungen 1 Jahr Garantie   der Verband übernahm die Verwaltung und leistete die Garantie, mindestens 25 Jahre lang in dieser Anlage zu bleiben.</i></p>	<p>t) give a guarantee</p>
<p><b>2</b> [mit indefinitem Pronomen oder nominalisiertem Adjektiv als direktem Objekt]: etwas vollbringen, erreichen: <i>Du kannst doch etwas leisten   Er hat in seinem Fach für die Wissenschaft Hervorragendes geleistet.</i></p>	<p>2 [with an indefinite pronoun or a nominalized adjective as direct object]: achieve something</p>
<p><b>3</b> (Technik) Leistung bringen: <i>Für den Heimgebrauch sollte ein Verstärker, der 50W leistet ... genügen   Das Auto leistet weniger wenn die Gegenkraft (durch Reibung und Luftwiderstand) höher ist.</i></p>	<p>3 deliver</p>
<p><b>4</b> sich [Dativ] etwas leisten: <b>a)</b> sich die Freiheit zu etwas nehmen; sich etwas erlauben: <i>er kann es sich leisten, zu spät zu kommen   er hat sich {eine Frechheit   einen schlechten Scherz   einen Seitensprung} geleistet.</i></p>	<p>4 a) take the liberty of allowing oneself to do something</p>

(continued)

## Appendix: Continued

<p>• <b>sich etwas leisten können:</b> <i>sie kann sich die jugendliche Frisur leisten</i> (=die Frisur steht ihr)</p>	<p>• be able to stand something (=the hairstyle suits her)</p>
<p><b>b)</b> die finanziellen Mittel zu etwas haben: <i>Langfristig wird erwartet, daß sich zumindest schon ein kleiner Teil der Bevölkerung bald ein Auto leisten kann.</i></p>	<p><b>b)</b> be able to afford (something)</p>
<p><b>c)</b> sich etwas gönnen: <i>Er leistete sich einen persönlichen Fitnesstrainer   sie leistete sich einen Eisbecher   heute abend wollen wir uns einmal etwas Gutes leisten.</i></p>	<p><b>c)</b> treat oneself to something</p>

Source: Hanks et Al. (2006) p. 151-154

L'entrée lexicale pour le verbe allemand *leisten* a une structure plus complexe. La définition de la valeur des Npréd, 'normally with nouns denoting an activity' demande plus de détail par les schémas énumérés a-t. Pour le Npréd *Schwur* sous f, il fallait créer un schéma pour ce substantif qui ne connaît pas d'alternances à l'intérieur de cette collocation. Par contre, le Npréd *Garantie* demande quatre sous-catégories (1-4b) dont les significations se divergent non seulement au niveau plan sémantique mais aussi sur au niveau syntaxique pour 4a et b. Les critères de catégorisation utilisés sont entre autres la signification de la collocation entière, les propriétés des compléments, les constructions Vsupp ayant un objet physique sont distingué des objets d'entité abstraite, ainsi que la structure syntaxique distinguant entre noms d'évènement +objet direct, verbe réflexif et nom abstrait spécifique dénotant une quantité (Hanks et al : 445). La conception de cette nouvelle entrée aurait pu être mieux expliquée dans l'article puisque dans cet exemple, il est difficile de retracer de quelle manière les données quantitatives ont été utilisées.

Il s'avère que Hanks et al. (2006) prennent des décisions de catégorisation dans ces schémas premièrement de manière arbitraire, à l'aide de principes généraux, et deuxièmement à partir des valeurs sémantiques de emplois divers. Selon la méthodologie, il se peut qu'un schéma montre des alternances lexicales, à condition que ce soient des collocations prototypiques et non pas des emplois exploités. On peut conclure de l'entrées des Vsupp *erteilen* et *leisten* sont très élaborées en ce qui concerne le nombre d'alternances lexicales affichées. Il en résulte un affichage exhaustif. La méthodologie de Hanks (2004) favorise par contre une entrée lexicale réduite à un ou plusieurs schémas. En outre, l'aspect quantitatif de la méthodologie n'est pas justifié dans la concrétisation des entrées lexicales. Ainsi, il s'avère que malgré la conclusion que *erteilen* a un seul schéma, dix sous-schémas sont affichés ici dans a-j. Par contre, les alternances lexicales sont affichées entre crochets { X | Y } de manière concise en nombre d'exemples par emploi prototypique, un seul exemple est affiché.

Dans l'analyse de corpus que je veux exécuter, j'adopterai partiellement la théorie des normes et d'exploitations afin de formuler les emplois prototypés qui doivent être insérées dans la nouvelle entrée lexicale. Comme j'ai critiqué aussi bien l'ampleur de l'entrée lexicographique que la méthode utilisée pour l'affichage dans Hanks et. al (2006), je vise à intégrer moins de données dans un ordre plus hiérarchisé que dans les entrées lexicales des verbes *leisten* et *erteilen*.

### 3.2.3. Schroten (1989, 1999b, 2002)

En 1989, un nouveau dictionnaire bilingue espagnol-néerlandais est publié, où la méthode d'affichage porte attention en particulier aux verbes en contexte. Le Vsupp espagnol *hacer* (doen) est analysé plus en détail :

#### **Table 3 : La conception d'une entrée lexicale pour les Vsupp espagnols *hacer***



**hacer** vi 1 iets doen, bezig zijn; — *bien*, er goed aan doen; — *como que*, doen alsof; — *y deshacer*, steeds opnieuw beginnen; alles beslissen; naar eigen inzicht beslissen; — *que hacemos*, iets te doen hebben; net doen of we druk bezig zijn; — *por* —, zomaar wat doen, nutteloos bezig zijn; *a medio* —, halfafgemaakt, onvolledig; *dar que* —, te doen geven, last bezorgen; *dejar* —, zijn gang laten gaan, laten begaan; *por lo que hace a*, wat betreft, voor wat ... betreft; *¡qué le vamos a —!*, niks aan te doen!; *tener que* —, het druk hebben, veel te doen hebben; 2 zijn, bv — *un buen juez*, een goede rechter zijn; — *el quince en la cola*, de vijftiende in de rij zijn; 3 zijn, bv *dos y dos hacen cuatro*, twee en twee is vier; 4 zich gedragen als, doen, als in — *el tonto*, zich als een dwaas gedragen, dom doen; 5 ontlasting hebben, poepen; — *de vientre*, ontlasting hebben, poepen; 6 — *para* (plus inf), — *por* (plus inf), proberen om, zijn best doen om; 7 in — *bien (con)*, goed passen (bij); — *mal (con)*, slecht passen (bij), slecht staan (bij); 8 in — *bien* (plus ger), — *bien en* (plus inf), er goed aan doen om; — *mal* (plus ger), — *mal en* (plus inf), er verkeerd aan doen om; 9 kunnen schelen, uitmaken, bv *no me hace*, het kan me niet schelen, het maakt me niet uit; *¿te hace que vayamos al cine?*, zullen we naar de bios gaan?; 10 — *a*, geschikt zijn voor, passen

op; vt 1 maken, scheppen; 2 — (plus adj), maken, bv — *feliz*, gelukkig maken; 3 maken tot, bv *la hizo su mujer*, hij maakte haar tot zijn vrouw; 4 — *de*, maken van, bv — *camisas de la tela*, hemden v d stof maken; 5 doen, verrichten, in orde maken, klaarmaken, bereiden (maaltijd), pakken (koffer), stellen (vraag), knopen (stropdas), geven (opdracht, speech), afleggen (bezoek), doen (boodschap), opmaken (bed), doen (haren, nagels); 6 kunnen bevatten (inhoud); 7 (m verschillende werkwoorden vertaald) bv — *camino*, de weg vrij maken; — *mala cara*, een lelijk gezicht trekken; — *la comunión*, ter communie gaan; — *un desprecio*, een vernedering aandoen; — *buen efecto*, een goede uitwerking hebben; — *hambre*, honger opwekken; — *rumbo*, koers zetten; 8 (in combinatie m e lijdend voorwerp, als één werkwoord vertaald) bv — *burla*, bespotten; — *caricias*, liefkozen, strelen; — *la guerra*, bestrijden, tegenwerken; — *mimos*, verwennen; — *mención*, vermelden; 9 — (plus inf), laten, doen, bv — *entrar*, laten binnenkomen; — *llegar*, doen toekomen; 10 — *que* (plus subj), laten, er voor zorgen dat, als in *haz que entre*, laat hem binnenkomen, zorg ervoor dat hij binnenkomt; *hizo que no entrase nadie*, hij liet niemand binnenkomen, hij zorgde ervoor dat er niemand binnenkwam; 11 (in uitdrukkingen v weersgesteldheid) zijn, bv *hace frío*, het is koud; *hace una primavera fría*, het is een koude lente; *hace sol*, het is zonnig, de zon schijnt; *hace viento*, het waait; 12 (alleen i d 3e persoon) geleden, als in *hace un año*, een jaar geleden; *hacia dos años que no nos habíamos visto*, het was al twee jaar geleden dat we elkaar gezien hadden; 13 veronderstellen, denken dat, houden voor, als in *yo te hacía en París*, ik dacht dat je in Parijs was; *yo le hacía más listo que todo eso*, ik dacht dat hij wijzer was; 14 — *con*, — *de*, voorzien van, verschaffen; 15 in —*la*, iets uithalen, iets flikken, een loer draaien, een hak zetten; *¡ya la has hecho!*, daar heb je het al!; 16 laten wennen, gewennen; *hacerse*, vr 1 zich vormen, ontstaan; 2 zich ontwikkelen, groeien, rijpen; 3 gaar worden; 4 gebeuren; 5 worden, bv — *abogado*, advocaat worden; — *de noche*, nacht worden; — *viejo*, oud worden; 6 zich voordoen als, doen alsof, als in — *el tonto*, zich van den domme houden; — *de nuevas*, doen alsof je iets voor het eerst hoort; 7 zich voordoen als, uithangen, als in — *el gracioso*, de grapjas uithangen; — *el interesante*, interessant doen, zich aanstellen; 8 doen, als in *¿qué te haces?*, wat doe je?, wat voer je uit?; *no saber qué* —, niet weten wat te doen; 9 het in zijn broek doen, in zijn broek poepen; — *encima*, het in zijn broek doen, in zijn broek poepen; 10 lijken, als in *se me hace que*, het lijkt me dat; 11 — (plus adj), vallen, bv — *difícil*, moeilijk vallen; 12 zich vormen, zich maken; — *una idea*, zich een idee vormen; — *ilusiones*, zich illusies maken; 13 — *a*, wennen aan, bv — *al frío*, wennen a d kou; 14 in — *allá*, opzij gaan, opschuiven; — *atrás*, naar achteren gaan, achteruit gaan; — *a un lado*, opzij gaan; 15 — *con*, te pakken krijgen, bemachtigen; 16 — *con*, voor zich winnen, voor zich innemen; 17 — *de*,

Source : Schroten, J. Spaans-Nederlands woordenboek (1989)

Premièremment, Schroten distingue entre verbes intransitifs (vi), verbes transitifs (vt) et verbes réflexifs (vr) dans la langue-source. Les vi du verbe *hacer* ont été catégorisés dans 15 sous-catégories qui ont été définies par traduction du Vsupp (schémas 1 à 6) ou par une description :

19. *Verschillende werkwoorden vertaald (7), in combinatie m e lijdend voorwerp als één werkwoord vertaald (8), +inf (9), +subj(10), in uitdrukkingen van weersgesteldheid (11), alleen i d 3<sup>e</sup> persoon (12) ou expressions idiomatiques (13-16)*

Les schémas 1 à 4 se caractérisent par leur traduction en langue-cible:

20. *Maken/scheppen, ADJ maken, maken tot, maken van*

A partir du schéma 5, d'autres traductions sont introduits: *doen/ verrichten/ in orde maken, kunnen bevatten*. Sous 5, avec les traductions *doen, verrichten, in orde maken* se trouvent les constructions Vsupp avec *hacer* qui font obstacle à une seule traduction de base. Pour donner une traduction en contexte, les Npréd en néerlandais ont été ajoutés :

21. *Bereiden (maaltijd), doen (boodschap)*

Sous 7, les verbes sont affichés différemment, par leur Npréd dans la langue-source et par une traduction de toute la construction Vsupp. Sous 8, une méthode d'affichage clair met ensemble les traductions qui prennent ensemble le verbe et le COD en français et le remplace par un seul verbe transitif en néerlandais:

22. *Hacer burla - bespotten, hacer caricias - liefkozen*

Le même auteur traite le problème des verbes supports de point de vue contrastif et lexicographique dans deux articles (Schroten, 1999b et 2002)<sup>35</sup>. Les idées présentées contribuent à la discussion sur les critères qui peuvent être mis en avant pour un affichage lexicographique qui prête plus d'attention aux constructions Vsupp. Dans le premier article, 'Equivalence and mismatch of semantic features', Schroten (1999b) signale que les dictionnaires monolingues traitent les Vsupp de manière divergente et non-systématique par rapport aux dictionnaires bilingues. Il se demande pourquoi on n'a pas encore élaboré une analyse des constructions Vsupp pertinente pour la lexicographie bilingue. Dans l'article 'Light verb constructions in bilingual dictionaries' (2002), Schroten aborde la question si les constructions Vsupp doivent être interprétées de manière compositionnelle, mais signale que cette question n'est pas pertinente pour les lexicographes bilingues : l'équivalence de traduction (*translational equivalence*) de la construction Vsupp dans la langue-cible peut être optimale, et ne doit pas être nécessairement totale. (Schroten, 2002 : 85).

La notion d'*équivalence* a été développée par l'anthropologue, mathématicien et traducteur de la bible, E. Nida. Dans son article, 'The science of translating' (1964), il explique que la l'équivalence partielle ou formelle se présente au moment où il est question d'équivalence formelle quand les constructions dans deux langues ne se ressemblent pas.<sup>36</sup> Il est seulement possible de traduire correctement la signification et pas la forme syntaxique, quand la langue-cible ne fournit pas les moyens lexicaux et syntaxiques semblables à la langue-source. Comme dans l'exemple des traductions intransitives du Vsupp *hacer*, l'équivalence formelle se présente comme l'équivalence de traduction optimale. La question se pose si on pourrait également suggérer que chaque construction Vsupp n'a qu'une seule traduction optimale. Je retournerai à la notion d'*équivalence* dans le chapitre 5.2.

L'absence d'une possibilité de choix pour verbe léger dans la langue cible pose également des problèmes d'affichage. Schroten prend l'exemple des constructions actives et passives où l'espagnol ne connaît d'équivalent passif pour le couple de Vsupp *geven/krijgen : dar/\*recibir*. Le verbe *recibir* ne fonctionne pas comme verbe léger en espagnol. Il constate que les particularités des verbes légers ne sont pas insérées de manière structurelle et que 'les lexicographes dans les dictionnaires bilingues présentent souvent les collocations avec un Vsupp comme des équivalences totales, même si les Npréd se combine avec plusieurs verbes.' (Schroten, 1999c : 85) Les dictionnaires monolingues sont critiqués de la même manière : les constructions Vsupp sont également traitées de façon inconséquente.

---

<sup>35</sup> Schroten, J. 'Equivalence and mismatch of semantic features' *Evidence for linguistic relativity*. (1999b) Eds. Niemeier, S. and Dirven, R. Amsterdam/Philadelphia: John Benjamins, p. 29-51.

Schroten, J. 'Light verb constructions in bilingual dictionaries' *De la lexicologie à la lexicographie*, Melka & Auguste (eds), UIL-OTS juin 2002, p.83-92

<sup>36</sup> Nida, E. 'Towards a Science of Translating: with special reference to principles and procedures involved in bible translating.' Leiden: E.J. Brill.

### 3.3 Méthodologie et méthodes de recherche dans les corpus

#### 3.3.1. Méthodologie

Il relève des sections 3.1 et 3.2, qu'il est difficile à évaluer dans quelle mesure l'analyse de corpus a été intégrée dans Van Dale 2010. L'objectif de notre analyse est de comparer les résultats des recherches de corpus et les données dans le dictionnaire bilingue Van Dale français-néerlandais en ce qui concerne les collocations qui se combinent avec le Vsupp *commettre* et afin d'améliorer l'entrée lexicale dans un dictionnaire français-néerlandais.

Notre analyse du verbe *commettre* comprend 3 étapes. Premièrement, nous analyserons la structure et la distribution lexicale de collocations qui contiennent le Vsupp *commettre*. A l'aide d'une série de tests (voir la section 1.4), nous déterminons d'abord quels sont les emplois légers du verbe et quelles constructions syntaxiques ne sont pas prises en compte. Du fait que le dictionnaire électronique offre plus de possibilités de recherche, toutes les occurrences de *commettre* dans le dictionnaire seront inventoriées comme supplément aux données de l'entrée lexicale du verbe. Les occurrences issues de Van Dale sont analysées dans la section 3.5.

Dans le chapitre 4, l'analyse des résultats des corpus monolingues constitue la deuxième étape. Après avoir établi une comparaison entre les données de Van Dale et les occurrences dans Frantext, deux autres corpus monolingues sont introduits, à savoir un inventaire des occurrences et fréquences d'un corpus issu du corpus Europarl et un corpus de textes issu de la base de données LexisNexis. Le corpus Europarl est décrit dans le chapitre 3.4.1. Dans l'annexe 2, les résultats d'une recherche quantitative après les collocations du Vsupp *commettre* dans trois corpus monolingues français sont affichés par corpus : les corpus issus de Frantext, de LexisNexis et de la partie français du corpus EU.

Dans la troisième étape, nous analyserons les versions néerlandaises des collocations françaises trouvées avec le Vsupp *commettre*. Cette étape se concentrera surtout sur l'analyse qualitative afin de compléter les données sur le plan de la sémantique lexicale. Ensuite, les constructions Vsupp sont catégorisées et hiérarchisées selon leur pertinence dans une nouvelle entrée lexicographique de *commettre*. Nous n'adoptons partiellement la théorie des normes et exploitations de Hanks(2004) afin d'attribuer des statuts aux collocations à l'aide des résultats qui relèvent des fréquences des collocations. Deux conceptions pour une nouvelle entrée lexicale sont présentées pour un dictionnaire français-néerlandais, dont les éléments les plus pertinents seront utilisés pour la conception finale d'une entrée du Vsupp *commettre*.

#### 3.3.2. Méthodes de recherche dans les corpus

La première étape constitue une recherche après la manière dont les informations lexicales du verbe *commettre* sont distribuées et affichées dans le dictionnaire Van Dale. Ainsi, il est possible d'analyser la mise à jour de ce dictionnaire : a-t-il suivi la tendance dans la linguistique en utilisant de nouvelles approches numériques afin d'améliorer la distribution et l'affichage, en particulier dans l'entrée lexicale des Vsupp? Bien que la langue française soit prise comme point de départ, le domaine de recherche bilingue et l'encadrement dans la traductologie exigent également l'incorporation des données du dictionnaire néerlandais-français afin de déterminer la distribution **directionnelle** : l'affichage français-néerlandais, diffère-t-il de l'affichage néerlandais-français? De cette comparaison on peut déterminer également le public visé : quel type d'utilisateur est visé dans la conception des entrées lexicographiques ? Pour la deuxième étape, nous utiliserons des séquences de recherche semblables afin de former les corpus constituant des collocations de type [forme Vsupp *commettre*] + Npréd. Les Vsupp peut prendre des compléments mais ils sont seulement pris en compte si le complément change la sémantique lexicale de la collocation du schéma existant. Dans la troisième étape, l'ensemble de données issues de Van Dale servira lui-même d'un ensemble de données et donc pas seulement comme point de départ pour la conception lexicale.

Bien que toutes les conjugaisons du verbe *commettre* soient valables, plusieurs formes ont été supprimées dans les séquences de recherche afin de rendre les résultats plus adéquats. Le verbe *commettre* connaît 38 graphies différentes, les homonymes comptés comme une seule graphie. Les graphies ci-dessus ont été supprimés pour de raisons différents. *Commissions* a posé des problèmes en particulier dans le vocabulaire des textes de l'Union Européenne comme les textes du corpus EU. L'homonyme du nom commun *commission* au pluriel est très fréquent et par conséquent supprimé de la séquence de recherche dans tous les corpus EU. *Commis* a son homonyme dans le nom commun *commis*, dans le sens de clerc ou agent. Les participes passés *commis*, *commise*, *commises* ont également une fréquence haute mais ne montrent pas de collocation équivalent de type [forme Vsupp commettre] + Npréd. Pourtant, elles sont souvent traduites par un seul substantif dans les versions néerlandaises des textes. Les textes de nature longs et oraux sont traduits à l'oral et de manière simultanée de telle façon qu'une traduction concise a la priorité. Par conséquent, il se peut que la traduction soit plutôt du type Vsupp + Npréd -> N(préd). Quoiqu'il en soit, les formes ont été supprimées afin de constituer un corpus plus apte aux besoins de traductions adéquates avec un plus grand nombre de traductions du Vsupp en néerlandais. Ci-dessous se trouve une table contenant les corpus utilisées, et les formes qui ont été supprimées.

**Table 4 : graphies supprimées**

corpus	Ft	LN	C558
nombre de résultats	477	424	395
graphies supprimés ?	-	<i>commissions</i>	<i>commissions</i> <i>commis(e)(s)</i>

Il s'avère que pour l'utilisation dans la deuxième étape, l'analyse quantitative, les formes nominales de *commis* ne posent pas de problèmes parce que cet emploi n'est pas fréquent dans le corpus Ft. Cet emploi nominal n'est pas insérée dans la table des résultats mais l'occurrence est un argument en faveur de son insertion éventuelle dans la nouvelle entrée lexicale.

Le nombre de graphies au total dans Frantext relève de 71 graphies (homonymes inclus) qui sont toutes insérées avec la fonction *&ccommètre*, ce qui indique *toutes les graphies de commettre*. Dans le corpus, la dernière délimitation est faite par l'exclusion de déterminants afin d'essayer à supprimer l'emploi nominal de *commis* ou *commissions*. Du fait que l'ancienne version de Frantext propose une recherche avec plusieurs séquences, la recherche est ainsi formulée :

23. Séquence 1 : *&e(g !=D)&ccommètre*

Séquence 2 : *se exclu*, avec les options '1 APRES 2' et 'distance maximale 5'

L'occurrence du verbe *commettre* dans la forme réflexive a été supprimé seulement dans Frantext.

Comme les bases de données EU et LexisNexis ne sont pas annotées, on a procédé par une séquence de recherche booléenne.

Le corpus C558 a été sélectionné du corpus EU dans ParaConc, un concordancier de recherche. ParaConc fournit un espace numérique pour aligner les textes et pour chercher par mots clés (*keywords in context, KWIC*) et par fréquences de mots qui se trouvent à distance d'au maximum 4 mots du pivot, à l'égard de leur place relative dans la phrase. La recherche booléenne s'effectue par les conditions de 'recherche avancée' avec 'regular expression' et une énumération des graphies :

24. [*\bcommet\b | \bcommets\b | \bcommettaient\b | \bcommettais\b | \bcommettait\b | \bcommettant\b | \bcommettent\b | \bcommette\b | \bcommettes\b | \bcommettez\b | \bcommettiez\b | \bcommettions\b | \bcommettons\b | \bcommettra\b | \bcommettrai\b | \bcommettraient\b | \bcommettrais\b | \bcommettrait\b | \bcommettras\b | \bcommettre\b | \bcommettrez\b | \bcommettriez\b | \bcommettrions\b | \bcommettrons\b | \bcommettront\b | \bcommîmes\b | \bcommîrent\b | \bcommisse\b | \bcommissent\b | \bcommisses\b | \bcommissiez\b | \bcommît\b | \bcommît\b | \bcommîtes\b*]

C558 est le résultat du total des graphies de commettre y exclus *commissions* et *commis(e)(s)*. Le corpus C558 est aussi bien utilisé pour l'analyse quantitative après les fréquences des collocations que l'analyse qualitative après les traductions des collocations. Les traductions sont automatiquement alignées et immédiatement affichées en correspondance avec la phrase en français.

Pour LexisNexis, le rapatriement des données est effectué par une sélection des 500 premiers résultats. La recherche booléenne est exécutée dans la période 17/05/2008-17/05/2010 dans tous les journaux français avec la séquence :

25. commet OR commets OR commettaient OR commettais OR commettait OR commettant OR commettent OR commette OR commettes OR commettez OR commettiez OR commettions OR commettons OR commettra OR commettrai OR commettraient OR commettrais OR commettrait OR commettras OR commettre OR commettrez OR commettriez OR commettrions OR commettrons OR commettront OR commîmes OR commirent OR commis OR commise OR commises OR commisse OR commissent OR commisses OR commissiez OR commit OR commît OR commîtes

A l'aide des méthodes de recherche établies dans cette section, on peut déterminer dans la section suivante comment les corpus ont-ils été composés et quel est leur fonction relative aux autres corpus et à l'objectif final lexicographique.

### **3.4. La composition et fonction des corpus**

#### **3.4.1. L'accessibilité et le choix des corpus**

L'objectif général pour l'analyse des trois corpus est de mettre en valeur la contribution des analyses de corpus à la restructuration de l'entrée lexicale pour des verbes supports, à l'exemple du verbe *commettre*. Elle mènera à une réflexion générale sur l'utilisation de l'analyse corpus aux fins lexicographiques.

Les différents types de corpus ont été définis dans le chapitre 2.1.2. Notre analyse de corpus prendra la forme d'une comparaison d'un ensemble de données issu d'un dictionnaire bilingue avec des données issues de plusieurs corpus monolingues, et une recherche dans corpus comparable bilingue. Finalement les résultats bilingues seront comparés aux données de Van Dale.

La difficulté pour exécuter une recherche réside dans l'accessibilité de corpus parallèles, dont les corpus norvégien-anglais et suédois-anglais font l'exemple : au fait des restrictions par copyright, les textes ne sont accessibles qu'à un groupe restreint de chercheurs. (Granger, 2003: 23) A présent, plusieurs corpus parallèles sont en train d'être conçus, notamment dans les pays scandinaves. Les corpus multilingues constitués de textes parlementaires comme le corpus parallèle Europarl ne contient pas de textes-cibles et textes-sources (voir le chapitre 3.4.2) Pour notre recherche, la neutralité des textes issus du corpus parallèle Europarl est une caractéristique positive : ce facteur permet une recherche plus neutre dans le cadre d'une analyse contrastive. Le plus grand désavantage de ce corpus est sa limitation de registre.

Avec une sélection précise de ces corpus, deux objectifs sont servis en même temps : en premier lieu, les corpus sont aptes au sujet de recherche : l'analyse de l'affichage d'un phénomène linguistique qui pose des problèmes dans la lexicographie bilingue ainsi que la de linguistique contrastive. Il se peut que l'affichage des collocations au Vsupp soit déjà problématique dans un dictionnaire monolingue, et par conséquent, je préfère utiliser une méthode de recherche de l'analyse contrastive pour obtenir des données bilingues équivalentes. Deuxièmement, l'ordre de l'analyse s'étend vers un corpus bilingue mais ne s'établit qu'en relation étroite avec le support des corpus monolingues français. Dans cette

section, la fonction et les caractéristiques pour l'utilisation comme base de données sont déterminées pour chaque corpus à part.

Les corpus issus de Frantext (appelé **corpus Ft**) et de LexisNexis (appelé **corpus LN**) servent comme une première vérification des données dans Van Dale français-néerlandais. En outre, ils constituent un ensemble de collocations issues de l'emploi langagier envers l'ensemble de données issu de Van Dale. Comme base de données renommée, *Frantext catégorisé* se caractérise par son accessibilité et est la seule base de données annotée par lemme. Le corpus Ft est issu de Frantext catégorisé, un corpus dynamique auquel chaque mois, il y a de nouvelles mises à jour<sup>37</sup>. Les 319 textes sont issus de la période 1960-1997 et consultés en avril 2010. La plupart des textes est d'origine littéraire, bien que parmi les textes des années 1960, plusieurs textes législatifs ont été intégrés. J'utilise l'ancienne version de Frantext en raison de sa recherche liant plusieurs séquences de recherche.

La base de données LexisNexis ajoute des collocations dans un domaine absent dans Frantext, celui des quotidiens. Dans les options de recherche, toutes les journaux ont été sélectionnés afin d'obtenir des données de plusieurs domaines et registres<sup>38</sup>.

Le corpus issu du **corpus parallèle Europarl**<sup>39</sup> est au centre d'intérêt quand il s'agit de l'aspect traductologique. Il s'agit de textes des actes de conférence en 11 langues de la durée d'une année de la période 1996-2009. Le français, l'italien, l'espagnol, le portugais, l'anglais, le néerlandais, l'allemand, le danois, le suédois, le grec et le finnois ont été alignés à l'anglais dans des corpus bilingues parallèles. Le corpus français-néerlandais qui a été créé<sup>40</sup>, est dénoté à partir de maintenant par **corpus EU**. Le corpus EU contient 29.664.819 mots en français versus 27.250.397 mots en néerlandais dans les textes alignés. Pour la deuxième étape de notre recherche, l'analyse quantitative, la partie française est également utilisée comme il s'agit d'un corpus monolingue, dénotée par **la partie française du corpus EU**. Le **corpus C558** est ainsi issu du corpus EU. Les données français du corpus C558 ont toutes été alignées à leur version néerlandaise et seront soumises à une recherche qualitative dans le chapitre 5.

La recherche d'un grand corpus bilingue nécessite l'utilisation d'un environnement de recherche automatique créé pour les corpus multilingues. Comme concordancer, j'ai utilisé le programme ParaConc.<sup>41</sup> Notre corpus est apte à l'utilisation de ParaConc par l'alignement au niveau de la phrase. Le programme offre de bonnes méthodes de recherche et de catégorisation. Les textes ont été pré-édités par l'alignement.

Un deuxième objectif de l'utilisation des corpus est de découvrir, en exécutant une recherche empirique, comment un corpus peut être constitué en fonction d'une recherche au sein de la traductologie. La réflexion sur le développement de la traductologie contribuera au développement de la recherche empirique ainsi que à l'analyse de corpus encadrée par la linguistique contrastive. Après la présentation de la conception lexicographique dans le chapitre 5.3, des conclusions seront tirées à un niveau plus abstrait.

### 3.4.2. Le statut des textes dans le corpus EU

---

<sup>37</sup> Base textuelle Frantext ATILF (Analyse et traitement informatique de la langue française), Frantext catégorisé. [www.frantext.fr](http://www.frantext.fr)

<sup>38</sup> LexisNexis academic NL <http://academic.lexisnexis.nl>

<sup>39</sup> European Parliament Proceedings parallel corpus, voir: <http://www.statmt.org/europarl/>

<sup>40</sup> Une fois de plus, je remercie Bert le Bruyn d'avoir effectué l'alignement de ces deux langues.

<sup>41</sup> Barlow, M. 'ParaConc and Parallel Corpora in Contrastive and Translation Studies'. P. 9-16, 2008.

D'origine, le corpus parallèle Europarl a été tiré des actes de conférence du Parlement Européen. Comme les 11 langues européennes sont considérées comme équivalentes, aucune langue ne peut être désignée comme langue-source. Par conséquent, je considère les textes comme versions ayant un contenu identique, en définissant 'version' par un texte au statut de ni texte-source et ni texte-cible par rapport aux textes alignés. Les traces de traduction éventuelles sont considérées comme négligeables. Dans cette recherche, il est adopté que supprimer la langue qui lie le français au néerlandais, l'anglais, n'aura pas de conséquences négatives pour la qualité de l'alignement des textes français aux textes néerlandais.

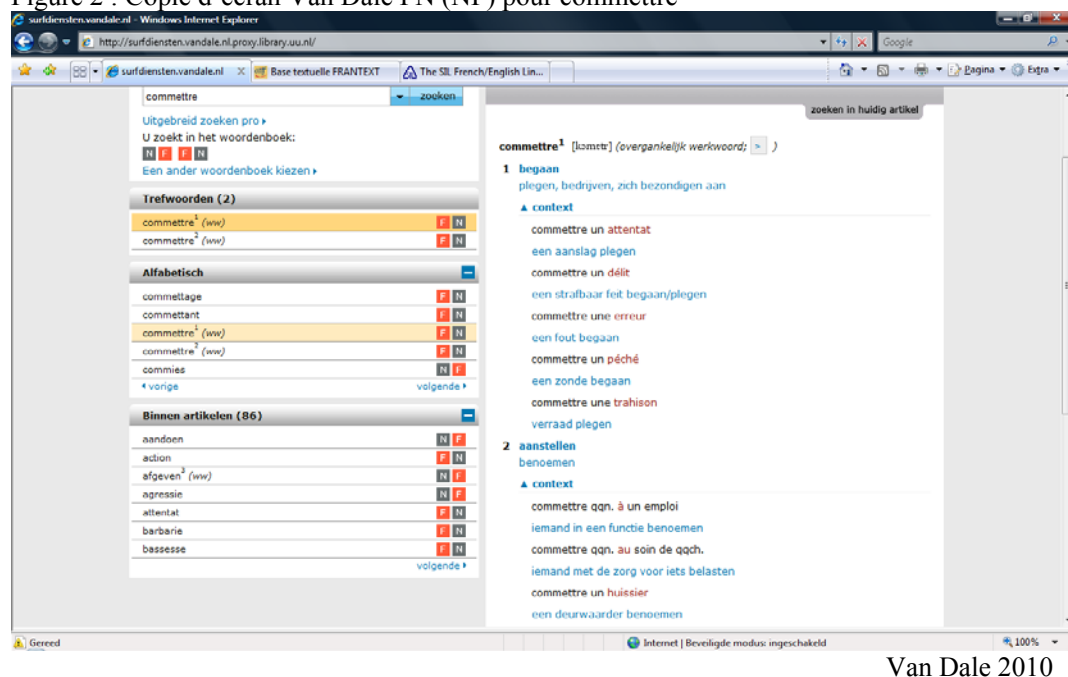
### 3.5 Analyse de l'entrée lexicale d' un Vsupp

L'analyse des données dans le Van Dale, s'effectue premièrement par l'entrée électronique de Van Dale français-néerlandais, et deuxièmement par l'entrée néerlandais-français de *commettre*.

#### 3.5.1. La composition de l'entrée lexicale

Le dictionnaire bilingue et bidirectionnel Van Dale professionnel constitue notre point de départ. Comme la version électronique fournit plus de possibilités de recherche, j'ai adopté la version 2010 électronique.

Figure 2 : Copie d'écran Van Dale FN (NF) pour commettre



Van Dale 2010

La copie d'écran montre les possibilités de recherche : par les différents dictionnaires (N-F ou FN), par une liste de *trefwoorden*. Les deux lexèmes de *commettre* se distinguent par leur forme syntaxique. *Commettre*<sup>1</sup> est le verbe transitif, tandis que *commettre*<sup>2</sup> est le verbe réflexif.

Au milieu se trouve la colonne *alfabetisch*, dans laquelle les langues se trouvent mêlées, et en bas la colonne *binnen artikelen* qui contient également des entrées français et néerlandais. L'affichage électronique permet de chercher en cliquant sur les entrées affichées dans les colonnes. Par contre, dans les colonnes on ne peut naviguer que par *vorige/volgende*.<sup>42</sup> De ce fait, la recherche d'un Npréd

<sup>42</sup> Dans la version électronique 2008, cette colonne était encore fournie d'un navigateur.



*attentat* est facilement à trouver, bien que *zonde* se trouve en naviguant jusqu'à la fin de la liste. Pour *commettre*, les entrées ne sont pas très nombreuses, mais si on prend le verbe *mettre* comme exemple, Van Dale donne 1857 entrées dans lesquelles la navigation n'est pas pratique puisque tous les lexèmes, non tenu compte de la langue, se trouvent dans la même colonne de manière alphabétique. En outre, les mots clés affichés en rouge marquent souvent le Npréd, mais ils ne sont pas fournis d'un lien hypertexte.

Ensuite, le verbe est divisé par différents emplois prototypiques, pour lesquels j'adopte la notion de schémas (voir la section 3.2.2). Il y a la possibilité d'exemplifier les schémas pour les emplois en contexte ou pour des idiomes uniquement.

### 3.5.2. L'emploi léger du verbe commettre

La première partie de cette section est consacrée au verbe *commettre* comme Vsupp dans la littérature. La deuxième partie comprend une délimitation du choix du verbe *commettre* pour notre analyse. La troisième partie comprend une analyse de données à l'aide d'un ensemble sélectionné de tests du chapitre 1.4.

Sous direction de G. Gross, une édition de la série *Linguistica Investigationis* est entièrement consacrée à l'analyse syntaxique et sémantique des verbes supports.<sup>43</sup> Bien que l'aspect de traduction soit surtout analysé sous l'analyse contrastive, l'analyse du Vsupp contribue à notre compréhension de l'emploi du verbe. Dans l'article 'Les verbes supports sans peine', Mel'cuk distingue une catégorie intermédiaire entre l'emploi principal ou entier et l'emploi léger.<sup>44</sup> La caractéristique de 'sémantique vide' qui est attribuée aux verbes comme *commettre* est discutée:

*Un Vsupp peut avoir un sens même dans la collocation en cause. Ainsi, dans commettre un assassinat, subir une sanction ou écoper une amende, les Vsupp commettre, subir et écoper conservent beaucoup de sens. On peut dire : Qu'est-ce qu'il a commis ? ou De combien a-t-il écoperé ?, où commettre et écoper sont sélectionnés d'après leur signifié (les expressions citées ne sont pas des collocations). Cependant, dans les collocations ci-dessus, ces Vsupp ne sont pas choisis pour leur sens par le mécanisme de lexicalisation - ils sont introduits par les règles de la syntaxe exactement pour 'supporter' les lexèmes assassinat, sanction et amende, choisis au préalable. (Mel'cuk : 205)*

Le choix du verbe n'est pas basé sur l'entrée dans Van Dale mais sur les remarques ci-dessus et les recherches d'essai tirées dans la base de données Frantext. Le verbe *commettre* est un verbe d'un type/token ratio modéré et fournira un nombre faisable de données. Je m'attends à un nombre modéré de **schémas catégoriels** au niveau sémantique et compositionnelle. L'entrée lexicale n'est pas prise comme point de départ afin de se concentrer sur le Van Dale, mais afin d'encadrer notre analyse de corpus dans la branche appliquée de la traductologie. Avec l'analyse critique de l'affichage d'un seul verbe, l'objectif est de développer un regard critique mais général envers les méthodes d'affichage dans un dictionnaire bilingue français-néerlandais.

Suivant Hanks et al (2006), la conception d'une entrée lexicale adaptée aux Vsupp commence par la constatation que le DUW, le *Deutsches Universalwörterbuch*, doit être considéré comme un dictionnaire pré-corpus. Dans le chapitre 3.2.2., on a conclu que l'analyse de corpus dans Hanks et al. sert ici à améliorer les méthodes d'affichage dans un dictionnaire monolingue la combinatoire lexicale et l'emploi du verbe en contexte. La variation syntaxique des collocations reçoit moins d'attention. La

---

<sup>43</sup> Verbes supports, *Linguistica Investigationis*, 27 :2 John Benjamins Publishing Company (2004)

<sup>44</sup> Mel'cuk, I. 'Les verbes supports sans peine', *Linguistica Investigationis*, 27 :2 John Benjamins Publishing Company (2004) p. 203-218.

première étape de notre analyse consiste à établir une hypothèse semblable au dictionnaire bilingue Van Dale dans le chapitre suivant.

Par rapport à l'analyse de Hanks et al, dans notre recherche la priorité est donnée à deux aspects interdépendants l'un de l'autre : l'affichage et le public visé. L'affichage des possibilités de collocation combine les caractéristiques lexicales et sémantiques du lexème. Notre public visé ne consiste pas seulement d'utilisateurs d'origine néerlandaise mais aussi d'utilisateurs allophones, des apprenants de néerlandais langue étrangère. Dans le cadre des fins traductologiques, je voudrais délimiter encore plus le public idéal vers un public de traducteurs. Ainsi, l'intérêt à l'information combinatoire recevra plus d'attention par rapport à la compréhension de lexèmes séparés. L'affichage avant tout doit servir le traducteur qui n'a pas nécessairement le néerlandais comme langue maternelle.

Comme notre sujet de recherche constitue les constructions V<sub>supp</sub> dans deux langues, l'entrée lexicale de *commettre* servira également de base de données statique. Cette base de données servira comme support à l'analyse des occurrences du V<sub>supp</sub> en collocation. Sa composition est analysée dans le chapitre 5.3.3.

Dans cette troisième section, je reprendrai les tests sélectionnés dans le chapitre 1.4 pour déterminer la légèreté de l'emploi de *commettre* un V<sub>supp</sub>.

La pronominalisation s'effectue de deux façons : externe à la construction, par une phrase avec que :

26. OK Il commet une incorrection que personne ne comprend.

L'autre possibilité était la pronominalisation interne :

27. OK Il regrettait la faute qu'il avait commise.

La condition de la variabilité du N<sub>préd</sub> est remplie :

28. OK Il a commis un meurtre, le meurtre le plus sensationnel, les assassinats les plus spectaculaires du 20<sup>e</sup> siècle. (Langer : 176)

Le même principe vaut pour la négation, les pronoms possessifs et l'emploi d'attributs, et la coordination de deux N<sub>préd</sub> :

29. OK Il n'a commis aucun crime

30. OK Il a commis son assassinat

31. OK Jean commettait une grave incorrection

32. OK Anne a commis deux hold-up et un cambriolage

Le N<sub>préd</sub> permet la passivation comme avec le verbe *commettre*, il s'agit également d'un N<sub>préd</sub> en position de objet direct (Langer : 177)

33. OK. Un crime a été commis

Le test Zeugma n'est pas bien délimité par rapport au test de la coordination de deux N<sub>préd</sub> comme dans l'exemple 32 : la subordination de deux N<sub>préd</sub> comme *deux hold-up et un cambriolage* se passe bien à condition que la même signification de *commettre* soit dénotée par le V<sub>supp</sub>. La subordination aléatoire de deux N<sub>préd</sub> qui vont avec le verbe *commettre* résulte dans un ensemble non-univoque :

34. a. \*Il commet un péché et des imprudences.

b. \*Il commettait des violations et un oubli.

c. \*Il commettait un impair et un vol.

d. Il commet deux fautes et l'imprudence de demander après l'âge de l'interlocuteur.

e. Il commet une bévue et une autre sottise

f. Il a commis des violences et des vols.

g. Il a commis un oubli et une inconvenance.

Personnellement, j'accepte les coordinations dans les exemples 12 à 15 et je rejette les exemples 9 à 11. Le test Zeugma pourrait être un test valable pour la sous-catégorisation des constructions V<sub>supp</sub> : à l'intérieur des constructions V<sub>supp</sub> sélectionnées par d'autres tests en raison des résultats non-équivoques, ce test peut être utilisé pour déterminer les sous-catégories. Bien qu'il ne soit pas sûr si la catégorisation établit la même catégorisation des normes, je reprendrai le test Zeugma dans le chapitre 5 pour la sous-catégorisation.

Finalement, le test de substitution du Vsupp par un autre Vsupp s'applique seulement à un certain nombre de constructions Vsupp. Les catégories formées par ce test ne correspondent pas aux catégories formées par le test Zeugma. L'utilisation de ce test pour la sous-catégorisation fournira une autre catégorisation, mettant probablement plus d'attention à la variation verbale.

### 3.5.3. Les occurrences dans Van Dale F-N et N-F

La première étape de recherche dans Van Dale constitue l'analyse de l'entrée du verbe *commettre* dans le dictionnaire français-néerlandais, affichée dans l'annexe 1.a.

Dans l'entrée de *commettre*<sup>1</sup>, six emplois prototypiques sont mis en avant dont deux emplois 'archaïques', à savoir 'toevertrouwen' en 'in gevaar brengen' un emploi spécifique 'zich bezondigen aan het schrijven van' avec l'exemple *commettre sa réputation - zijn goede naam in gevaar brengen* qui pourrait mieux être classifié sous le paradigme 'in gevaar brengen'. Par conséquent, les emplois 5 et 6 pourraient être mis ensemble en suivant le principe de rassemblement d'occurrences dans les schémas non-récurrents. L'emploi de *aanstellen/benoemen* est suivi de six exemples qui montrent surtout la variation syntaxique autour deux paradigmes de 'iemand benoemen' et la combinaison fixe *commis d'office*. Il ne sera pas soumis à une analyse qualitative mais leur occurrence dans les corpus monolingues sera prise en compte.

En vue de la comparaison avec les résultats des corpus, le domaine de la recherche est limité au premier emploi avec la traduction de base 'begaan, plegen, bedrijven, zich bezondigen aan'. Les cinq exemples ne reflètent pas toutes les traductions proposées (trois fois *plegen* et trois fois *begaan*). Les six exemples ont tous une forme transitive suivie par la traduction avec un verbe transitif. Les exemples n'affichent pas d'autres variations de traduction sauf *commettre un délit - een strafbaar feit begaan/plegen*. Les exemples sont affichés dans l'ordre alphabétique. Au niveau de l'affichage, les Npréd insérées ne montrent que 6 exemples de l'emploi du Vsupp. Pour l'utilisateur, il peut être difficile de déterminer quelles traductions parmi ces quatre propositions se combinent avec le Npréd recherché. L'insertion d'au minimum toutes les traductions du verbe pourrait apporter une vue d'ensemble de la variation de la traduction en néerlandais, en particulier à l'utilisateur qui n'a pas le néerlandais de langue maternelle. Un affichage électronique fourni d'hyperliens fait partie des améliorations possibles sans que l'entrée devienne un affichage de l'ampleur monolingue du type *leisten*: je veux éviter que l'ampleur total l'entrée lexicale bilingue devienne plus long que la proposition allemande pour un dictionnaire monolingue.

Dans la deuxième étape toutes les occurrences dans la colonne 'binnen artikelen' ont été rassemblées dans l'annexe 1.b. afin de compléter la table 5 pour ce qui en est les données de Van Dale. Premièrement, la copie d'écran dans 3.5.1. montre des mots marqués en rouge qui ne servent pas comme des hyperliens mais comme une méthode de recherche peu développée. Dans la table 5 se trouve un inventaire des Npréd qui occurrent avec *commettre* ainsi que les traductions des Npréd et les caractéristiques spécifiques d'affichage de la collocation.

**Table 5 : les collocations avec le Vsupp commettre dans Van Dale**

occurrence	source	alternance Vsupp?	Remarques	Traduction
acte(s)	vD NF		pluriel ex. +reprehensible	'zondigen' Een slechte daad begaan, een misstap doen
action	vD FN		ex. + mauvais	een lage daad begaan
adultère	vD NF		bij laag +vile	'echtbreken'
adultère	vD NF			Ontrouw
aggression	vD NF		ex. +acte de	agressie plegen
assassinat	vD NF		Pluriel	'moorden'

attentat	vD FN		contre quelqu'un	Een aanslag op iem. plegen
barbarie	vD FN		ex. +acte de	wreedheid, onmenselijke daad
bassesse	vD FN			Een gemene daad, laagheid begaan
bévue	vD FN			Een flater slaan
bévue	vD NF			'blunderen'
confusion	vD FN		ex. + grossière	'de zaken ernstig door elkaar halen'
coup d'état	vD NF			Staatsgreep
crime	vD NF			'misdrijven'
daderschap	vD NF		idioom 'amener qqn à commettre un crime	'iem tot het daderschap leiden'
débauche	vD NF		ex. +acte de	ontucht plegen'
délit(s)	vD NF			'misdrijven'
dievenpad	vD NF		idioom 'se mettre à commettre des vols'	'het dievenpad opgaan' Stommiteiten uithalen,
énormité	vD FN		pluriel	'flaters slaan'
erreur	vD FN	faire		Een vergissing maken
étourdie	vD FN			iets doms doen
faute	vD NF		ex. +grave	Een misstap begaan
faute	vD NF			'misdrijven'
faute(s)	vD FN	faire		Een overtrading begaan, 'zondigen'
félonie	vD FN		Ex. + acte de	Trouwbreuk plegen
forfait	vD FN			Een misdad, misdrijf begaan
fraude(s)	vD FN		ex. + à un examen	'fraude plegen bij een examen' Een stommiteit begaan, 'een bok schieten', 'blunderen'
gaffe	vD FN	faire		
geweldig	vD NF		comme adj	geweldige stommiteiten uithalen lomp optreden, zich onhebbelijk gedragen
goujaterie	vD FN			
horreurs	vD FN		pluriel	afschuwelijke dingen doen 'een flater slaan, een bok schieten
impair	vD FN			
impolitesse	vD FN			een onbeleefdheid begaan 'een onvoorzichtigheid/ onverstandigheid begaan
imprudence	vD FN			
impudicité	vD FN		[formeel] pl	Onkuisheden begaan
inceste	vD FN			Incest plegen
inconvenance	vD FN			een onbeleefdheid begaan
incorrection	vD FN		ex.+ grave	'heel onbeleefd/onbehoorlijk zijn'
indélicatesse	vD FN			bedrog plegen, oneerlijk handelen
indiscrétion	vD FN			'uit de school klappen'
infraction	vD FN			een strafbaar feit plegen
infraction	vD NF			Overtreding
injustice	vD NF		ex. +envers qqn	'iem onrecht aandoen'
irrégularité(s)	vD FN			'in strijd met de wet handelen'
laag	vD NF		comme adj	een lage daad begaan
larcin	vD FN	faire		(kruimel)diefstal

maladresse(s)	vD FN			een flater begaan
malhonnêté	vD FN			iets oneerlijks doen, bedrog plegen
méfait	vD NF			'misdrijven'
meurtre(s)	vD FN			een moord plegen/begaan
misbruiken	vD NF	abuser de	ex +abus de pouvoir	commettre un abus de pouvoir, abuser de son pouvoir
oubli	vD FN	faire		iets verzuimen (te doen)
péché	vD FN			een zonde begaan
péché	vD NF		ex. +article défini	'zonde bedrijven'
plagiat	vD NF		Ex. art. indéfini	plagiaat bedrijven'
sabotage	vD NF		ex. +acte de	'sabotage plegen'
school	vD NF		une indiscretion'	'uit de school klappen'
sottise	vD FN			een stommititeit uithalen, een dwaasheid begaan
trahison	vD FN			verraad plegen
turpitude(s)	vD FN		[formeel of ironisch]	'zich schaamteloos gedragen'
uithalen	vD NF		ex +des énormités	stommititeiten uithalen
violence	vD FN		pluriel ex. + sur qqn	iemand geweld aandoen
vol	vD NF		ex +avec infraction	'diefstal met braak plegen'
zonde	vD NF	pécher		zonde began

La première colonne contient à part des noms prédicatifs également des entrées verbales et mots clés néerlandais. Il s'agit de verbes pour lesquels la traduction par *commettre* est proposée. Les mots clés néerlandais dénotent des expressions idiomatiques insérées comme dans Van Dale néerlandais-français sous les entrées suivantes :

*daderschap (het)*

¶ *Idioom iem. tot het daderschap leiden - amener qqn. à commettre un crime*

*dievenpad (het)*

¶ *Idioom het dievenpad opgaan - se mettre à voler, à commettre des vols*

*school (de)*

¶ *Idioom uit de school klappen jaser, bavarder- commettre une indiscretion*

Ils ne sont pas considérés comme des Npréd mais je ne propose pas de les supprimer. En outre, la traduction *uit de school klappen* constitue une variation pour la collocation *commettre une indiscretion*, qui peut être insérée au moment où *indiscretion* est insérée dans l'entrée de *commettre*. Je considère tous les noms prédicatifs comme des Npréd qui occurrent en collocation avec le Vsupp *commettre* et qui remplissent les critères posés dans le chapitre 1.4.

En général, moins d'attention est prêté aux alternances sémantiques que pour les alternances syntaxiques : Parmi les entrées des V et N qui occurrent dans les constructions Vsupp, il se trouve peu de paradigmes. Ainsi '*misdrijven*' est marqué par la caractéristique 'transitif' :

35. *Misdrijven - commettre (un méfait, un délit, un crime, une faute)*

Dans la discussion des résultats, on a discuté la transitivité du verbe '*misdrijven*' en néerlandais : il s'avère qu'il est difficile à substituer *wat* ou *iets* par un Npréd concret :

36. OK *Wat heeft hij misdreven ? - Hij heeft iets ?misdreven/OKmisdgaan.*

Dans Van Dale néerlandais-français, le verbe est inséré comme un verbe transitif.

Tout comme dans l'entrée lexicale de *commettre*, le nombre de traductions est limité à un ou deux exemples. En outre, les emplois en direction français-néerlandais proposent plus souvent une traduction transitive qu'intransitive. Dans les entrées lexicales de *faute*, *gaffe/bévue* et *méfait/délit* un verbe intransitif correspondant est proposé comme traduction en néerlandais : *zondigen*, *blunderen* et *misdrijven*. Ces variations ne sont pas insérées dans l'entrée du verbe. En outre, l'emploi de *commettre*

est alterné par le Vsupp *faire* seulement. L'affichage des possibilités de construction Vsupp ainsi que leur traduction est donc limité au niveau sémantique, à une approche unidirectionnel et à une tendance d'afficher les possibilités de collocation en majorité dans les entrées de Npréd en question.

De la table 5, il s'avère que l'ensemble d'informations contribue à l'élargissement de l'ensemble des collocations du français. Après la justification dans l'introduction de plus nouvelles éditions de Van Dale dictionnaires bilingues, les alternances lexicales sont affichées dans les entrées des Npréd plutôt que dans les entrées des verbes. Les constatations qui suivent ci-dessous sont exemplaires pour le traitement des constructions Vsupp dans les cas particuliers.

Un des mots clés qui est marqué d'une fréquence haute dans Van Dale est *acte*. L'élément disparaît dans les collocations néerlandaises. En outre, l'article sur *acte* dans Van Dale n'a inséré aucun exemple avec l'emploi de 'commettre un acte de [N]' Pour toutes les collocations, il s'agit d'actes négatifs. Dans une recherche en dehors les corpus utilisées pour *commettre une bonne action*, il s'avère qu'il existe bien une connotation positive mais qu'il est peu fréquent<sup>45</sup>. Contraire à ce que j'ai supposé dans le chapitre 3.5.2, le verbe *commettre* peut combiner avec des compléments dans certaines constructions Vsupp:

37. Commettre un attentat sur quelqu'un / commettre des violences sur quelqu'un

Le test du déplacement des compléments vérifie si le complément fait partie de la construction Vsupp:

38. a. C'est contre Luc que Max commet un crime. (Langer: 186)

b. ?C'est pour un travail que Max a été commis.

c. C'est contre le parti socialiste que l'armée a commis un attentat.

d. C'est sur les filles que le l'accusé a commis des violences.

La collocation dans l'exemple 2 est marquée par le sens de *embaucher* et n'est pas considérée comme une construction Vsupp. Il faudra trouver des moyens afin d'intégrer des remarques syntaxiques. En outre, l'alternance lexicale dans les traductions proposées doit être bien délimitée par rapport aux combinaisons permises en néerlandais:

39. Commettre une bévue

a.= een blunder begaan/blunderen

b.= een flater begaan /\*flateren

L'analyse de Van Dale mène aux conclusions suivantes : Le dictionnaire a employé une façon d'affichage sélective pour le nombre de collocations possibles ainsi que sur le plan de la direction de traduction. Le traitement de *misdriven* comme verbe transitif est significatif pour la difficulté de rassembler des occurrences dans des paradigmes ainsi que d'afficher une traduction prototypique. En outre, l'entrée lexicale actuelle est trop restreinte au niveau sémantique pour satisfaire aux besoins d'un traducteur qui n'a pas le néerlandais comme langue maternelle. La description est surtout visée sur la compréhension chez le traducteur d'origine néerlandais de la signification et l'occurrence du verbe seul sans rendre compte de schémas fréquents avec par exemple *action* et *acte*. Il s'avère que la distribution syntaxique a eu la priorité. En outre, l'affichage par Npréd n'est pas toujours pratique comme montre le mode de recherche dans 'binnen artikelen'. L'élargissement de cette entrée constituera l'insertion de plusieurs alternatives pour le verbe commettre et plus de proposition de traduction avec des verbes intransitives, à condition que les limites de la traduction par verbe intransitif soient bien tirées. Je m'attends à une contribution considérable de collocations possibles dans les corpus afin d'insérer plus de propositions pour la traduction. Finalement, l'aspect quantitatif a pour but la modification de l'affichage des possibilités dans une nouvelle entrée lexicale.

---

<sup>45</sup> Une recherche élargie dans Frantext (entre 1900 en 2000) montre 49 occurrences trouvées dont 4 constructions peuvent être décrit comme positives.

## Chapitre 4 : Résultats des corpus monolingues

La fonction des corpus monolingues est définie comme la vérification des collocations affichées dans Van Dale avec le Vsupp *commettre*. J'analyserai des corpus de nature différents : le corpus Ft est un corpus littéraire (4.1), le corpus EU forme un corpus de textes d'affaires (4.2.) et le corpus LN est constitué d'articles de presse (4.2.). L'analyse dans ce chapitre se concentre sur la distribution lexicale et quantitative des Npréd dans ces corpus avec le Vsupp en question. Les trois corpus monolingues sont comparés dans le chapitre 4.3, suivi par une comparaison de l'ensemble des tokens avec l'ensemble des données issues de Van Dale dans le chapitre 4.4. Dans cette dernière section se trouvent également des propositions personnelles au niveau de la traduction de constructions Vsupp du français en néerlandais.

Il s'agit dans tous les corpus de l'ensemble de résultats tirés des recherches qui remplissent les conditions et limitations posées dans le chapitre 1.4. Pour les résultats intéressants qui se trouvent en dehors du corpus, les coordonnées seront données.

### 4.1. Les résultats issus de Frantext

#### 4.1.1. Résultats rapatriés

Le corpus Ft contient 477 résultats. Les formes verbales réflexives de *commettre* ainsi que les occurrences de *commis* utilisées comme substantif ou adjectif ont été supprimées manuellement :

40. Cependant la victoire fut gâchée par les organisateurs qui s'obstinèrent à nommer «Sacré-Coeur» le temple de la Révolution et par les parents qui confisquèrent les chapeaux que nous avions gagnés, sous prétexte qu' on ne **se commet** pas avec la réaction.  
R757/ MORDILLAT.G /VIVE LA SOCIALE/1981 page 95 / 29
41. La loyauté des **commis de l' état** à l' égard de tout gouvernement, comme la compréhension par tout gouvernement que le métier public comporte d' avoir servi les gouvernants qui l' ont précédé et de servir ensuite ceux qui lui succéderont, sont les deux conditions nécessaires pour écarter la tentation du " système des dépouilles " .  
P292/ BELORGEY.G /GOUVERN. ET ADMINISTR. FRANCE/1967 page 241 / -
42. Qui plus est, c' était, non un chimiste, mais un étudiant en droit, qui travaillait comme **commis greffier**.  
P322/ ROUSSEAU.P /HIST. TECHNIQUES ET INVENTIONS/1967 page 299 / -

La question se pose si les occurrences nominales *commissions* et *commis* doivent être insérées par une référence (électronique) dans la nouvelle entrée de *commettre*.

#### Table 6 : Les résultats issus de Frantext par ordre décroissant

<b>Npréd</b>	<b>inséré dans Van Dale?</b>	<b>nombre de résultats</b>
erreur	vD FN	97
faute	vD FN, vD NF	84
crime	vD NF	44
péché	vD NF, vD FN	34
imprudence	vD FN	15
acte	vD NF	11
meurtre	vD FN	9
délit	vD NF	8
infraction	vD NF	8
vol	vD NF vD FN, vD	8
action	NF	7
forfait	vD FN	7
indiscrétion	vD FN	7
injustice	vD NF	7
bêtise		6
assassinat	vD NF	5
impair	vD FN	5
sottise/sotie	vD FN	5
abus		4
bévue	vD FN, vD NF	4
connerie		4
dégât		4
folie		4
larcin	vD FN	4
méfait	vD NF	4
abomination		3
atrocité		3
attentat	vD FN	3
gaffe		3
mensonge		3
trahison	vD FN	3
adultère	vD NF	2
confusion	vD FN	2
contresens		2
cruauté		2
excès		2
geste		2
horreur	vD FN	2
oubli	vD FN	2
pas		2
sacrilège		2
violence	vD FN	2



agression	vD FN	1
bassesse	vD FN	1
cambrilage		1
cercle		1
châtiment		1
circoncision		1
déprédation		1
écrit		1
effraction		1
énergie		1
enlèvement		1
espièglerie		1
étourd(er)ie	vD FN	1
exaction		1
fait		1
fausse note		1
faux pas		1
galoup		1
graffiti		1
imposture		1
inceste	vD FN	1
inconvenance	vD FN	1
indécence		1
indignité		1
infamie		1
irréremédiable		1
lâcheté		1
ladrerie		1
livre		1
lourdeur		1
mal		1
maladresse	vD FN	1
malversation		1
manquement		1
massacre		1
sabotage	vD NF	1
sacrifice		1
sournoiserie		1
stupidité		1
vandalisme		1
vilenie		1
viol		1
barbarie	vD FN	0
coup d'état	vD NF	0
débauche	vD NF	0

énormité	vD FN	0
félonie	vD FN	0
fraude	vD FN	0
goujaterie	vD FN	0
impolitesse	vD FN	0
impudicité	vD FN	0
incorrection	vD FN	0
indélicatesse	vD FN	0
irrégularité	vD FN	0
malhonnêteté	vD FN	0
plagiat	vD NF	0
turpitude	vD FN	0
<b>Total:</b>		<b>477</b>

Dans la table 6, les Npréd issus du corpus Ft et de l'ensemble de données de Van Dale sont insérés, ainsi que la direction de traduction, dans l'ordre décroissant de fréquence. Dans la section 4.1.2. se trouve une comparaison entre Van Dale et Frantext, dans cette présente section les résultats rapatriés de Frantext sont présentés. Bien que plusieurs Npréd se combinent au Vsupp au pluriel, tous les Npréd sont insérés dans la table au singulier. La table 6 contient 49 occurrences issues de Van Dale, auquel 50 Npréd issus du corpus Ft ont été ajoutés. Parmi les nouveaux ajouts, les Npréd *acte*, *infraction* et *imprudence* se trouvent parmi les Npréd qui ont une occurrence de >10 dans le corpus Ft :

43. La cour **committ d'autres imprudences** en s'attachant à l'étiquette qui humiliait le tiers état : on assigna un costume particulier à chacun des ordres

P342/ LEFEBVRE.G /LA REVOLUTION FRANCAISE/1963, page 128 / -

Le Npréd *acte* a deux emplois syntaxiques : la construction Vsupp suivi par un adjectif se qui est formalisé par *commettre un acte* +ADJECTIF, et la construction Vsupp suivi par le complément de et un syntagme nominal. Cette dernière construction est formalisé par *commettre un acte* +DE.

44. Mon père me rétorque violemment que j'ai dû mériter ce traitement, **commettre un acte répréhensible**, une agression, ou plutôt non, je suis trop lâche, une sournoiserie.

S326/ SALVAYRE.L /LA PUISSANCE DES MOUCHES/1995, page 115 / 14

45. Charlotte pérorait: «Gnagnagna, **un acte de... vandalisme** insensé gnagna d'être **commis**, messieurs.

R727/ BAYON /LE LYCEEN/1987, page 40 / HENRI-IV

Les deux constructions sont mises ensemble pour la fréquence absolue.

A moitié des occurrences du corpus Ft, le Npréd *infraction* est marqué également par la présence d'un complément +A :

46. Il y a faute de service si le joueur **commet une infraction à ces règles**, s'il manque la balle en essayant de la frapper, si, avant de toucher le sol la balle ne passe pas au-dessus du filet.

P839/ \*Anonyme /JEUX ET SPORTS/1967, page 1371 / -

47. Les infractions spécialement définies pour la protection des parcs nationaux ainsi que **les infractions commises** dans ces parcs en matière forestière, de chasse et de pêche sont constatées par des agents assermentés, [...]

P763/ JOCARD.L-M /TOURISME ET ACTION DE L'ETAT/1966, page 182

Les Npréd *injustice*, *bêtise* (qui est plus fréquent que le Npréd similaire *bévue*), *abus*, *connerie/folie* et *dégâts* représentent un assez grand nombre (4-7 occurrences) des Npréd insérés dans les résultats.

48. Voilà que la quatrième génération se lance dans **les mêmes folies** que mes enfants et moi avons déjà **commises**.  
S300/ ORMESSON.J D' /TOUS LES HOMMES SONT FOUS/1986, Page 117 / III  
PANDORA ! PANDORA !

49. [...] puisqu'on a **commis une terrible injustice** en arrachant ces gens à leur Afrique originelle, il serait tout à fait logique qu'on les y renvoyât.  
R527/ BLONDIN.A /MA VIE ENTRE LES LIGNES/1982, page 173 / ENTRE  
1949 ET 1955, L'ONCLE TOM ET L'ONCLE ATOME

Dans la table 6, le Npréd *faux pas* est affiché séparément de *fausse note*:

50. Ce sont les anciennes pauvres qui **commettent des faux pas**, des impairs, ce sont elles qui vous outragent ou vous ridiculisent, à celles-là, se disait Léo, on ne pardonne pas des coups comme celui qu' elle vient de me faire avec son chauffeur.  
R462/ RIVOYRE.CH DE /LES SULTANS/1964, page 193 / DEUXIÈME PARTIE,  
LE LENDEMAIN

51. J'écoute parler nos candidats maîtres et le rouge me vient aux joues, comme si ma fille **commettait une fausse note** dans un concert de collégiennes, ou mon fils, une ânerie dans un de ses premiers dîner d'adulte.  
S247/ NOURISSIER.F / LETTRE A MON CHIEN / 1975 page 156 /

Il se trouve aussi des Npréd dans le corpus Ft du type *écrit et livre* :

52. Ergo une condamnation prononcée par ces Français simples à l'encontre d'un écrivain français pour des actes ou **écrits antifrancçais commis** par lui relève de l'infamie la plus abjecte.  
S121/ TOURNIER.M /LE VENT PARACLET/1977 page 88 / II Le roi des Aulnes

#### 4.1.2. Une comparaison entre Van Dale et Frantext

Parmi les six Npréd qui sont insérées dans l'entrée lexicale Van Dale (voir l'annexe 1.a), seulement *erreur* et *péché* se trouvent parmi les 10 Npréd les plus fréquents. *Attentat* (3 fois) et *trahison* (3 fois) sont beaucoup moins fréquents que *crime* (44 fois). *Délit* est également peu fréquent par rapport aux autres Npréd avec une fréquence >10. Les 15 occurrences en bas de la table 6, *barbarie à turpitude*, ne se trouvent pas dans le corpus Ft. Ce groupe de zéro occurrences est repris dans les chapitres 4.4 et 5.2.

Malgré le fait que les Npréd de *écrit et livre*, sont peu fréquents, ces Npréd représentent la catégorie de '*zich bezondigen aan het schrijven van*', le cinquième schéma dans l'entrée de commettre dans Van Dale. D. Gaatone propose que pour un tel emploi du Vsupp 'tout objet de ce verbe prend *ipso facto* le sens d'*acte répréhensible*, même si son sens lexical est parfaitement neutre de ce point de vue.' (Gaatone, 2003 : 249) L'emploi particulier peut ainsi prendre un effet stylistique. Le schéma se trouverait dans la même sphère que l'emploi général et pourrait en effet être catégorisé comme une exploitation avec une variation du Npréd semi-productive. Nous déciderons après la fréquence de ce type de Npréd dans tous les corpus si le schéma proposé sera inséré ou non dans le chapitre 5.2.

#### 4.2. Résultats issus de la partie française du corpus EU et de LexisNexis

Du corpus EU, le corpus 558 a été rapatrié qui fournit les 558 occurrences d'une forme du Vsupp *commettre* où les graphies *commissions* et *commis(e)(s)* ont été supprimées de la séquence de recherche. Après la sélection manuelle, deux tiers des occurrences (395 résultats) font partie du corpus final et figurent dans la table 2 ci-dessous. Les fréquences de la partie française du corpus EU se trouvent dans l'a table 2.

La sélection manuelle est motivée par les conditions de l'analyse qualitative dans le chapitre 5.1. Quand un résultat n'a pas de traduction verbale pour le Vsupp, le résultat n'est pas pris en compte dans les fréquences du corpus C558<sup>46</sup> :

53. (*src*)="2135"> Mais ils l' ont fait , en **commettant finalement des actes de violence criminelle**

(*trg*)="2440"> Toch hebben ze dat gedaan, op het laatst duidelijk gepland en met crimineel geweld.

---

<sup>46</sup> Le corpus EU est marqué irrégulièrement avec des tags du type (*src*)="2135">. A défaut d'une meilleure façon de référence, les tags sont insérés dans les tokens aussi bien dans la partie française ainsi que la partie néerlandaise.

**Table 7 : les résultats du corpus C558 par fréquence décroissante**

occurrences	C558				
erreur	108	bévue	1	effraction	0
crime	44	folie	1	énergie	0
acte	44	trahison	1	enlèvement	0
infraction	25	confusion	1	espièglerie	0
délit	23	contresens	1	étourderie	0
violation	21	excès	1	fait	0
fraude	14	dégradations	1	fausse note	0
attentat	12	homicides	1	galoup	0
injustice	10	pillage	1	graffiti	0
abus	10	tuerie	1	imposture	0
génocide	9	vol	0	inceste	0
faute	8	indiscrétion	0	inconvenance	0
atrocité	8	bêtise	0	indécence	0
meurtre	6	impair	0	indignité	0
péché	5	connerie	0	infamies	0
action	5	dégât	0	irréremédiable	0
méfait	4	larcin	0	lâcheté	0
assassinat	3	abomination	0	ladrerie	0
exaction	3	gaffe	0	livre	0
imprudence	2	mensonge	0	lourdeur	0
forfait	2	cruauté	0	mal	0
faux pas	2	geste	0	manquement	0
adultère	2	horreurs	0	sabotage	0
oubli	2	sacrilège	0	sacrifice	0
violence	2	agression	0	sournoiserie	0
maladresse	2	bassesse	0	stupidité	0
massacre	2	cambriolage	0	vandalisme	0
atteinte	2	cercle	0	vilenie	0
illégalité	2	châtiment	0	viol	0
malversation	2	circoncision	0		
sottise/sotie	1	déprédations	0		
		écrits	0		
				<b>Total d'occ :</b>	<b>395</b>

Au total, le corpus C558 élargit l'ensemble de constructions par 8 nouveaux Npréd dont *violation*, *génocide* et *fraude*. Le dernier Npréd se combine aussi bien avec l'article défini, qu'avec l'article indéfini et se trouve également au pluriel, ayant chaque fois une nuance différente dans la version néerlandaise :

54. [...] lutte contre **la fraude** , où qu' elle se **commette**.  
(trg)="366"> Wij moeten allen deelnemen aan de strijd tegen fraude , waar ook gepleegd .
55. ... qui prétendent utiliser le système pour **commettre une fraude** . (trg)="2894">  
[...] tegen diegenen die de bedoeling hebben het systeem te gebruiken en misbruik te maken van de wet .
56. ... le gouvernement de mon pays , le Royaume Uni , **commettent des fraudes** sur le dos tant des régions Européenne [...]

(trg)="2382"> Ik moet zeggen dat zowel de Europese regio's en de EU zijn misleid door enkele nationale regeringen, en niet in de laatste plaats door die van mijn eigen land , het Verenigd Koninkrijk .

En général, le corpus C558 est moins varié que le corpus Ft, ayant plus de Npréd d'une fréquence haute et moins de Npréd qui occurrent un ou deux fois dans le corpus Ft. Il se trouve très peu d'occurrences de *faute* (8 fois) et très peu de résultats avec *vol* (0 fois) et *péché* (5 fois) :

57. ... les autres questions pour savoir si elles **commettent le même péché** .

(trg)="3344"> Ik wil u dus vragen , mijnheer de Voorzitter , of ik mijn vraag kan stellen, en of u onderzocht heeft of de overige vragen ook waardeoordelen bevatten .

58. [...] elle **commet le péché capital** de ne pas être une stratégie ...

(trg)="875"> De Lissabon-strategie bestaat uit een serie prachtige bedoelingen , maar we hebben de doodzonde begaan dat we er geen Gemeenschapsstrategie van hebben gemaakt [...]

Afin d'étendre le genre de textes avec des articles de presse, le corpus LN est également constitué. Ce corpus issu du corpus LexisNexis, est un corpus de textes tirés dans la base de données LexisNexis de la période 17-05-2008 au 17-05-2010. Au total, il s'agit de 414 résultats dont les fréquences sont affichées dans la table 7.

**Table 8 : les résultats du corpus LN par fréquence décroissante**

occurrence	LN				
erreur	85	forfait	2	couac	1
faute	77	larcin	2	débauche	1
crime	27	excès	2	destruction	1
infraction	22	maladresse	2	diffamation	1
attentat	16	massacre	2	embarquée	1
agression	15	geste	2	en-avant	1
délit	13	homicides	2	imprécision	1
violence	13	tuerie	2	incorrection	1
acte	12	exhibition sexuelle	2	opuscule	1
exaction	12	tentative	2	ouvrage	1
vol	11	turpitude	2	perte de balle	1
fait	7	imprudence	1	poème	1
meurtre	6	action	1	recueil	1
atteinte	6	mensonge	1	suicide	1
viol	5	illégalité	1	tube	1
atrocité	4	dégradations	1	injustice	0
faux pas	4	écrits	1	génocide	0
cambrilage	4	espièglerie	1	assassinat	0
hold-up	4	étourderie	1	méfait	0
irréparable	4	pillage	1	indiscrétion	0
péché	3	bavure	1	bêtise	0
impair	3	bogey	1	sottise/sotie	0
livre	3	bourde	1	bévue	0
violation	2	braquage	1	folie	0
abus	2	clic	1	adultère	0
fraude	2	contrefaçon	1	connerie	0
				dégât	0

oubli	0	déprédations	0	lâcheté	0
trahison	0	effraction	0	ladrerie	0
abomination	0	énergie	0	lourdeur	0
confusion	0	enlèvement	0	mal	0
contresens	0	fausse note	0	manquement	0
gaffe	0	galoup	0	sabotage	0
malversation	0	graffiti	0	sacrifice	0
cruauté	0	imposture	0	sournoiserie	0
horreurs	0	inceste	0	stupidité	0
sacrilège	0	inconvenance	0	vandalisme	0
bassesse	0	indécence	0	vilenie	0
cercle	0	indignité	0	<b>total</b>	<b>414</b>
châtiment	0	infamies	0		
circoncision	0	irrémissible	0		

Le corpus LN ne comprend que 70 types de Npréd au total, mais affirme les tendances d'une distribution lexicale avec une répartition inégale de fréquences, c'est-à-dire qu'il se trouve une dizaine de Npréd très fréquents et beaucoup plus de formes (45 Npréd) peu fréquents (1-2 fois). Les occurrences dans le corpus LN montrent également de la variation syntaxique comme pour *acte* avec Npréd +ADJECTIF et Npréd+ DE mais aussi sémantiquement : il se trouve plus de variation aux compléments qui coexistent avec *commettre un acte de*, en ajoutant dans l'exemple 22 la préposition *contre*:

59. ...la télé-réalité doivent être dénoncées lorsque celle-ci conduit les personnes à accepter de **commettre des actes aussi graves et dégradants**, pour elles-mêmes et pour les autres, qu'elles ne commettraient pas en d'autres circonstances.

*L'Humanité, 27 mars 2010*

60. ...leur condamnation "de **tous les actes de terrorisme**, où qu'ils se déroulent et quiconque **les commette**" et déploré "la mort d'innocents" qui en découle.

*Agence France Presse, 1 avril 2010*

61. **Le nombre d'actes de violence commis** par des écoliers russes contre leurs professeurs a plus que doublé depuis l'an dernier, ...

*Agence France Presse, 1 octobre 2008*

Les occurrences sont marquées par leur origine par l'emploi des Npréd, témoignent de registre de rapportage dans la presse par la fréquence haute de Npréd comme *agression, délit* et *atteinte* mais aussi d'une variation de Npréd qui dénotent des crimes (*pillage, braquage*), ainsi que des termes issus d'articles de sport (*faux pas, perte de balle, en-avant*).

#### 4.3. Le Corpus Ft, la partie français du corpus EU et le corpus LN : une comparaison

Au total, les corpus ensemble relèvent à un nombre de 1286 tokens. Il s'agit de 118 types dont la moitié (57 types) ont une occurrence de 3 fois ou plus. Bien que le nombre total des occurrences soit considérable, le top-10 représente déjà 69% des tokens et le top-20 82% du total des tokens :

**Table 9 : les résultats de l'ensemble des corpus monolingues par fréquence décroissante**

occurrences	Ft	C558	LN	total
erreur	97	108	85	290

faute	84	8	77	169
crime	44	44	27	115

acte	11	44	12	67
infraction	8	25	22	55
délit	8	23	13	44
péché	34	5	3	42
attentat	3	12	16	31
violation	0	21	2	23
meurtre	9	6	6	21
vol	8	0	11	19
imprudence	15	2	1	18
injustice	7	10	0	17
violence	2	2	13	17
abus	4	10	2	16
agression	1	0	15	16
exaction	1	3	12	16
fraude	0	14	2	16
atrocité	3	8	4	15
action	7	5	1	13
forfait	7	2	2	11
faux pas	3	2	4	9
génocide	0	9	0	9
assassinat	5	3	0	8
atteinte	0	2	6	8
fait	1	0	7	8
impair	5	0	3	8
méfait	4	4	0	8
indiscrétion	7	0	0	7
bêtise	6	0	0	6
larcin	4	0	2	6
sottise/sotie	5	1	0	6
viol	1	0	5	6
bévue	4	1	0	5
cambriolage	1	0	4	5
excès	2	1	2	5
folie	4	1	0	5
maladresse	1	2	2	5
massacre	1	2	2	5
adultère	2	2	0	4
connerie	4	0	0	4
dégât	4	0	0	4
geste	2	0	2	4
hold-up	0	0	4	4
irréparable	0	0	4	4
livre	1	0	3	4
mensonge	3	0	1	4
oubli	2	2	0	4

trahison	3	1	0	4
abomination	3	0	0	3
confusion	2	1	0	3
contresens	2	1	0	3
gaffe	3	0	0	3
homicides	0	1	2	3
illégalité	0	2	1	3
malversation	1	2	0	3
tuerie	0	1	2	3
cruauté	2	0	0	2
dégradations	0	1	1	2
écrits	1	0	1	2
espièglerie	1	0	1	2
étourderie	1	0	1	2
exhibition sexuelle	0	0	2	2
horreurs	2	0	0	2
pillage	0	1	1	2
sacrilège	2	0	0	2
tentative	0	0	2	2
turpitude	0	0	2	2
bassesse	1	0	0	1
bavure	0	0	1	1
bogeys	0	0	1	1
bourde	0	0	1	1
braquage	0	0	1	1
cercle	1	0	0	1
châtiment	1	0	0	1
circoncision	1	0	0	1
clic	0	0	1	1
contrefaçon	0	0	1	1
couac	0	0	1	1
débauche	0	0	1	1
déprédations	1	0	0	1
destruction	0	0	1	1
diffamation	0	0	1	1
effraction	1	0	0	1
embardée	0	0	1	1
en-avant	0	0	1	1
énergie	1	0	0	1
enlèvement	1	0	0	1
fausse note	1	0	0	1
galoup	1	0	0	1
graffiti	1	0	0	1
imposture	1	0	0	1
imprécision	0	0	1	1



inceste	1	0	0	1
inconvenance	1	0	0	1
incorrection	0	0	1	1
indécence	1	0	0	1
indignité	1	0	0	1
infamies	1	0	0	1
irrémediable	1	0	0	1
lâcheté	1	0	0	1
ladrerie	1	0	0	1
lourdeur	1	0	0	1
mal	1	0	0	1
manquement	1	0	0	1
opuscule	0	0	1	1
ouvrage	0	0	1	1
perte de balle	0	0	1	1
poème	0	0	1	1
recueil	0	0	1	1
sabotage	1	0	0	1
sacrifice	1	0	0	1
sournoiserie	1	0	0	1
stupidité	1	0	0	1
suicide	0	0	1	1
tube	0	0	1	1
vandalisme	1	0	0	1
vilenie	1	0	0	1
	<b>477</b>	<b>395</b>	<b>414</b>	<b>1286</b>

En comparant les dix fréquences les plus hautes des occurrences dans les corpus C558 et LN aux résultats de Ft, les Npréd *faute* et *péché* sont beaucoup moins fréquents. Il est pareil pour les Npréd qui ne désignent pas de crimes comme *bêtise*, *bévue*, *folie*, *impair* et *connerie*. Comme dans la comparaison entre Van Dale et les occurrences de Ft, du corpus LN, il s'avère aussi si le Npréd prend un complément. En recherche de l'emploi prototypique, les occurrences de *commettre des violences* sont rassemblées ci-dessous :

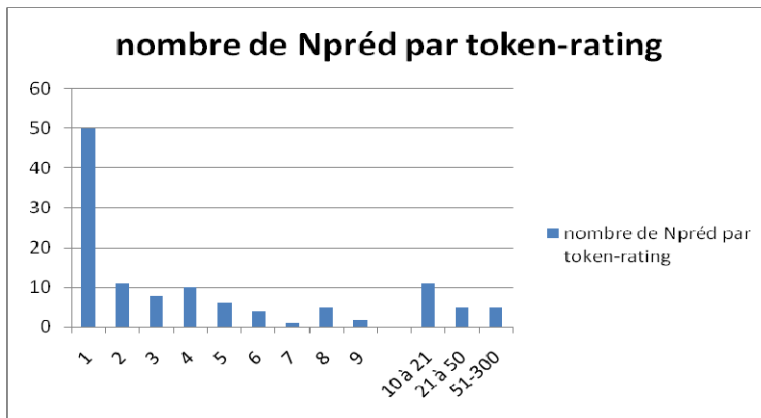
62. Pendant très longtemps, nous l' avons vu, les crimes proprement dits étaient des "crimes de sang " c' est-à-dire des **actes de violence commis contre des personnes ou contre des biens**, le plus souvent par des individus situés au bas de l' échelle sociale que les difficultés de l' existence et une certaine déficience intellectuelle et morale.  
P337/ GURVITCH.G /TRAITE DE SOCIOLOGIE T.2/1968, page 222 / -
63. **La violence commise contre son enfant** par cette malheureuse ressemble fort à celle d'une bête se rongant un membre pour s'arracher à la trappe où la cruauté des hommes l'a fait choir, dit amèrement le philosophe.  
S003/ YOURCENAR.M /L'OEUVRE AU NOIR/1968, page 815 / TROISIÈME PARTIE : LA PRISON, LA VISITE DU CHANOINE
64. « On se focalise souvent sur **les violences commises sur la voie publique**, or la sphère privée et familiale est quasiment aussi dangereuse », commente ...  
*Les Echos*, 19 novembre 2008
65. "Je souhaite que les **violences commises sur un agent de l'Education nationale** constituent à l'avenir des circonstances aggravantes [...].  
*Libération.fr* 17 avril 2009
66. ...constatons davantage de cas de mutilation, d'extrême violence et de torture **dans les violences sexuelles commises contre les femmes et les filles**, et un bien plus grand nombre des victimes sont ...  
*News Press*, 15 novembre 2009
67. ... l' on tente de résoudre le problème **en commettant des violences à l' encontre de la population** ...  
*Corpus EU*

Au total, de nombre des résultats relève de 17 occurrences du Npréd *violence(s)*, sept occurrences prennent un complément. Les résultats montrent trois prépositions différents : *contre* (3 fois), *sur* (3 fois) et *à l'encontre de* (1 fois).

Bien que cet exemple du Npréd *violence(s)* seul mène à la conclusion que l'ampleur du corpus limite les conclusions tirées de ce nombre modéré de résultats, il s'avère aussi que la construction Vsupp *commettre des violences sur quelqu'un* n'est pas la seule combinaison avec des compléments. Par conséquent, le support des corpus utilisés n'est pas convaincant afin de généraliser l'emploi de Vsupp *commettre* une seule construction incluant une préposition. A cela s'ajoute que les fréquences relatives des constructions Vsupp avec une préposition ne se distinguent pas l'une des autres Prép. Les analyses de Fellbaum et al. (2007) et Hanks et al (2006) ne relèvent pas des chiffres derrière les schémas proposés, mais on peut supposer que leurs propositions pour l'affichage lexicographique sont soutenues par l'analyse de corpus à une plus grande échelle.

On constate que les fréquences des types trouvés pourraient contribuer à l'amélioration de l'entrée lexicale, mais également que le support quantitatif ne se présentera pas comme un facteur prédominant dans la conception d'une nouvelle entrée pour le verbe *commettre*. En outre, l'ensemble des résultats issu du corpus ne constitue pas un support stable à tous les Npréd.

**Table 10 : la distribution des tokens : le nombre de Npréd par le nombre de tokens**



La table 10, qui affiche la distribution des fréquences par le nombre de Npréd versus nombre d'occurrences, montre bien la distribution inégale parmi les types : il existe beaucoup de Npréd qui ont un seul token, bien que les types els plus fréquents ont plus de 20 tokens. Le nombre limité d'occurrences par type (Npréd) ne permet pas de sélectionner les Npréd pour l'entrée lexicale de manière fondée, quand il s'agit seulement d'un nombre d'occurrences moins de 10.

#### 4.4. Les corpus monolingues vs. Van Dale

Cette section est réservée aux remarques qui regardent les données issues des corpus monolingues et la relation avec l'ensemble de données issues de Van Dale à savoir six pistes de réflexion qui résultent de l'analyse dans ce chapitre :

- la délimitation de la nouvelle entrée lexicale dans notre recherche
  - les Npréd de zéro occurrence
  - l'insertion des informations concernant le nombre et l'article en construction Vsupp
  - l'insertion de l'exploitation *commettre un faux pas*
  - l'insertion d'un schéma de la combinatoire du Vsupp avec des Npréd du type *écrit(s)*
  - la contribution des Npréd marqués comme exploitations à notre recherche
- Il s'agit dans tous les cas propositions personnelles pour la nouvelle entrée lexicale.

Il s'est avéré de la section 4.1.1. et 4.2. qu'il existe un nombre considérable de formes non-verbales de commis et commissions. Le nombre d'occurrences de *commis* comme adjectif dans le corpus Frantext est un argument en faveur de l'insertion d'un schéma sur l'emploi '*benoemd*', comme montent les exemples dans le chapitre 4.1.1. Cet emploi ainsi que 'emploi substantivé ne se trouve pourtant pas dans les autres corpus et peut donc être considéré spécifique au registre littéraire. Par contre, la forme *commission* est assez fréquente mais la fréquence haute est la conséquence des méthodes de recherche automatique. L'affichage de ce schéma peut se limiter à deux exemples pour lesquels je propose un seul exemple de l'emploi adjectival et une référence au substantif pour la nouvelle entrée français-néerlandais de *commettre*, en adoptant en partie le schéma 5 de l'ancienne entrée (annexe 1a) :

68. Commis d'office - toegevoegd advocaat  
 → Commis (N)

Comme résultat à la comparaison entre les corpus monolingues et Van Dale, il reste 11 Npréd qui se trouvent dans la colonne '*binnen artikelen*' pour *commettre* dans Van Dale et qui n'ont aucune occurrence dans l'ensemble des corpus. Ces Npréd sont *barbarie*, *débauche*, *félonie*, *goujaterie*, *énormité*, *coup d'état*, *indélicatesse*, *malhonnêteté*, *plagiat*, *impolitesse*, *impudicité* et *irrégularité*. Il se présente plusieurs possibilités ou alternatives à l'insertion de ces Npréd. L'occurrence de *barbarie* dans Van Dale n'est pas comme Npréd mais fait partie de la construction *commettre un acte de barbarie*. *Débauche* et *félonie* sont du même type de constructions. Les autres Npréd cités ci-dessus ne sont pas des Npréd pertinents quant à leur

distribution quantitative dans les corpus. *Enormité* fait également partie des Npréd avec zéro occurrences. Même dans le Van Dale NF, l'entrée de *stommitteit* affiche la variation des Npréd mais ne fournit qu'un exemple, *faire des bêtises*. Pour les deux entrées, le nombre de possibilités pour la traduction est restreint à un seul exemple bien qu'il existe une variation de Vsupp, en français par *commettre/faire* et en néerlandais on dirait *begaan/uithalen*. Pour *coup d'état*, l'entrée de Van Dale FN fournit seulement le Vsupp *faire*, bien que l'entrée lexicale NF de *staatsgreep* propose *commettre* et *perpétrer un coup d'état*. Dans une recherche après la distribution lexicale de *coup d'état* dans Frantext et LexisNexis, les Vsupp utilisés sont *faire*, mais aussi *exécuter* et *opérer*. En passant d'une perspective monolingue à une perspective bilingue, je verrai dans le chapitre 5.2 de quels schémas leur traduction peut être dérivée et si les Npréd trouvent portant une place dans l'entrée lexicale.

Je me suis posé la question si la combinatoire du nombre et de l'article des Npréd doivent être insérée. Dans le corpus Frantext, certains Npréd donnent du sens au Vpréd en combinaison avec l'article défini :

69. [...] pour ceux qui voudraient **commettre la folie de** cloner l'être humain.  
(trg)="18"> De bij dit experiment betrokken wetenschappers hebben zelf erkend dat zij degenen die zo dwaas zouden willen zijn om mensen te klonen, de weg hebben  
gewezen.

Il se peut également que la construction Vsupp soit plus flexible, comme montre le Npréd *fraude(s)*. Quand on part du principe que tous les Npréd sont flexibles syntaxiquement, il n'est pas nécessaire d'ajouter une remarque sur la variation du déterminant, au moins du point de vue monolingue.

70. ... qui prétendent utiliser le système pour **commettre une fraude**.  
(trg)="2894"> Maar de garanties moeten zich eveneens, en op dezelfde manier en met strengheid, richten tegen diegenen die de bedoeling hebben het systeem te gebruiken en misbruik te maken van de wet.
71. ... core , nous pourrions bien les inciter à **commettre ces fraudes** ou ces abus .  
(trg)="3896"> Erger nog, we zouden met ons beleid wellicht zelfs uitnodigen tot fraude en misbruik.

Par contre, un certain nombre des Npréd se trouve surtout au pluriel pour lesquels il faudra dévier de notre méthode d'affichage des résultats dans les tables de ce chapitre, et insérer une remarque comme (*vaak in mv*). Le Npréd *exaction* par exemple, ne se combine qu'au pluriel avec *commettre* :

72. ...à sa récente expérience de diplomate en Afrique de l'Ouest. Il y raconte les **exactions commises** par les commandos islamistes dans cette région. Le roman s'intitulera  
Katiba, mot arabe appartenant au ...  
Le Figaro économie, 9 février 2010

Ayant un nombre limité de tokens par Npréd , je ne considère pas insérer des informations en supplément sur l'emploi des déterminants.

Bien que *faux pas* soit une collocation fixe dans le Van Dale, le dictionnaire FN ne fournit pas l'option de *faux pas* comme Npréd en collocation avec *commettre* : le verbe léger *faire* seulement sert comme Vsupp. Il serait intéressant de savoir quelle est la fréquence de *commettre un faux pas* par rapport à *faire un faux pas*. Une recherche dans l'ensemble de textes dans Frantext catégorisé dans la période 1960-2000 montre que *commettre un faux pas* se trouve une seule fois envers une occurrence de 112 fois de *faire un faux pas*. Je propose donc d'insérer le verbe comme une variation dans les entrées lexicales NF *misstap*, *uitglijder* et *maken* dans le dictionnaire FN dans l'entrée de *commettre* comme une traduction en contexte.

Les Npréd tu type *livre* et *écrits*, sont marqués par une fréquence basse mais il s'agit dans mes yeux d'un schéma à part et pertinent pour la nouvelle entrée lexicale. Je propose une reconsidération de ces Npréd après l'analyse bilingue afin de constituer un schéma.

Dans le chapitre 3.2, on a déterminé des conditions quantitatives et qualificatives fondées sur la recherche de corpus pour l'insertion des types qui représentent l'emploi normal (N) et les alternances (A) des constructions Vsupp uniquement. A la base des données quantitatives, les types d'une occurrence de <10 n'entreraient pas en ligne de compte pour l'insertion dans l'entrée lexicale. Pourtant, les occurrences constituent une valeur de contrôle pour les futurs emplois normaux, à partir desquels l'utilisateur devrait être capable de trouver une traduction pour les constructions Vsupp qui ne sont pas insérées dans l'entrée lexicale. En effet, l'ensemble de tous les tokens est pris pour l'analyse de traductions du verbe dans les constructions Vsupp. Dans l'affichage concentré aux Npréd, j'ajoute également des propositions pour la traduction de ce dernier emploi, les exploitations, et pour les expressions idiomatiques.

Les Npréd qui ont une occurrence modérée sont également d'importance de point de vue bilingue. En prenant comme exemple *geste* et *contresens*, deux Npréd peu fréquents dans les corpus monolingues, les Npréd ne se trouvent pas dans Van Dale en collocation avec le Vsupp *commettre*. Ces Npréd sont également intéressants par rapport à la traduction en néerlandais, en raison de la traduction possible '*maken*' :

73. Je lui répondis que je serais incapable de **commettre ce geste** et donc que je déclinai sa demande.

R720/ GUIBERT.H /A L'AMI QUI NE M'A PAS SAUVE LA VIE/1990, page 38 /13

*Traduction proposée: een gebaar maken, doen*

74. Il semble qu'on soit venu à cette pratique à la suite d'un **contresens commis** dans l'interprétation des sceaux antiques : ce n' était pas la figure du titulaire qui était gravée sur les intailles, mais celle d' un personnage illustre, qu' il appartienne à l' histoire ou à la légende, à la famille ou seulement à la patrie.

P653/ \*Anonyme /L'HISTOIRE ET SES METHODES/1961, pages 408-409 /

*Traduction proposée: een vergissing maken, begaan*

On entrera plus en détail sur l'aspect bilingue dans le chapitre suivant.

## Chapitre 5 : Résultats du corpus bilingue EU

Comme l'analyse qualitative constitue le deuxième centre d'intérêt, dans ce présent chapitre nous inventorions la contribution du support des corpus à l'analyse qualitative des constructions Vsupp. Au niveau de l'analyse de corpus, nous passons de la partie française du corpus EU aux versions néerlandaises des constructions Vsupp sélectionnées. Les données du corpus affichées dans le chapitre 5.1 augmentent l'ensemble des données quantitatives sur le plan de la sémantique lexicale et en même temps, constituent la vérification des données de Van Dale dans une analyse contrastive entre le français et les constructions correspondantes en néerlandais. Avec la conception d'une nouvelle méthode d'affichage, notre objectif est d'hierarchiser les collocations dans l'entrée lexicale et d'insérer une structure apte à l'utilisateur. Nous adoptons partiellement la méthodologie de TNE développé par Hanks et Pustejovsky (2004,2005) afin d'attribuer un statut aux collocations à l'aide des fréquences des collocations. Dans les chapitres 5.2 et 5.3, deux conceptions pour une nouvelle entrée lexicale sont présentées dans lesquelles nous prenons en compte les résultats d'aussi bien l'analyse des dictionnaires dans le chapitre 3 que des résultats qualitatives et quantitatives. Finalement, nous aborderons une réflexion plus générale sur la contribution de l'analyse de corpus dans le domaine de la traductologie dans la conclusion de la section 5.4.

### 5.1. Résultats de la recherche parallèle dans le corpus EU

D'abord, les résultats bilingues sont présentés dans la table 11. Le support bilingue est exclusivement issu du corpus C558, du fait que les versions néerlandaises des actes de conférence ont été alignées aux textes français. Les conditions de recherche ont été optimisées afin d'obtenir plus de traductions adéquates des constructions Vsupp : les constructions Vsupp qui ont été traduites par un seul substantif ne sont pas prises en compte. Pour le Npréd *faux pas*, les traductions contextuelles par un seul substantif ou par une construction plus libre ne peuvent pas contribuer à l'élargissement des données sur la traduction de *commettre* :

75. [...] même lorsque cette dernière commet des faux pas ou persiste dans l'erreur ...  
(trg)="1088"> Is dit Parlement soms in het leven geroepen **om de Commissie in bescherming te nemen** ?
76. ... agriculteurs et aux petites laiteries s' ils commettent le moindre faux pas administratif .  
(trg)="1029"> Waarom is de Commissie zo inschikkelijk en laat zij grote agrifood-bedrijven ongestraft , terwijl de Commissie zeer streng is bij het opleggen van boetes **voor simpele administratieve voorvallen** aan boeren en ambachtelijke zuivelbedrijven?

Le point d'exclamation dans la table 11 marque un doute personnel sur la traduction qui peut être le résultat de l'absence (d'une partie) du contexte. Dans la table se trouvent uniquement les traductions et les remarques sur la combinatoire et la syntaxe.

**Table 11 : les traductions des constructions Vsupp dans les versions néerlandaises**

Npréd	traductions dans le corpus C558
abus	(+sur des enfants) kindermisbruik, (iem) misbruiken, hun boekje te buiten gaan, rommelen (intrans), zich bezondigen aan, slecht behandelen,
abus	(+de pouvoir) er vindt machtsmisbruik plaats, zich bezondigen aan machtsmisbruik
acte +ADJ	ADJ daden begaan,

	overgaan tot ADJ daden verrichten, zich schuldig maken aan, (acte malveillant) kwaad aandoen, iets ADJ doen, ADJ gedrag vertonen, zich voordoen (acte illégal) de wet overtreden
acte +DE	de violence: geweld plegen, de torture: martelen, d'ingérence: er vindt inmenging plaats
acte +DE +contre	Een ADJ handeling verrichten tegen
action	daad plegen/begaan, acties ondernemen
adultère	overspel plegen
assassinat	moorden (intrans), moorden begaan
atteinte	( +à qqn): een aanslag doen op
attentat	aanslag plegen, bloedbaden aanrichten/uitvoeren(!)
atrocité	gruwelen/verschrikkingen/wreedheden begaan
bévue	een flater slaan
confusion	zichzelf in verwarring brengen
contresens	radicaal langs het doel heen schieten
coup d'état	Een staatsgreep plegen
crime	er vindt criminaliteit plaats, een misdrijf plegen/begaan, criminaliteit bedrijven (!), schade aanrichten
dégât	(iem) schaden
dégradation	(iem) schaden
délit	misdaden/delict begaan, strafbare handelingen verrichten, (délits corporels) toebrengen van lich. letsel, zich schuldig maken aan
dévastation	schade aanrichten
discrimination	(optioneel. iem) discrimineren
embardée	Blunder begaan
erreur	Fouten/vergissingen maken, in de fout lopen, fouten maken , zich vergissen, fout lopen (intrans), fouten gebeuren(!), het bij het verkeerde eind hebben, een vergissing begaan, de plank mislaan, de fout ingaan, een misvatting maken
excès	kwistig zijn met
exaction	gewelddaden plegen, misdrijven begaan,
fait	zich schuldig maken aan (het gepleegde feit)
faute	iets verkeerd doen, een fout maken, fout begaan, zijn,

folie	in een fout vervallen, grote vergissing begaan, (ART DEF) zo dwaas zijn om
forfait	criminele activiteiten uitvoeren
fraude	frauderen (intrans) fraude plegen, misbruik maken van, zich bezig houden met fraude
génocide	zich schuldig maken aan volkerenmoord, genocide plegen, volkerenmoord uitvoeren, genocide toepassen,
infraction	in overtreding zijn, overtreding/misdrijf begaan, iets misdoen, een overtreding begaan, delicten plegen, iets verkeerd doen,
injustice	onrechtvaardig behandelen, onrecht plegen (!) onrecht begaan, onrechtvaardigheid plegen (!)
illégalité	(context) iemand tot outlaw maken, iets onrechtmatig doen
imprudence	gevaarlijk bezig zijn, onvoorzichtig zijn,
infamie	misdaden plegen
injustice	onrechtvaardigheden begaan
irrégularité	Er vinden onregelmatigheden plaats
maladresse	een onhandigheid begaan
massacre	bloedbaden aanrichten
méfait	wandaden plegen, misdaden begaan
meurtre	moorden, gedood worden, een moord plegen, vermoorden
oubli	het zou nalatig zijn van mijn kant, zaken worden vergeten (!)
péché	zondigen, (péché capital) een doodzonde begaan
pillage	aan het plunderen slaan
sottise	een absurd voornemen uitvoeren
trahison	verraad plegen
tuerie	slachtingen aanrichten
viols	verkrachtingen begaan
violation	Er vinden bruitheden plaats, (violations des droits de l'homme) er vinden mensenrechtenschendingen plaats, Er vinden gewelddadigheden plaats, (du droit) overtreden, schenden misdaden plegen,
violence	(violences des droits de l'homme)schendingen van de mensenrechten begaan, uitoefenen van geweld, geweld gebruiken tegen



**Légende**

+ est un complément  
 intrans = verbe intransitif  
 (iem) = COD humain du V(supp) néerlandais  
 ART DEF = article défini  
 ADJ = adjectif  
 X = Npréd

Malgré le nombre de 558 collocations, le corpus ne fournit que pour 51 Npréd une ou plusieurs traductions adéquates avec au moins un V(supp) en néerlandais. Toutefois, les traductions montrent en général une grande variété. Le nombre restreint de traductions est en partie la conséquence du choix du verbe. Le verbe *commettre* est relativement peu utilisé dans un corpus de plus de 29 millions de mots.

Dans la table 11, j'ai distingué entre les compléments qui se combinent occasionnellement avec les Npréd et ceux qui sont plus pertinents pour la traduction prototypique et qui se trouvent dans la colonne gauche : par rapport à *acte + ADJ* qui est traduit de plusieurs manières, le paradigme *acte + DE* constitue un paradigme à part avec les N *violence, torture* et *ingérence*. Par conséquent, les deux variations syntaxiques ont été séparées dans cette colonne. Par contre, *abus de pouvoir* se produit seulement avec le N *pouvoir* en se trouve par conséquent dans la colonne droite parmi les traductions en contexte. Selon les données du corpus LN, le Npréd *atteinte* peut se combiner avec différentes prépositions (*sur, contre, à*). Le corpus fournit seulement une traduction de la combinaison avec la préposition *à* de ce qui s'avère que le nombre de combinaisons dans le corpus EU est très limité. Les données quantitatives servent ainsi à la vérification des constructions Vsupp en français aussi bien dans Van Dale que dans les corpus français.

Occasionnellement, on trouve un emploi particulier de l'article défini, par exemple dans *commettre la folie de*, comme on a trouvé dans Frantext l'occurrence *commettre le péché de* avec un article défini obligatoire mais il ne semble pas nécessaire d'insérer de schéma général pour [ART DEF Npréd DE] pour ces Npréd.

Pour le Npréd *illégalité*, les traductions en contexte ne sont pas utiles pour la conception générale de son emploi:

77. ...que de voir ses propres agents de la paix commettre des illégalités .  
 (trg)="2398"> President Bush wil internationaal soms cowboy spelen maar dreigt zijn **eigen peace makers tot outlaws te maken**.

78. ...personne ne va prendre la responsabilité de commettre des illégalités .  
 (trg)="1915"> Het is zonneklaar dat dat onrechtmatig zou zijn en niemand zal het voor zijn verantwoording willen nemen om **iets te doen wat onrechtmatig is**.

La deuxième traduction peut être considérée comme une stratégie de traduction, mais il n'est pas possible de dériver un paradigme comme *iets ADJ doen* n'est pas spécifique à la collocation de *commettre des illégalités*.

Plusieurs traductions peuvent être vues comme libres : les V(supp) néerlandais *aanrichten, uitvoeren, verrichten, plaatsvinden, gebeuren* :

79. ... e sont précisément les États membres qui commettent beaucoup d'erreurs.  
 (trg)="320"> Er is ook inderdaad nood aan actie in de regionale en de nationale parlementen, want het is juist in de lidstaten dat er heel wat **fouten gebeuren**.

80. ...je puis vous dire qu' on y commet les pires violations des droits de l' homme...  
 (trg)="568"> (...) ik kan u zeggen dat daar **zeer zware schendingen van de mensenrechten plaatsvinden**.

Un verbe comme *plaatsvinden* est relativement fréquent dans les données néerlandaises du corpus C558, mais la question se pose si la traduction est pertinente pour l'insertion dans un

dictionnaire en raison de son sens générale. La traduction par un verbe de sens général est significative pour la délimitation sémantique non-fixe du Vsupp qui interfère dans la traduction dans une langue avec d'autres délimitations sémantiques.

Il s'avère que dans ce genre de textes du corpus EU que la signification de *violences* et *violations* se ressemblent surtout en contexte des droits de l'homme.

**Table 12 : violence et violation**

occurrences	Ft	C558	LN	total
violence	2	2	13	17
violation	0	21	2	23

Quand les données quantitatives sont comparées, il s'avère que *violence* a une fréquence haute dans le corpus LN, bien que *violation(s)* soit plus fréquent dans le corpus EU. Quand les données du corpus LN sont analysées, la signification générale de *violence* reste plus proche du sens propre :

81. Alors que le conseil des droits de l'homme de l'ONU se penche aujourd'hui sur le cas de la RD-Congo, deux rapports mettent en évidence **les violences commises par l'appareil d'État**.

La Croix, 28 novembre 2008

82. La France a « condamné les arrestations arbitraires et les **violences commises contre de simples manifestants**, venus défendre leur droit à la liberté d'expression et leur aspiration...

Figaro magazine, 29 novembre 2009

Pour la traduction des exemples 81 et 82 en néerlandais, une construction nominale suffit. Le Npréd *violation* dans le corpus C558 prend un deuxième sens qui ressemble à *infraction à la loi* et les traductions par *overtreden (wet)* et *schenden (recht)* :

83. Si nous agissions autrement ici, nous commettrions une violation du droit .

(trg)="3107"> Wanneer wij op dit vlak op een andere manier zouden handelen , zouden we ontegenzeggelijk **de wet overtreden**.

84. ... s' agissait d' un séquestre et le Canada commettait une violation du droit international . ...

(trg)="1796"> Zoals de commissaris indertijd zeer terecht gezegd heeft , ging het om een kaping en Canada **schond** daarbij **het internationaal recht**.

Pour le Npréd *oubli*, le corpus fournit deux autres traductions qui peuvent servir comme alternatif dans la traduction aux propositions de Van Dale pour *verzuimen*: *nalatig zijn* et *vergeten worden*.

85. (src)="612"> Je commettrais un oubli en ne parlant pas de la mascarade...

(trg)="743"> **Het zou nalatig zijn van mijn kant** indien ik niets zou zeggen over de farce van de Conventie...

86. ... nécessaire de réécrire le préambule, car il commet deux oublis essentiels.

(trg)="625"> Tenslotte zal de preambule herschreven moeten worden aangezien er twee essentiële zaken in **worden vergeten**.

Il se peut bien qu'une telle traduction en contexte résulte dans une traduction plus naturelle qu'une traduction correcte mais construite. Contraire à *oubli*, le corpus C558 ne fournit pas de traductions pour les Npréd du type *livre*. Le genre de textes se prête moins aux emplois littéraires du verbe commettre.

On peut conclure que l'analyse qualitative dans un corpus bilingue relève des résultats intéressants pour compléter les données quantitatives. Il faut néanmoins remarquer que la portée du corpus est relativement restreinte quant à la variété des Npréd. Avec une base de données plus ample et variée, les constatations seraient fondées sur un support plus stable que

dans notre analyse exécutée dans le corpus C558. Le manque de variation de Npréd nous laisse avec un nombre limité de variations de collocations et avec relativement peu de traductions. En exécutant une analyse plus élargie, il faudrait également porter plus d'attention aux compléments afin d'en savoir plus sur la combinatoire des Vsupp avec les propositions. Finalement, les résultats dans la table 11 et les constatations peuvent contribuer aussi à la catégorisation par Npréd qu'à une catégorisation par verbe.

Pour l'application de résultats qui va suivre, il est important de souligner dans quelle mesure les données quantitatives et qualitatives motivent la sélection des Npréd pour l'entrée du dictionnaire. Vu le total d'occurrences et la répartition dans les corpus différents, les résultats monolingues sont représentatifs au niveau du français, mais il est impossible de considérer le corpus bilingue comme adéquate au niveau du registre. La nature des textes et le nombre modéré de traductions fournies par la partie néerlandaise du corpus C558 ont pour conséquence que les propositions sont un essai envoi d'une restructuration de l'entrée lexicale. L'application prend donc la forme d'un pilot au niveau lexicographique. En outre, les jugements personnels seront décisifs aux données quantitatives et qualitatives au moment où les résultats de recherche n'offrent pas assez de support pour motiver la sélection des Npréd. La dernière remarque est que la nouvelle conception lexicale se concentrera sur l'emploi de *commettre* comme verbe transitif dans la langue-source. Les emplois réflexifs en les expressions fixes comme on a vu pour 'commis d'office' dans le chapitre 4.1.2. ne sont pas pris en compte dans les propositions, et également pas dans la proposition finale.

## 5.2. Une proposition de catégorisation par Npréd

Comme on passe à la catégorisation afin de concevoir une entrée lexicale, je me servirai des constatations du chapitre 3.2.2. et 3.2.3 : les propositions de Hanks et al. (2006) sont marquées par une divergence entre la théorie, la TNE, à savoir la sélection stricte des Npréd qui reflètent l'emploi normal et la conception de l'entrée concrète qui est très précise mais et en même temps la réalisation des schémas exhaustives. Dans ce chapitre, la TNE est adaptée partiellement aux objectifs de cette recherche : par l'élargissement des catégories qui sont pertinentes pour l'insertion dans l'entrée lexicale afin d'afficher aussi des alternances fréquents, une entrée lexicale bilingue plus compacte est envisagée. Je veux également montrer de quelle manière les données quantitatives ont été affichées en remplaçant l'ordre alphabétique par une structure hiérarchisée sur pertinence.

Comme les traductions comprennent un champ limité des constructions Vsupp sélectionnées, l'entrée lexicale appuie également sur des jugements personnels sur l'emploi langagier en néerlandais. La méthode d'insertion des Npréd de Schroten(1999) est adoptée : les traductions du Vsupp seront accompagnées du Npréd en langue-cible correspondant pour expliciter la traduction en contexte :

87. Bereiden (maaltijd), doen (boodschap)

Il s'est avéré du chapitre 3.5. que le traitement des Npréd qui vont avec *commettre* est non-systématique : les constructions Vsupp se trouvent parfois dans le Van Dale français-néerlandais et parfois dans l'entrée néerlandais-français. La distribution lexicale n'est également pas intégrée dans les entrées de manière structurelle. En outre, la conception de l'entrée lexicale de Van Dale est concentrée sur un traducteur ayant le néerlandais comme langue maternelle (voir 3.5) ce qui a pour conséquence une façon sélective d'afficher des exemples clés. Si les exemples donnés dans l'entrée du Vsupp ne ressemblent pas la collocation recherchée, le dictionnaire ne fournira pas de solutions adéquates à l'utilisateur-traducteur. Notre restructuration du dictionnaire vise également un utilisateur qui n'a pas le néerlandais comme langue maternelle.

Je lance d'abord une proposition de restructuration de l'entrée lexicale qui prend comme répartition la catégorisation des Npréd. Dans la section 5.3., elle est suivie par un affichage par verbe afin d'élaborer une comparaison entre les deux méthodes d'affichage et finalement

de combiner les éléments les plus pertinentes dans la proposition finale. La théorie des normes et exploitations adhère le mieux au modèle catégorisant les Npréd : la conception de schémas réside dans l'aspect de la variation sémantique parmi les collocations avec le même type de Vsupp. La catégorisation s'accomplit par la conception de schémas ayant des significations différentes.

**Table 13 : Les Npréd qui ont une occurrence de >10**

Npréd	Traduction proposée par Van Dale ou fournie du corpus C558	Fréquences			total	catégorisation
		Ft.	C558	LN		
<b>erreur</b>	een vergissing maken	97	108	85	<b>290</b>	N
<b>faute</b>	een overtreding begaan, zondigen	84	8	77	<b>169</b>	A
<b>crime</b>	commettre un crime	44	44	27	<b>115</b>	N
<b>acte</b>	commettre des actes répréhensibles	11	44	12	<b>67</b>	N
<b>infraction</b>	(jur.) een strafbaar feit plegen	8	25	22	<b>55</b>	N
<b>délit</b>	commettre un délit	8	23	13	<b>44</b>	A crime
<b>péché</b>	commettre un péché/pécher	34	5	3	<b>42</b>	N
<b>attentat</b>	een aanslag plegen	3	12	16	<b>31</b>	A crime
<b>violation</b>	C: geweld gebruiken tegen, geweld plegen		21	2	<b>23</b>	A infraction
<b>meurtre</b>	een moord plegen/begaan	9	6	6	<b>21</b>	A crime
<b>vol</b>	diefstal (met braak) plegen	8	0	11	<b>19</b>	A crime
<b>imprudence</b>	een onvoorzichtigheid/onverstandigheid begaan	15	2	1	<b>18</b>	N
<b>violence</b>	iemand geweld aandoen (+SUR qqn)	2	2	13	<b>17</b>	N
<b>injustice</b>	ongerechtigheden aandoen	7	10	0	<b>17</b>	N
<b>abus</b>	C: misbruiken	4	10	2	<b>16</b>	N
<b>fraude</b>	fraude plegen bij een examen		14	2	<b>16</b>	A crime
<b>agression</b>	commettre un acte d'agression	1		15	<b>16</b>	E acte DE
<b>exaction</b>	C: gewelddaden plegen, misdrijf begaan	1	3	12	<b>16</b>	A violence
<b>atrocité</b>	C: verschrikkingen begaan	3	8	4	<b>15</b>	A violence
<b>action</b>	een slechte daad begaan, een lage daad begaan	7	5	1	<b>13</b>	A acte
<b>forfait</b>	een misdaad, misdrijf begaan	7	2	2	<b>11</b>	A crime

Dans le chapitre 3.2, les trois catégories abstraites N, A, et E ont été définies afin d'attribuer un statut provisoire à l'aide de la catégorisation des Npréd. Dans la colonne la plus à droite de la table 13, le statut de N (norme), A (alternance) ou E (exploitation) est attribué aux 21 Npréd les plus fréquents dans les corpus inventoriés. Comme la catégorisation appuie sur les caractéristiques non seulement dans la langue-source mais aussi dans la langue-cible, la deuxième colonne à droite montre une traduction proposée par Van Dale, à trouver dans la table 5 au chapitre 3.5, ou par défaut, par le corpus C558 dont les résultats se trouvent dans la table 11 au chapitre 5.1, marquée par C. Ainsi, la traduction au sens judiciaire de *commettre une infraction*, ici traduit par *een strafbaar feit plegen* est fournie par Van Dale, bien que le C558 montre plusieurs alternatifs, qui ne sont pas affichés mais qui sont bien pris en compte. Le support de l'analyse quantitative est inséré de manière qu'en général, le Npréd avec la fréquence la plus haute reçoit le statut de N : les Npréd dans la même catégorie reçoivent à l'instant le statut d'alternance et forment les alternances dans le schéma. De la liste ci-dessus, la colonne la plus à droite affiche neuf emplois normaux, ce qui est un assez grand nombre de N déjà. Il s'avère que les Npréd comme *abus*, *infraction*, *injustice* ne peuvent pas être insérées dans la catégorie de *crimes* sans ajouter de plus amples informations sur leurs caractéristiques sémantiques et que plusieurs Npréd au statut N sont pertinents dans l'entrée lexicale. Dans le chapitre 3.5.2, j'ai utilisé le test Zeugma pour analyser le verbe commettre et une des

conclusions était que le test peut être également utile à la catégorisation des alternances. Selon le test Zeugma, les Npréd du même type, adhérant au même schéma, peuvent subir la coordination. Ce test est donc applicable pour la à la sélection des alternances.

Les schémas sont affichés de manière structurelle, pour lequel j'ai adopté la méthode d'affichage de Hanks et al (2006), les entrées pour les verbes *leisten* et *erteilen* dans la section 3.2 mais également l'affichage dans la table 11 dans 5.1.

Pour chaque schéma, d'abord le Npréd prototypique X, le Npréd qui a reçu le statut de N, est affiché entre crochets : [X]. Afin de choisir la bonne traduction en langue cible, le Npréd qui correspond est affiché entre parenthèses (Y). Chaque Npréd néerlandais et immédiatement suivi de la traduction du Vsupp adéquate en néerlandais. Si nécessaire, les Npréd au français peuvent également être affichés entre parenthèses, entre autres pour fournir une traduction de la construction Vsupp par un verbe intransitif. Troisièmement, il se trouve des remarques syntaxiques entre parenthèses pour donner plus de contexte à un niveau plus abstrait. Comme la langue-cible est le néerlandais, les méta-informations sont en néerlandais. Les traductions peuvent reprendre des marques syntaxiques abrégées comme : *iets ADJ doen*, où ADJ est l'abréviation de *adjectief*. Il se peut aussi que la construction Vsupp est élargie vers un contexte avec un complément. Dans ce cas, la description est suivie par une traduction dans ce contexte. Un slash : A/B marque deux traductions identiques et échangeables dans le contexte du Npréd donné. Les alternances sont affichées entre accolades, séparées {A | B | C}. Pour un schéma avec beaucoup d'alternances fourni, j'ai choisi manuellement les alternances 'prototypiques' ce qui résulte dans une représentation compacte des alternances sur le Npréd prototypique.

La méthode d'affichage par Npréd résulte dans 14 schémas et un groupe de cas restants de traductions plus libres qui sont présentées ci-dessus. Il s'agit de l'ordre hiérarchisé dans lequel les schémas ultérieurs ont été définis par rapport aux schémas précédents des Npréd les plus fréquents. Bien que les propositions soient motivées par les résultats de recherche, il s'agit de jugements personnels sur la pertinence des résultats issus des corpus.

Bien que *erreur* et *faute* ont une distribution lexicale qui peut différer, je les ai mis ensemble dans un seul schéma. J'ai remarqué une tendance dans les résultats que *faute* seul peut s'étendre la signification de *overtreding begaan*. Il est difficile de saisir cette impression personnelle dans l'entrée lexicale du Vsupp d'une entrée bilingue et par conséquent, je n'ai pas distingué entre *faute* et *erreur* ici.

**Schéma1: [erreur, faute] (fout, vergissing) maken/begaan, (overtreding) begaan,**

*Commettre un crime* est la norme qui prend le plus grand nombre d'alternances et d'exploitations dans les résultats de la recherche que j'ai limitée à deux traductions, *plegen*, qui se combine avec tous les Npréd, et *begaan* que j'ai limité au Npréd néerlandais *misdaad* seulement. J'ai sélectionné entre les alternances pour les types de différents crimes et ajouté un alternatif intransitif pour qui a une occurrence pertinente pour l'insertion.

**Schéma 2 [crime] {délit | meurtre | vol | viol | larcin | trahison | adultère | fraude } plegen, (misdaad) begaan, (assassinat) begaan, moorden**

*Commettre un acte* est devenu la norme pour les Npréd *action* et *faits*. Le dernier Npréd, j'ai inséré au pluriel en raison de la fréquence haute du pluriel. Il faut remarquer que *faits* se trouve plus éloigné de *acte* au niveau sémantique et combinatoire : *faits* s'emploie souvent seul selon les données issues du corpus LN et est moins fréquent, tandis que *acte* a un grand nombre d'occurrences avec le complément +DE et par conséquent cette remarque sur la combinatoire est insérée. En outre, malgré les remarques faites sur l'emploi très général de *iets ADJ doen*, j'ai décidé de l'insérer parce que il représente une stratégie de traduction, se qui peut être utile au traducteur. Aussi, *faits* réfère plus souvent aux crimes commis (e.a. dans le domaine lexical de la presse) mais cela relève de l'emploi en contexte de *gepleegde feiten*.

Notre corpus de constructions Vsupp ne fournit pas d'exemples d'actes positifs et comme il s'est avéré dans une recherche élargie que les actes positifs sont peu fréquents pour l'insertion dans ce schéma. **Schéma 3: [acte] (+DE) (violence, agression) geweldddaden plegen, (+ADJ) ADJ daden begaan, verrichten, iets ADJ doen {action | faits } (goede daad) verrichten, gepleegde (feiten)**

J'ai distingué *infractio*n de la sous-catégorie *crime* en raison de sa nature différente : les significations dans cette catégorie ne sont pas limitées aux crimes (*infractio*n à la loi) mais constituent plusieurs domaines. Le Npréd *illégalité* avec la traduction contextuelle *plaatsvinden* est également inséré dans ce schéma. Parmi les traductions des constructions Vsupp dans C558, j'ai marqué *plaatsvinden* comme une stratégie de traduction, à utiliser à défaut d'une traduction plus adéquate. Cette stratégie d'une traduction neutre adhère à mon avis le mieux au schéma de *infractio*n, du fait que le N et les alternances insérées sont plus abstraits que les Npréd qui dénotent des crimes, fautes et destructions.

**Schema 4: [infractio] (wet) overtreden, (recht) schenden, (overtreding) begaan, in overtreding zijn, iets misdoen (+A) (droit) het recht schenden {violation | atteinte | illégalité} (onwetmatigheden) plaatsvinden.**

Les deux Npréd *péché* et *abus* constituent deux schémas à part. Le complément SUR pour le dernier ne change pas la signification du verbe en contexte mais indique le COD *iets/iem* et est par conséquent inséré en combinaison avec le verbe *misbruiken*.

**Schéma 5: [péché] (een) zonde begaan, zondigen (intrans)**

**Schéma 6: [abus] (misbruik) plegen, +DE misbruik maken van +SUR (iets/iem) misbruiken**

Comme on a vu que la ressemblance de *violence* et *violation* dépend du contexte des textes dans le corpus C558, les deux Npréd constituent un schéma chacun. Les Npréd qui dénotent des faits de violence qui sont presque toujours au pluriel sont par conséquent distingués des Npréd du type *crime*. En outre, les Npréd peu fréquents *abominations*, *cruautés* et *dégradations* sont des alternances qui n'ont pas été insérées.

**Schéma 7: [violences] (eerder geweld- dan misdaden, vaak mv) plegen (+SUR/+CONTRE) tegen {atrocités | exactions | horreurs } (verschrikkingen) begaan/plegen**

Les Npréd *imprudence* et *injustice* ne sont pas compatibles avec les catégories précédentes ni compatibles aux traductions des verbes précédents seuls : les traductions *ADJ zijn* et *iemand X aandoen* sont ajoutés :

**Schéma 8: [imprudence] onvoorzichtigheid begaan, onvoorzichtig zijn**

**Schéma 9: [injustice] onrecht begaan, iemand onrecht aandoen**

Pour les Npréd du schéma sémantique de *impair*, plusieurs traductions se présentent qui sont liées à la traduction du Npréd :

**Schéma 10: [impair] blunder begaan, blunderen (intrans) { bêtise | maladresse | bévue | gaffe} (bévue) een flater slaan**

Pour le schéma de *dégâts*, un schéma des Npréd au pluriel est constitué. Le deuxième critère est constitué par leur traduction par *aanrichten*. La traduction par le verbe *schaden* n'est pas pertinente puisqu'elle ne permet pas des compléments qui peuvent se présenter avec ce type de Npréd.

**Schéma11: [dégâts] (vaak mv) (schade, verwoestingen) aanrichten { destructions }**

Contraire à ce que fournit Van Dale, la traduction *verzuimen* pour *commettre un oubli* met ensemble le COD et le Vsupp et n'est donc plus intransif dans mon schéma, sinon le transitif est suivi par un complément non-obligatoire +DE.

**Schéma 12: [oubli] verzuimen (intrans), nalatig zijn (intrans)**

Le schéma de *zich schuldig maken aan* est emprunté du deuxième schéma de Van Dale et grâce à sa distribution, il constitue également une place dans la nouvelle entrée. L'argument en faveur de cette insertion est que le nombre d'occurrences et la diversité d'occurrences est non-négligeable. Dans l'ensemble des constructions Vsupp trouvées, une dizaine d'occurrences se trouvent dans la sphère d'un Npréd de *livre*.

**Schéma 13: [livre] zich schuldig maken aan het schrijven van { ouvrage | recueil | poème }**

Le Npréd prototypique *inconvenance* représente une catégorie de constructions Vsupp liées au comportement. Dans l'analyse de corpus monolingues, il s'est avéré que ce type de Npréd est surtout utilisé dans le domaine littéraire. La présence dans ces corpus a contribué à déterminer la pertinence pour insérer un schéma sur le contexte de comportement.

*Incorrection* et *impolitesse* ont aucune occurrence dans l'ensemble des corpus mais je ne les ai pas encore supprimé afin d'avoir plusieurs exemples.

**Schéma 14: zich ADJ gedragen [inconvenance] zich onbeleefd gedragen, onbeleefd zijn { incorrection | impolitesse }**

Finalement, il reste un ensemble de Npréd classifiés comme exploitation qui est trop grand pour afficher en entier. J'ai sélectionné des Npréd français qui introduisent de nouvelles traductions plus libres, mais qui peuvent constituer des suggestions pour le traducteur afin de traduire les Npréd qui ne ressemblent pas les exemples dans les schémas. J'ai librement choisi pour un ensemble limité dans cette catégorie, afin d'éviter une répétition de Vsupp néerlandais et de Npréd français du même type. Un Vsupp que je n'ai pas affiché en particulier, qui était pourtant très fréquent dans les données du corpus C558, est *plaatsvinden*. L'absence d'un agent en néerlandais ne correspond pas en général avec la transitivité du verbe *commettre* mais je l'ai finalement inséré dans le schéma 4.

**(excès) zich aan excessen te buiten gaan, (indiscrétion) uit de school klappen (idiom.), (contresens) een misvatting maken, (confusion) in verwarring brengen, (imprécisions) onnauwkeurig zijn**

Par contre, la traduction *zich schuldig maken aan* est trop fréquent pour être inséré en bas de l'hierarchisation et constitue une traduction qui vaut pour un nombre considérable de Npréd en construction Vsupp avec *commettre*. Dans l'entrée lexicale de Hanks, une traduction de base est affiché à la tête des schémas. Les constructions Vsupp des Npréd dans les schémas 1 à 3, 5 à 7, 10, et 12 à 14 sont compatibles avec la traduction *zich schuldig maken aan*. Par contre, les schémas 4, 8 et 9 ne le sont pas. Par conséquent, l'insertion de cette traduction de base n'est pas recommandée.

Le corpus LN en particulier fournit un grand nombre de Npréd qui peuvent être considérés comme des exploitations:

88. ...identités. Il suffit d'entrer son compte utilisateur Facebook et son mot de passe avant de **commettre le clic fatal** qui désactive votre compte. Et rejoindre le vaste réseau social des suicidés ...  
Libération, 21 décembre 2009
89. Au ministère de l'Environnement, la mezzo-soprano amatrice de Verdi **commit de bruyants couacs** sur le nucléaire et les intempéries avant que ses confidences sur le ...  
Les Echos, 26 juin 2008
90. Un scooter obstrue une piste cyclable et amène un vélo à **commettre une embardée**. Si le cycliste percute un piéton, le conducteur du scooter doit être mis ...  
Le Monde, 24 août 2008
91. S'il n'y a pas de recette pour **commettre un tube**, certains ingrédients ont quand même fait leurs preuves : un zeste de mélodie tout ...  
Le Figaro, 27 février 2010

La signification des constructions Vsupp avec ces Npréd peut être déduit des schémas définis.

Dans le chapitre 4.4., onze Npréd issus de Van Dale sont marquées par une zéro occurrence dans les corpus sélectionnées. On verra ici si ces exploitations peuvent être catégorisées dans un des schémas proposés afin de tester la catégorisation :

*Goujaterie* est un Npréd qui est au nom seul peu fréquent (neuf fois) dans un corpus du même période (1960-2000) de Frantext. Comme le paradigme de Npréd '*zich X gedragen*' se trouve dans le schéma 14, je propose ne pas mettre le Npréd ni dans l'entrée, ni comme hyperlien qui mène à l'entrée de *goujaterie*. *Enormité* fait partie du paradigme du schéma 10. Dans une version électronique un hyperlien aux Npréd dans cette catégorie aiderait à trouver la traduction *stommiteiten uithalen*.

*Coup d'état* peut être considéré comme une alternance et *indélicatesse* comme *trahison*, *plagiat* et *malhonnêteté* sont des exploitations sur *crime* (schéma 1). Le nombre de Npréd catégorisé sous *crime* est trop grand afin d'insérer ces Npréd si peu fréquents. *Impudicité* et *impolitesse* peuvent être catégorisées sous le schéma 14, *zich X gedragen*. Pour *irrégularité*, je propose la traduction '*in strijd met de wet handelen*' qui le rend semblable à *commettre une infraction à la loi* bien que ce Npréd ne prend pas de complément. Pourtant, il suffit quand le Npréd *infraction* est inséré dans l'entrée de *commettre*.

Quand les exemples utilisées dans 3.5.2. sont repris, il s'avère que les Npréd qui ne sont pas compatible à la coordination, appartiennent aux schémas différents :

92. a.\*Il commet un péché et des imprudences.
- b.\*Il commettait des violations et un oubli.
- c.\*Il commettait un impair et un vol.
- d. Il commet deux fautes et l'imprudence de demander après l'âge de l'interlocuteur.
- e. Il commet une bévue et une autre sottise
- f. Il a commis des violences et des vols.
- g. Il a commis un oubli et une inconvenance.

Pourtant, les exemples 19 et 20 mettent en doute s'il s'agit de schémas différents pour violence et vol et oubli et inconvenance, parce le test Zeugma est compatible, contraire à mes attentes.

Dans l'analyse par Npréd, j'ai évalué la possibilité de regroupement éventuel des schémas qui ont la même traduction *begaan* et *plegen*. On peut conclure que l'insertion de plus grands schémas ou même d'une traduction de base à la tête de tous les schémas constitue une piste de discussion intéressante. Il s'est avéré qu'un regroupement a pour conséquence une généralisation qui ne laisse pas de place pour l'insertion de remarques pour les traductions individuelles. Dans la sélection des exemples et la composition des schémas, j'ai aussi bien pris en compte les caractéristiques lexicales des Npréd que la traduction du Vsupp dans la langue-cible. Il s'avère qu'une catégorisation entièrement nominale est difficile à réaliser et qu'elle ne sert pas toujours l'objectif de la création d'une entrée verbale dans le dictionnaire. Par conséquent, comme l'accent dans ces schémas est mis sur les Npréd, les traductions du verbe peuvent être sous-représentées, en particulier les types de traductions comme les traductions de la construction Vsupp par un seul verbe. La question est comment les traductions par un verbe de ce type peuvent être affichées de manière plus claire. On passe donc à une deuxième analyse, celle par verbe.

### 5.3. Une proposition de catégorisation par verbe

Pour une catégorisation par traduction du verbe, la TNE est moins apte tandis que la méthode d'affichage dans le dictionnaire bilingue espagnol-néerlandais de Schroten (1999) donne la priorité aux verbes.



Premièrement, Schroten constate que la problématique pour les lexicographes bilingues réside dans la conception des propositions afin d'obtenir l'équivalence de traduction optimale. Du chapitre 3.2.3, il ressent que les possibilités combinatoires sont limitées dans le lexique des langues en question. (Schroten 2002 : 85) La syntaxe de la langue-source ne peut pas toujours être suivie dans la langue-cible. Il s'avère du chapitre 5.2 que les Vsupp posent en général le problème d'équivalence formelle, là où '*begaan, plegen*' en néerlandais ne suffit pas comme traduction de base aux constructions Vsupp avec le verbe *commettre*. Dans la discussion des résultats, la construction *zich schuldig maken aan* se présentait comme une alternative à *begaan/plegen*. Avec l'analyse par Vsupp, on peut nuancer les résultats de la section 5.2. et tester plusieurs traductions générales pour le Vsupp *commettre*.

Schroten introduit les premiers schémas selon leur traduction dans la langue-cible, à savoir *maken, doen, verrichten, in orde maken*. Dans ce chapitre je verrai comment les Npréd peuvent être catégorisés selon les Vsupp en néerlandais. Tout comme la catégorisation précédente, les jugements personnels ont mené aux traductions insérés dans **l'annexe 3 Catégorisation par Vsupp**. Les Npréd sont classés par traduction 'autre' dans la dernière colonne afin de montrer la variation verbale.

Les catégories examinées sont *begaan, plegen, aanrichten*, la catégorie *intransitif* (traduction par un verbe intransitif) et une catégorie désignée par *constr Vsupp -> V(supp)*, qui constitue les traductions qui prennent le Npréd et le Vsupp ensemble dans un verbe transitif :

93. Commettre un abus sur quelqu'un -> iemand misbruiken

Dans l'annexe 3, les traductions marquées par \* sont considérées comme des possibilités en combinaison avec le Npréd.

D'abord, *begaan* se présente comme une possibilité pour 46 Npréd, bien que *plegen* prenne un nombre considérable de 37 Npréd. Comme alternatif *aanrichten*, il s'avère que peu de verbes se catégorisent par ce verbe. Par contre, tout un groupe de Npréd prend cette traduction au lieu des verbes *begaan* ou *plegen* :

94. dégâts, dégradations, délits, déprédations, destructions, dévastations

Ce groupe correspond au schéma 11 et confirme donc la pertinence de ce schéma. La traduction *verrichten* prend cinq Npréd et se présente surtout comme une traduction dans les cas où *commettre* est suivi d'une action positive (avec *acte +DE, +DE +CONTRE, acte ADJ* ou *action*). Il pourrait former un schéma qui peut être sous-catégorisé dans le schéma 3.

Dans la catégorie des traductions par un verbe intransitif se trouvent sept verbes qui sont spécifiques pour le Npréd en question, à l'exception de *moorden* qui a deux Npréd :

95. *blunderen, discrimineren, frauderen, martelen, moorden, rommelen, zondigen*

Les Vsupp *gebeuren* et *plaatsvinden* n'ont pas été sélectionnés pour cette catégorie. Dans le corpus C558, *gebeuren* s'emploie avec *erreur*, ce que j'ai mis en doute, mais avant tout, ces verbes ne peuvent pas prendre un agent aussi bien qu'un thème. En effet, la traduction par un de ces verbes n'est pas optimale. Les verbes intransitifs s'emploient comme traduction optimale quand la construction Vsupp ne se combine pas avec des compléments. Les verbes *blunderen, discrimineren* et *frauderen* se prêtent en effet mieux à l'insertion comme alternatif aux constructions intransitives. .

Dans les chapitres 3.5 et 5.2, les variations sur *commettre une bétise* et ses traductions ont soulevé une discussion autour de la combinatoire dans le schéma 10. Si on prend la langue-cible comme point de départ, le néerlandais a les possibilités suivantes :

- 96. a. \*een blunder plegen
- b. Een blunder begaan
- c. een blunder maken
- d. blunderen

97. a. \*een fout plegen  
 b. een fout begaan  
 c. een fout maken  
 d. \*fouten
98. a. \*een stommiteit plegen  
 b. een stommiteit begaan  
 c. \*een stommiteit maken  
 d. een stommiteit uithalen  
 e. \*stommiteiten
99. a. \*een flater plegen  
 b. \*een flater begaan  
 c. \*een flater maken  
 d. een flater slaan  
 e. \*flateren

Les Vsupp néerlandais *plegen* n'est pas compatible avec les Npréd, *maken* se combine seulement avec *fout* et *blunder*, le verbe *begaan* ne se combine seulement pas avec *flater* et *uithalen* se combine uniquement avec *stommiteit*. Par conséquent, les combinaisons pertinentes pour l'affichage sont :

100. *blunder begaan, blunderen, (flater) slaan, (stommiteit) uithalen*

Ainsi, l'affichage montre toutes les possibilités et variations en néerlandais. L'insertion du dernier Npréd, *stommiteit*, dépend des fréquences. Le Vsupp *maken* est inséré sous le premier schéma de *faute et erreur*. Cette constatation correspond au schéma 10, à l'exception de *uithalen*. La construction Vsupp *stommiteit(en) uithalen* ne se trouve pas dans le corpus C558 mais en raison de sa tendance idiomatique en néerlandais, l'insertion de cette construction Vsupp aura néanmoins une contribution à l'entrée lexicale.<sup>47</sup>

Comme on a vu dans le chapitre 3.2.3. dans l'entrée lexicale conçue par Schroten du verbe *hacer*, la méthode d'affichage met ensemble les constructions Vsupp qui se traduisent par un verbe seul. La catégorie qui est définie par Schroten par '*in combinatie m e [met een] lijdend voorwerp als één werkwoord vertaald*' :

101. *hacer burla - bespotten, hacer caricias - liefkozen*

Tout comme Schroten, les constructions Vsupp qui se traduisent par un seul verbe sont rassemblées sous la colonne *constr Vsupp-> V(supp) transitif*. Cette catégorie n'en prend que 4 résultats :

102. *commettre des discriminations -discrimineren, commettre des tortures - martelen, commettre des abus - misbruiken, commettre des dégradations - schaden*

Dans les schémas du chapitre 5.2, seulement *misbruiken* a été inséré. Pour l'insertion de (*iemand*) *discrimineren*, il manque le support de données quantitatives mais aussi qualitatives : les occurrences issues du corpus 558 ne prennent pas de complément, qui se traduirait en position COD en néerlandais. Les Npréd *torture* et *dégradations* sont donc les deux des alternances trop peu fréquentes.

Dans la catégorie 'autres' se trouvent entre autres les traductions *zich schuldig maken aan het schrijven van* (classifié dans le schéma 13), *maken* (schéma 1), *ADJ gedrag vertonen* (qui peut être ajouté au schéma 14), *ADJ zijn* (onvoorzichtig, nalatig zijn) mais aussi *zich schuldig maken aan* qui se prête à un très grand nombre des Npréd. Au moins, cette construction Vsupp devrait être insérée à une place plus visible dans l'entrée lexicale. La question qui se pose après cette analyse des résultats pour *zich schuldig maken aan* est de savoir s'il faut

<sup>47</sup> Pour vérifier la fréquence de *stommiteit(en) uithalen*, il faudrait un corpus monolingue au néerlandais comparable aux corpus français.

insérer une traduction de base pour *commettre*. Après l'analyse de ce chapitre et une évaluation des résultats, je suis d'avis que pour *commettre*, il ne faut pas insérer une traduction de base à la tête de l'entrée. Au niveau de la distribution de *begaan*, *plegen* et *zich schuldig maken aan*, un affichage général constituerait un affichage inconsistant par rapport aux Npréd qui se traduisent autrement en collocation avec le Vsupp. La nature du verbe *commettre* s'avère plus légère dans cette recherche qu'on a attendue au préalable. En outre, la catégorisation dans les schémas a pour conséquence un affichage plus clair et logique qui prend la hiérarchisation selon la fréquence décroissante des Npréd. Pour les emplois normaux, le traducteur n'ayant pas le néerlandais comme langue maternelle peut bénéficier d'une description plus détaillée sans que le sens général interfère avec les suggestions fautives de combinatoires et de la traduction.

Afin de concevoir la proposition finale pour une nouvelle entrée lexicale, il relève des remarques que la catégorisation par Npréd peut le mieux servir de base de la nouvelle entrée lexicale : le regroupement dans les schémas a été confirmé, comme pour le schéma 11, testé, comme le schéma 10 et nuancé, comme il s'avère des remarques sur les verbes intransitifs. La catégorisation par traduction du verbe ne s'avère pas le meilleur moyen de catégorisation mais constitue un bon outil pour vérifier la cohérence des schémas des Npréd semblables. L'analyse verbale a finalement constitué l'argument décisif contre l'insertion de trois traductions de base à la tête de la nouvelle entrée lexicale.

En outre, l'analyse par Vsupp ajoute encore trois points d'adaptation. Premièrement, plusieurs verbes intransitifs pourraient être insérés mais leur fréquence basse dans l'ensemble des corpus ne suffit pas pour la sélection. Pourtant, la construction Vsupp n'est pas suivie par un complément, ce type de traduction est plus spécifique qu'une traduction avec *begaan*, *plegen*, *verrichten*, ou autre et a donc la priorité sur une traduction plus générale. Deuxièmement, les traductions 'libres' se sont avérées utiles pour donner plus de suggestions au traducteur. Les exemples en bas de l'entrée lexicale ont été sélectionnés à partir de jugements personnels : les traductions insérées contribuent à l'ensemble en proposant une nouvelle traduction du Vsupp. Troisièmement, dans la conception finale, il faut veiller à un affichage clair et le plus simple possible. Les méta-informations doivent guider l'utilisateur, et visent l'emploi de la construction Vsupp dans la langue-cible. Par conséquent, toutes les méta-informations seront donc affichées en néerlandais.

#### 5.4. La nouvelle entrée lexicale

##### Commettre (intransitief)

1

[erreur, faute] (fout, vergissing) maken/begaan, (overtreding) begaan

2

[crime] {délit | meurtre | vol | viol | larcin | trahison | adultère } plegen, (misdaad) begaan, zich schuldig maken aan, (assassinat) moord begaan, moorden, (fraude) frauderen (intrans), (discriminations) discrimineren (intrans)

3

[acte] (+DE) (violence, agression) gewelddaden plegen, (+ADJ) ADJ daden begaan, verrichten, iets ADJ doen {action | faits}(goede daad) verrichten, gepleegde (feiten)

4

[infraction] (wet) overtreden, (recht) schenden, (overtreding) begaan, in overtreding zijn, iets misdoen (+A) (droit) het recht schenden {violation | atteinte | illégalité} (onwetmatigheden) plaatsvinden

5

[péché] (een) zonde begaan, zondigen (intrans)

6

[abus] (misbruik) plegen, +DE misbruik maken van +SUR (iem) misbruiken

7

[violences] (eerder geweld- dan misdaden, vaak mv) plegen (+SUR/+CONTRE) tegen {atrocités | exactions | horreurs } (verschrikkingen) begaan, plegen

8

[imprudence] onvoorzichtigheid begaan, onvoorzichtig zijn

9

[injustice] onrecht begaan, (iemand) onrecht aandoen

10

[impair] blunder begaan, blunderen (intrans) { bêtise | maladresse | bévue | gaffe} (bévue) een flater slaan, (énormités) stomiteiten uithalen

11

[dégâts] (vaak mv) (schade, verwoesting) aanrichten {destructions }

12

[oubli] verzuimen (intrans), nalatig zijn (intrans)

13

[livre] zich schuldig maken aan het schrijven van {ouvrage | recueil | poème}

14

zich ADJ gedragen [inconvenance] zich onbeleefd gedragen, onbeleefd zijn {incorrection | impolitesse}

(excès) zich aan excessen te buiten gaan  
(indiscrétion) uit de school klappen  
(contresens) een misvatting maken  
(confusion) in verwarring brengen  
(imprécisions) onnauwkeurig zijn

## 5.5 Conclusion

Dans ce mémoire, j'ai élaboré aussi bien une analyse théorique sur la linguistique de corpus en relation avec la traductologie, qu'une recherche apte à mon sujet, les constructions à verbe support. Les deux questions principales étaient de savoir de quelle manière, l'analyse de corpus peut contribuer à la recherche empirique dans le domaine de la traductologie et comment les constructions à verbe support peuvent-elles être mieux affichées dans un dictionnaire bilingue. L'analyse de corpus exécuté après le verbe support *commettre* a constitué le facteur liant les deux volets.

Des deux premiers chapitres, il s'est avéré que le domaine de l'analyse de corpus est caractérisé par de nombreux nouveaux champs de recherche qui sont étroitement liés quant aux intérêts et au développement de la méthodologie. L'encadrement de ma recherche dans le domaine de la traductologie ne peut explicitement pas être considéré comme une analyse de corpus traditionnelle en raison de l'application lexicographique des résultats et se situe par conséquent dans le domaine appliqué de la traductologie, plutôt que des domaines théoriques de la traductologie. En exécutant cette recherche, j'ai voulu porter attention à cette filière interdisciplinaire qui peut bénéficier des approches linguistiques pour se développer comme domaine de recherche.

Comme ma deuxième question de recherche visait l'amélioration de l'affichage de l'emploi des constructions à verbe support dans un dictionnaire bilingue, l'objectif final constituait la conception d'une nouvelle entrée lexicale pour un dictionnaire bilingue français-néerlandais, résultant dans une analyse en plusieurs étapes.

Dans le premier chapitre, le phénomène de verbes support (*light verbs*) a été délimité et encadré dans le contexte de ma propre recherche aux objectifs lexicographiques. Une analyse de la recherche dans les deux filières a résulté dans une préférence pour une méthode de recherche empruntée à la linguistique contrastive.

Le deuxième chapitre est consacré à l'avènement de l'analyse de corpus. Il en sort que la nouvelle application de l'analyse de corpus dans la lexicographie exige également une nouvelle méthodologie qui est adaptée aux objectifs spécifiques. Contraire à une analyse de corpus visant les influences de la traduction, une analyse de corpus de textes dépourvus d'un statut de langue-cible ou de langue-source, à savoir des versions des textes de contenu identique en français et en néerlandais, constituait la meilleure choix pour la recherche visée

La conception de cette méthodologie qui combine la linguistique contrastive, la traductologie théorique et les objectifs lexicographiques a constitué la première partie du chapitre 3. L'entrée lexicale de *commettre* dans le dictionnaire électronique Van Dale 2010 français-néerlandais est soumise à une analyse détaillée et résulte dans un nombre considérable d'améliorations afin de mieux afficher les possibilités combinatoires des verbes support. En catégorisant les constructions à verbe support d'une entrée lexicale dans un dictionnaire bilingue, de trois corpus monolingues de registres divers et d'un corpus parallèle bilingue, les résultats sont analysés et comparés dans le chapitre 4. En outre, l'analyse quantitative ajoute une dimension importante aux résultats permettant d'y concevoir une entrée lexicale plus structurée et hiérarchisée par fréquence décroissante.

Comme projet pilote, une nouvelle entrée lexicale pour un dictionnaire bilingue est élaboré pour le verbe *commettre* dont le résultat final se trouve dans la section 5.4. Le corpus bilingue ayant un grand nombre de mots (29 millions) a fourni une variété modérée d'emplois de la construction Vsupp avec *commettre*. Une nouvelle proposition pour l'entrée lexicale dans un dictionnaire est finalement conçue à la base des conclusions tirées de l'analyse de Van Dale, les données et exemples des corpus mais également de jugements personnels afin de concevoir les schémas : avec un corpus de constructions Vsupp d'au total 1286 constructions

Vsupp, les résultats des corpus monolingues compris, le support des corpus n'a finalement pas toujours constitué le facteur décisif dans la sélection pour l'insertion dans l'entrée lexicale.

## **Bibliographie**

### **Bibliographie primaire**

- Baker, M. 'Corpora in translation studies: an overview and suggestions for future research', *Target* 7, 1995, John Benjamins Publishing Company: Amsterdam/Philadelphia p. 223-243
- Corbin, P. Gasiglia, N., « Changer les dictionnaires ? Une pluralité d'approches », *Lexique* 19 « Changer les dictionnaires ? », Pierre Corbin & Nathalie Gasiglia eds., 2009 b, p. 7-38.
- Fellbaum, C. (eds) *Idioms and collocations*. Clippingham, Wiltshire: CI Antony Rowe, 2007.
- Gaatoone, D. 'Ces insupportables verbes supports : le cas des verbes événementiels', *Linguistica Investigationis*, 27 :2 John Benjamins Publishing Company (2004) p. 239-252.
- Granger, S. 'The corpus approach : a common way forward for Contrastive linguistics and translation studies?' *Corpus-based approaches to contrastive linguistics and translation studies*. Granger, S. et al. Rodopi, Amsterdam (2003) p.17-30
- Gross, G. 'Introduction', *Linguistica Investigationis*, 27 :2 John Benjamins Publishing Company (2004) p. 167-170
- Hanks, P. Urbschat, A., Gehweiler, E. 'German light verb constructions in corpora and dictionaries' *International Journey of Lexicology*, 2006, p 439-457.
- Hanks, P. 'The syntagmatics of metaphor and idiom', *International Journal of Lexicography*, 2004, p.245-272.
- Hanks, P. and Pustejovsky, J. 'A pattern dictionary for Natural Language Processing', *Revue Française de Langue appliquée*, (2005), 10 :2
- Holmes, J.S. 'Wat is vertaalwetenschap?' *Nederlandse vertalingen en uitbreiding in: Denken over vertalen*, Amsterdam: Van Tilt, 2004. Origineel: 'The name and nature of translation studies' in *Translated ! Papers on translation and translation studies*. Amsterdam: Rodopi, 1988 [1972].
- Johansson, S. 'Contrastive Linguistics and corpora' *Corpus-based approaches to contrastive linguistics and translation studies*. Granger, S. et al. Rodopi, Amsterdam (2003) p. 31-44
- Langer, S. 'A linguistic test battery for delimiting support verb constructions', *Linguistica Investigationis*, 27 :2 John Benjamins Publishing Company (2004) p. 171-184
- Laviosa, S. 'Corpora and translation studies' *Corpus-based approaches to contrastive linguistics and translation studies*, Granger, S. et al. Rodopi, Amsterdam (2003) p. 45-56
- Sag, I.A., Baldwin, T., Bond, F., Copestake, A, Fickinger, D. 'Multiword expressions: a pain in the neck for NLP', *CICling 2002, LNCS 2276*, p. 1-15.
- Salkoff, M. 'Automatic translations of support verb constructions', 1990,p. 243-246.  
[www.aclweb.org/anthology/C/C90/C90-3043.pdf](http://www.aclweb.org/anthology/C/C90/C90-3043.pdf)
- Schroten, J. 'Equivalence and mismatch of semantic features' *Evidence for linguistic relativity*. (1999b) Eds. Niemeier, S. and Dirven, R. Amsterdam/Philadelphia: John Benjamins, p. 29-51.

Schroten, J. 'Light verb constructions in bilingual dictionaries' Eds. Melka, F. Augusto, C. From lexicology to lexicography, Utrecht, UiL OTS, 2002, 83-94.

Storrer, A. 'Corpus-based investigations on German support verb constructions' Idioms and collocations, eds. Fellbaum, C. CPI Antony Rowe, Wiltshire (2006). p.164-187.

Oholan, M. Introducing corpora in translation studies, Oxfordshire: Routledge, 2004.

### **Bibliographie secondaire**

Barlow, M. 'ParaConc and Parallel Corpora in Contrastive and Translation Studies'. P.9-16, 2008.

Gellerstam, Martin. "translationese in Swedish novels translated from English", in Lars Wollin & Hans Lindquist, Translation Studies in Scandinavia, CWK Gleerup, Lund, 1986, pp. 88-95.

Johansson, S. Loving and hating in English and Norwegian: A corpus-based contrastive study. Perspectives on foreign and second language pedagogy. Essays presented to Kirsten Haastrup on the occasion of her sixtieth birthday, Albrechtsen et al. (eds) 93-103. Odense: Odense University Press.

Johansson S., and Hofland, K., Towards an English-Norwegian parallel corpus. Dans: U. Fries, Editor, Creating and Using English Language Corpora, Rodopi, Amsterdam (1994), pp. 25-37.

Lindquist, H. English adverbials in translation: A corpus study of Swedish Renderings. Lund studies in English 80. Lund: Lund University Press, 1989.

Mel'cuk, I. 'Verbes supports sans peine', *Linguisticae Investigationis*, 27 :2, 2004, p. 203-228.

Mel'cuk, I. Lexical Functions : A tool for the description of lexical relations in a lexicon. Dans Wanner, L.(ed) *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1996.

Sinclair, J. 1991. *Corpus, concordance, collocation: Describing English language*. Oxford: Oxford University Press.

Toury, G. *Descriptive translation studies and beyond*. Amsterdam/Philadelphia: John Benjamins, 1995

Hansard Corpus: <http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC95T20>

### **Dictionnaires**

Van Dale Onlinewoordenboeken 2010 frans-Nederlands Nederlands-Frans  
<http://www.uu.nl/NL/Bibliotheek/vakgebieden/geesteswetenschappen/frans/Pages/default.asp>

Van Dale Groot Handwoordenboeken Frans-Nederlands en Nederlands-Frans 2008, 4e druk, red. Bogaards, P./Beerden, A.G.M.

Van Dale Groot Woordenboek Frans-Nederlands en Nederlands-Frans 2000, 3e druk, red Bogaards, P. / Beerden, A.G.M.



Schroten, J. (m.m.v. D. Bloemraad, M. Buursma, N. Gerbrandy, T. Ceelen, J. Hulst, J. Boone, B. van Maris, M. Inklaar). *Spaans-Nederlands woordenboek*. Muiderberg, Coutinho, 1989, xi+807 pags.

### **Corpus et bases textuelles**

Base textuelle Frantext ATILF (Analyse et traitement informatique de la langue française, Frantext catégorisé. (2010)  
[www.frantext.fr](http://www.frantext.fr)

LexisNexis academic NL (2010)  
<http://academic.lexisnexis.nl>

Europarl Parliament Proceedings Parallel Corpus 1996-2009 Version V5, lancé le 20 Janvier 2010.  
<http://www.statmt.org/europarl/>

### **Terminologie**

D. Gile, 'Présentation traductologie' (2009)  
<http://www.cirinandgile.com/ESIT%20home.htm>

'The French/English glossary of linguistic terms' (2010)  
[http://www.sil.org/linguistics/Glossary\\_fe/glossary.asp?entryid=12213](http://www.sil.org/linguistics/Glossary_fe/glossary.asp?entryid=12213)

## Annexe 1 : l'entrée lexicale *commettre* dans Van Dale

1a: L'entree lexicale Van Dale pour *commettre*

### **commettre**<sup>1</sup> (ww)

commettre<sup>1</sup> (overgankelijk werkwoord)

Begaan, plegen, bedrijven, zich bezondigen aan

[▲ context](#)

**commettre un attentat** een aanslag plegen

<sup>1</sup>**commettre un délit** een strafbaar feit begaan/plegen

**commettre une erreur** een fout begaan

**commettre un péché** een zonde begaan

**commettre une trahison** verraad plegen

Aanstellen, benoemen

[▲ context](#)

**commettre qqn. à un emploi** iemand in een functie benoemen

**commettre qqn. au soin de qqch.** iemand met de zorg voor iets belasten

<sup>2</sup>**commettre un huissier** een deurwaarder benoemen

**avocat commis d'office** toegevoegd advocaat

**commettre qqn. d'office à** iemand toewijzen aan(archaïsch)

**commettre qqn. pour un travail** iemand een werk opdragen

<sup>3</sup>(archaïsch)toevertrouwen

<sup>4</sup>(scheepvaart) touwslaan(kabel) slaan, stranden, twijnen

(schertsend)zich bezondigen aan het schrijven van

<sup>5</sup>[▲ context](#)

**commettre sa réputation** zijn goede naam in gevaar brengen

<sup>6</sup>(archaïsch)riskeren, in gevaar brengen

© 2010 Van Dale Uitgevers

### **commettre**<sup>2</sup> (ww)

se commettre<sup>2</sup> (wederkerend werkwoord)

zich compromitteren

<sup>1</sup>[▲ context](#)

**se commettre avec qqn.** zich met iemand compromitteren, zich met iemand inlaten

<sup>2</sup> begaan worden

## 1b : Données de l'option de recherche *binnen artikelen*

Comme il est difficile de montrer les résultats de la version électronique de Van Dale, les occurrences 'binnen artikelen' sont ici affichés en Word. Les données sont insérées dans le schéma 1 'Occurrences Npréd commettre'.

### Résultats français pour commettre 'binnen artikelen'

Dans le schéma suivant les Npréd sont fournis de leur signification de base et des collocations trouvés par l'option 'binnen artikelen' de Van Dale. En plus, Van Dale distingue seulement entre signification générale, idiomes, et des significations en contexte. Les chiffres réfèrent à la signification en question. Aucune hypothèse sur l'ordre d'affichage dans Van Dale ne peut être confirmée.

#### Npréd

##### action

Daad, handeling

▲ context

1 une bonne action een goede daad

**commettre une mauvaise action** een slechte daad begaan, een misstap doen  
action d'éclat schitterende daad

##### attentat

Aanslag

▲ context

1 attentat à la pudeur aanranding van de eerbaarheid, ontucht(ige handeling)

**commettre un attentat contre qqn.** een aanslag op iemand plegen

##### barbarie

Wreedheid, barbaarsheid

3 ▲ context

**commettre un acte de barbarie** een wreedheid begaan, een onmenselijke daad verrichten

##### bassesse

lage daadslaafse daad

2 ▲ context

**commettre une bassesse** een gemene daad, laagheid begaan

faire des bassesses à qqn. iemand de hielen likken

##### bévue

flaterbok, domme streek, stommiteit, blunder

1 ▲ context

**commettre une bévue** een flater slaan

##### confusion

verwisseling (van twee personen, zaken)

4 ▲ context

**commettre une grossière confusion** de zaken ernstig door elkaar halen

##### énormité

stommiteitflater, enormiteit

1 ▲ context

**commettre des énormités** stommiteiten uithalen, flaters slaan

dire des énormités stomme opmerkingen maken

erreur

vergissingfout, misrekening, abuis, dwaling

▲ context

1 erreur de calcul rekenfout, telfout, misrekening

cherchez l'erreur er klopt iets niet

**commettre/faire une erreur** een vergissing maken

étourderie

onbezonnen streek, domheid

2 ▲ context

**commettre une étourderie** iets doms doen

faute

Vergrijp, fout, overtreding (ook sport)zonde, misstap (ook letterlijk)

1 ▲ context

**commettre/faire une faute** een overtreding begaan, zondigen

félonie

trouwbreuk (van vazal)verraad

1 ▲ context

**commettre un acte de félonie** trouwbreuk plegen

forfait

Euveldaad misdaad, misdrijf, zwaar vergrijp

5 ▲ context

**commettre un forfait** een misdaad begaan, een misdrijf begaan

fraude

Fraudebedrog

▲ context

fraude électorale verkiezingsfraude

en fraude frauduleus, met ontduiking van de wet, zwart(figuurlijk) heimelijk

1 entrer en fraude binnengesmokkeld worden

prendre qqn. en fraude iemand betrappen op bedrog

marchandise qui passe la frontière en fraude koopwaar die over de grens gesmokkeld wordt

**commettre une fraude dans un examen** fraude plegen bij een examen

gaffe

flaterstommititeit, blunder, miskleun

1 ▲ context

**commettre/faire une gaffe** een stommititeit begaan, een bok schieten, blunderen

goujaterie

Ongemanierdheid, onbeschoftheid, ploertigheid

1 ▲ context

**commettre une goujaterie** lomp optreden, zich onhebbelijk gedragen

horreurs

Gruwelikheden, verschrikkingen, wreedheden

1 ▲ context

**commettre des horreurs** afschuwelijke dingen doen

Impair1

1 Flater, blunder, stommititeit

- ▲ context  
**commettre un impair** een flater slaan, een bok schieten
- impolitesse  
 Onbeleefdheid, ongemanierdheid  
 1 ▲ context  
**commettre une impolitesse** een onbeleefdheid begaan
- imprudence  
 Onvoorzichtigheid, onverstandigheid, onbezonnenheid  
 1 ▲ context  
**commettre une imprudence** een onvoorzichtigheid/onverstandigheid begaan
- impudicité (formeel)  
 onkuisheidonzedigheid, obsceniteit  
 1 ▲ context  
**commettre des impudicités** onkuisheden begaan
- inceste  
 Bloedschande, incest  
 1 ▲ context  
**commettre un inceste** incest plegen
- inconvenance  
 Ongepastheid, onbehoorlijkheid, onbetamelijkheid, onbeleefdheid, onwelvoeglijkheid, indecentie  
 1 ▲ context  
**commettre une inconvenance** een onbeleefdheid begaan
- incorrection  
 Ongepastheid, onbehoorlijkheid, onbeleefdheid  
 2 ▲ context  
**commettre une grave incorrection** heel onbeleefd/onbehoorlijk zijn
- indélicatesse  
 Oneerlijkheid  
 2 ▲ context  
**commettre une indélicatesse** bedrog plegen, oneerlijk handelen
- indiscrétion  
 Loslippigheid,praatzucht  
 2 ▲ context  
**commettre une indiscrétion** uit de school klappen
- infraction  
 (juridisch)strafbaar feitvergrijp, (wets)overtreding  
 2 ▲ context  
**commettre une infraction** een strafbaar feit plegen
- irrégularité  
 onregelmatigheidonwettigheid, wederrechtelijkheid, onrechtmatigheid  
 2 ▲ context  
**commettre une irrégularité** in strijd met de wet handelen

larcin

Kruimeldiefstal

1 ▲ context

**commettre/faire un larcin** een (kruimel)diefstal begaan

maladresse

Stommiteit, uitglijder

2 ▲ context

**commettre une maladresse** een flater begaan

malhonnêteté

Oneerlijkheid

1 ▲ context

**commettre une malhonnêteté** iets oneerlijks doen, bedrog plegen

meurtre

Moorddoodslag

▲ context

1 crier au meurtre moord en brand schreeuwen

meurtre avec préméditation moord met voorbedachten rade

**commettre un meurtre** een moord plegen/begaan

oubli

Verzuim, nalatigheid, leemte, weglating, lacune

2 ▲ context

il y a des oublis dans son récit zijn verhaal bevat/vertoont leemten

**commettre/faire un oubli** iets verzuimen (te doen)

sottise

dwaasheidstommiteit, domme streek

▲ context

1 avoir la sottise de (+ onbepaalde wijs) zo dom/dwaas zijn om

**commettre une sottise** een stommiteit uithalen, een dwaasheid begaan

dire des sottises onzin uitslaan, raaskallen

faire une sottise een stommiteit uithalen, een dwaasheid begaan

trahison (v.)

Verraad

1 ▲ context

**commettre une trahison** verraad plegen

turpitude (formeel of ironisch)

schandelijk/schaamteloos gedraglaagheid

2 ▲ context

**commettre des turpitudes** zich schaamteloos gedragen

violence

gewelddwang(meervoud) gewelddadigheden

▲ context

avoir recours à la violence zijn toevlucht nemen tot geweld

céder à la violence voor geweld zwichten(ironisch)

1 se faire une douce violence gaarne zwichten voor de aandrang

faire une douce violence à qqn. iemand met zachte dwang ergens toe brengen

se résoudre à employer la violence besluiten om geweld te gebruiken

se faire violence zichzelf geweld aandoen

faire violence à qqn. iemand geweld aandoen(archaisch)

faire violence à une femme een vrouw verkrachten  
film de violence geweldsfilm  
répondre à la violence par la violence geweld met geweld beantwoorden  
scènes de violence gewelddadige/brute scènes  
commettre des violences **sur** qqn. iemand geweld aandoen  
violence verbale verbaal geweld  
violence aveugle/gratuite zinloos geweld

### Résultats néerlandais pour commettre ‘binnen artikelen’

La première sélection s’effectuée par catégorie grammaticale. Ensuite, une recherche après les occurrences du verbe commettre ‘binnen artikelen’ montre que le dictionnaire donne des exemples ‘en contexte’ qui ne figurent pas dans le schéma 1. Par conséquent, l’affichage entière des exemples en contexte est reprise pour les résultats néerlandais. Les verbes en néerlandais portent d’abord une signification spécifique.

Aandoen

Berokkenen

faire

>▲ context

<sup>2</sup> dat kun je haar niet aandoen! tu ne peux pas lui faire cela!  
iem. onrecht aandoen **commettre une injustice envers qqn.**  
iem. een proces aandoen faire, intenter un procès à qqn.  
iem. verdriet aandoen faire de la peine à qqn.  
zich iets aandoen attenter à sa vie

zich afgeven<sup>3</sup> (wederkerend werkwoord)

zich inlaten(m.b.t. personen) **se commettre ≥ (avec)**(m.b.t. zaken) s’ingérer ≥ (dans)

1▲ context

zich afgeven met allerlei gespuis s’encanailler

bedrijven

faire ≥accomplir ≥, pratiquer ≥(pejoratief) commettre ≥, perpétrer ≥

▲ context

kwaad, onheil bedrijven faire le malcauser un désastre, faire un malheur

liefdadigheid bedrijven faire, pratiquer la charité

de ruitersport bedrijven faire du cheval, de l’équitation

struisvogelpolitiek bedrijven pratiquer la politique de l’autruche

zonde bedrijven **commettre le péché, pécher**

begaan<sup>3</sup> (overgankelijk werkwoord)

bedrijven

<sup>1</sup> commettre ≥accomplir ≥, consommer ≥, faire ≥

▲ context

(figuurlijk) een misstap begaan se fourvoyer, **commettre une faute grave**

benoemen (overgankelijk werkwoord)

<sup>1</sup> Aanstellen

nommer ≥appeler ≥, désigner ≥, faire ≥(juridisch) **commettre** ≥

blunderen (onovergankelijk werkwoord)

<sup>1</sup> **commettre ≥ une bétvue** gaffer ≥, faire ≥ une gaffe, mettre ≥ les pieds dans le plat

echtbreken (werkwoord, alleen onbepaalde wijs)

commettre ≥ un adultère

1 ▲ context

het echtbreken l'adultère

Klappen

Doorvertellen

rapporter ≥ bavarder ≥

4 ▲ context

uit de school klappen jaser, bavarder, commettre une indiscretion  
(België) uit de biecht klappen vendre la mèche

misdoen2 (overgankelijk werkwoord, ook absoluut)

zondigen(onovergankelijk)

commettre ≥ des actes répréhensibles accomplir ≥ de mauvaises actions, pécher

1 ≥ (overgankelijk) commettre ≥ accomplir ≥, consommer ≥

▲ context

wat heb ik misdaan, dat ... qu'est-ce j'ai fait au bon Dieu pour que ... (+ aanvoegende wijs)

misdrijven (overgankelijk werkwoord)

1 commettre ≥ (un méfait, un délit, un crime, une faute)

moorden (onovergankelijk werkwoord)

commettre ≥ des assassinats

1 ▲ context

erop los moorden semer la mort

plagiëren1 (onovergankelijk werkwoord)

1 plagiaat bedrijven

commettre ≥ un plagiat

plegen2 (overgankelijk werkwoord)

(iets ongeoorloofds) bedrijven

1 commettre ≥ perpétrer ≥

▼ context

doen, verrichten faire ≥

▲ context

een aanslag plegen op attenter à la vie de (qqn.), braquer (une banque)

2 diefstal met braak plegen commettre un vol avec effraction

een gewapende overval plegen op

attaquer (qqn., qqch.) à main armée, braquer (qqch.)

terreurdaden plegen se livrer à des actes de terrorisme

saboteren2 (overgankelijk werkwoord, ook absoluut)

1 sabotage plegen commettre ≥ un acte de sabotage saboter ≥ (qqch.)

Uithalen

Uitvoeren faire ≥ commettre ≥

3 ▲ context

streken uithalen faire des siennes

zich vergalopperen (wederkerend werkwoord) (figuurlijk)

(mond voorbijpraten) ± se trahir ≥

1 (blunder) gaffer ≥, commettre ≥ une bévue

1 (overijld handelen) agir ≥ sans réfléchir

▲ context



zich aan iets vergalopperen aller trop vite en besogne

zondigen (onovergankelijk werkwoord)

zonde begaan **commettre ≥ un péché pécher ≥**

**▲ context**

<sup>1</sup> tegen de tien geboden zondigen pécher contre les dix commandements

zwaar zondigen pécher gravement

### Résultats pour Npréd possibles

Plusieurs Npréd possibles demandent une analyse de plus proche. Par conséquent, l’affichage est reprise aussi pour les noms qui résultent d’une recherche ‘binnen artikelen’.

### Noms non-Npréd

Parfois, le dictionnaire réfère aussi aux compléments de verbes ou aux adjectifs dans ‘binnen artikelen’ :

Agressie

agressie plegen **commettre un acte d’agression (contre)**

laag

Gemeen

(bijvoeglijk naamwoord) bas, vil, méprisable

(bijwoord) bassement, vilement

<sup>6</sup> **▲ context**

lieden van het laagste allooi individus de bas etage

een lage daad begaan **commettre une bassesse, une action vile**

Geweldig

enorm, reusachtig

formidable énorme, gigantesque, colossal

**▲ context**

een geweldig applaus un tonnerre, une tempête d’applaudissements

een geweldig bedrag une somme colossale, faramineuse

leen geweldige eetlust un appétit effrayant, de gargantua, gargantuesque

een geweldig getier un vacarme terrible ↓ un boucan du diable, de tous les diables

een geweldig pak slaag une correction magistrale, ↓ une volée magistrale, une fessée

monstre

een geweldige persoonlijkheid une très forte personnalité, une nature

geweldige stommiteiten uithalen **commettre des énormités**

Les idiomes se trouvent également dans cette liste :

daderschap (het)

¶ Idioom

¶ iem. tot het daderschap leiden **amener qqn. à commettre un crime**

dievenpad (het)

¶ Idioom het dievenpad opgaan **se mettre à voler, à commettre des vols**

¶ Idioom uit de school klappen jaser, bavarder, **commettre une indiscretion**

## Annexe 2: occurrences, traductions et statuts des Npréd

Npréd	source	remarques	traduction vD	traduction autre	statut	occ. ds Ft	occ. ds C558	occ. ds LN	total
erreur	vD FN		vergissing begaan		N	97	108	85	290
faute	vD NF, vD FN		'misdrijven', overtreding, zondigen		N	84	8	77	169
crime	vD NF		'misdrijven'		N	44	44	27	115
acte	vD NF	pluriel ex. +répréhensibles	'zondigen'		N	11	44	12	67
infraction	vD NF		strafbaar feit, in vD NF overtreding		N	8	25	22	55
délit	vD NF vD NF, vD		'misdrijven'		N	8	23	13	44
péché	FN	ex. +article défini	'zonde bedrijven', een zonde begaan		N	34	5	3	42
attentat	vD FN		aanslag plegen		N	3	12	16	31
violation					A infraction	0	21	2	23
meurtre	vD FN		moord plegen/begaan		A crime	9	6	6	21
vol	vD NF	ex +avec infraction	'diefstal met braak plegen' 'een onvoorzichtigheid/onverstandigheid begaan'		A crime	8	0	11	19
imprudence	vD FN				N	15	2	1	18
injustice	vD NF	ex. +envers qqn	'iem onrecht aandoen'		N	7	10	0	17
violence	vD FN	pluriel ex. + à qqn	iemand geweld aandoen		N	2	2	13	17
abus				misbruik maken	N	4	10	2	16
agression	vD NF	ex. +acte de	agressie plegen		A acte A	1	0	15	16
exaction					violence	1	3	12	16
fraude	vD FN	ex. + à un examen	fraudebedrog 'fraude plegen bij een examen'		A crime A	0	14	2	16
atrocité				wreedheden plegen/begaan	violence	3	8	4	15
action	vD FN	ex. + mauvais et +vile (lage)	daad, idiomatisch		A acte	7	5	1	13
forfait	vD FN		misdaad, misdrijf begaan		A crime	7	2	2	11

faux pas				E	3	2	4	9
génocide				A crime	0	9	0	9
assassinat	vD NF	ex au pluriel	'moorden'	A crime	5	3	0	8
atteinte				A				
fait		ex. +de violence		infraction	0	2	6	8
impair	vD FN		'een flater slaan', 'een bok schieten'	A acte	1	0	7	8
méfait	vD NF		'misdrijven'	N	5	0	3	8
indiscrétion	vD FN		'uit de school klappen'	A crime	4	4	0	8
bêtise				E	7	0	0	7
larcin	vD FN		(kruimel)diefstal plegen	A impair	6	0	0	6
sottise/sotie	vD FN		een stommititeit uithalen, een dwaasheid begaan	A crime	4	0	2	6
viol				E	5	1	0	6
bévue	vD FN, vD NF		flater (in vD NF: misslag), blunderen	A crime	1	0	5	6
cambriolage				A impair	4	1	0	5
excès				A crime	1	0	4	5
folie				E	2	1	2	5
maladresse	vD FN		flater (begaan)	E	4	1	0	5
massacre				A impair	1	2	2	5
adultère	vD NF		'echtbreken', ontrouw plegen	A crime	1	2	2	5
connerie				A crime	2	2	0	4
dégât				E	4	0	0	4
geste				N	4	0	0	4
hold-up				E	2	0	2	4
irréparable				E	0	0	4	4
livre				E	0	0	4	4
mensonge				N	1	0	3	4
oubli	vD FN		iets verzuimen (te doen)	E	3	0	1	4
trahison	vD FN		verraad plegen	N	2	2	0	4
				A crime	3	1	0	4

abomination		pluriel		afgrijselijke daad	E	3	0	0	3
confusion	vD FN	ex. + grossière	verwisseling: ex.'zaken ernstig door elkaar halen'		E	2	1	0	3
contresens					E	2	1	0	3
gaffe	vD FN		stomiteit, 'een bok schieten', 'blunderen'		A impair	3	0	0	3
homicide					E	0	1	2	3
illégalité					A				
malversation					infraction	0	2	1	3
tuerie					E	1	2	0	3
cruauté					A crime	0	1	2	3
dégradations					E	2	0	0	2
écrits					E	0	1	1	2
espèglerie					A livre	1	0	1	2
étourderie	vD FN		'iets doms doen'		E	1	0	1	2
exhibition sexuelle					E	0	0	2	2
horreurs	vD FN	pluriel	'afschuwelijke dingen doen'		A				
pillage					violence	2	0	0	2
sacrilège					E	0	1	1	2
tentative		ex. +d'intimidation e.a.			E	2	0	0	2
turpitude	vD FN	[formeel of ironisch]	'zich schaamteloos gedragen'		E	0	0	2	2
bassesse	vD FN		gemene daad, laagheid		E	1	0	0	1
bavure					E	0	0	1	1
bogeys					E	0	0	1	1
bourde		syn. bévues +de débutant			E	0	0	1	1
braquage					E	0	0	1	1
cercle		ex. + vicieux			E	1	0	0	1
châtiment					E	1	0	0	1
circoncision					E	1	0	0	1

clic		ex. 'commettre le clic fatal'			E	0	0	1	1
contrefaçon		ex. +de marque			A	0	0	1	1
couac					E	0	0	1	1
débauche	vD NF	ex. +acte de	ontucht plegen'		E	0	0	1	1
déprédations				plunderingen	E	1	0	0	1
destruction					A dégât	0	0	1	1
diffamation				smaad	E	0	0	1	1
effraction					E	1	0	0	1
embardée					E	0	0	1	1
en-avant					E	0	0	1	1
énergie					E	1	0	0	1
enlèvement					E	1	0	0	1
fausse note	vD NF	ex. +grave	misstap begaan		E	1	0	0	1
galoup					E	1	0	0	1
graffiti					E	1	0	0	1
imposture					E	1	0	0	1
imprécision					E	0	0	1	1
inceste	vD FN		incest plegen		A crime	1	0	0	1
inconvenance	vD FN		'een onbeleefdheid ondergaan'		N	1	0	0	1
incorrection					N	0	0	1	1
indécence					E	1	0	0	1
indignité					E	1	0	0	1
infamies					E	1	0	0	1
irréversible					E	1	0	0	1
lâcheté	vD NF	comme adj	een lage daad begaan		E	1	0	0	1
ladrerie					E	1	0	0	1
lourdeur					E	1	0	0	1
mal		pl.			E	1	0	0	1
manquement					E	1	0	0	1

opuscule				E	0	0	1	1
ouvrage				A livre	0	0	1	1
perte de balle				E	0	0	1	1
poème				A livre	0	0	1	1
recueil				A livre	0	0	1	1
sabotage	vD NF	ex. +acte de	'sabotage plegen'	E	1	0	0	1
sacrifice				E	1	0	0	1
sournoiserie				E	1	0	0	1
stupidité				E	1	0	0	1
suicide				E	0	0	1	1
tube				E	0	0	1	1
vandalisme		ex +acte de		A crime	1	0	0	1
vilenie				E	1	0	0	1
					<b>477</b>	<b>395</b>	<b>414</b>	<b>1286</b>

### Annexe 3 une catégorisation par Vsupp

Npréd	begaan	plegen	aanrichten	intransitief	constr Vsupp-> V(supp) transitif	ander
dégats			*			aanrichten
dévastations			*			aanrichten
massacres	*	*				aanrichten
tube						aanrichten
tueries	*	*				aanrichten
turpitude						aanrichten
action	*	*				actie ondernemen, een goede daad verrichten
crime	*	*				bedrijven
attentat		*				bloedbaden aanrichten
comm un acte illégal						de wet overtreden, een illegale daad verrichten
atteintes		*				een aanslag zijn/doen op
bévue	*					een flater slaan
incorrection	*					een fout maken
geste						een gebaar maken
bassesse						een lage streek uithalen
faux pas	*					een mistap maken
sacrifice						een offer doen
tentative						een poging doen tot
imprudence	*					gevaarlijk bezig zijn, onvoorzichtig zijn
violence		*				geweld uitoefenen, geweld gebruiken tegen
goujaterie						ich ongemanierd gedragen
sottise/sotie	*					iets doms doen
infraction	*	*(delict)				iets misdoen/verkeerds doen, in overtreding zijn
illégalité						iets onrechtmatigs doen, tot outlaws maken
faute	*(vergissing)	*(fout)				iets verkeerds doen, maken, in een fout vervallen

indiscrétion					indiscreet zijn
couacs					maken
erreur	*(vergissing)				maken, (in de) fout lopen, de plank misslaan,
fraude(s)		*		frauderen	misbruik maken van,
indélicatesse					oneerlijk zijn
injustice	*	*			onrechtvaardig behandelen
violation	*	*			overtreden (wet)
infraction(s)					overtreden (wet), schenden (recht)
contresens					radicaal lang het doel heenschieten
forfait		*			uitvoeren
meurtre(s)	*	*		moorden	vermoorden
acte +DE			*	martelen (de torture)	verrichten
acte +DE +contre					verrichten
abomination					verrichten
acte +ADJ	*				verrichten, iets ADJ doen, ADJ verdrag vertonen
confusion					verwarring veroorzaken
abus +DE	*	*			zich bezondigen aan
faits		*			zich schuldig maken aan
délit(s)	*		*		zich schuldig maken aan
enlèvement					zich schuldig maken aan
graffiti					zich schuldig maken aan
lâcheté					zich schuldig maken aan
larcin					zich schuldig maken aan
pillages					zich schuldig maken aan
film					zich schuldig maken aan het schrijven van
livre					zich schuldig maken aan het schrijven van
opuscule					zich schuldig maken aan het schrijven van
ouvrages					zich schuldig maken aan het schrijven van





horreurs	*	*	
impair	*		
impolitesse	*		
imprécisions	*		
impudicité	*		
inceste		*	
inconvenance	*		
infamies		*	
injustice	*		
irrégularité(s)			gebeuren
irrémissible	*		
irréparable	*		
maladresse(s)	*		
malhonnêteté			
méfait	*	*	
mensonge	*		
oubli		*(verzuim plegen)	nalatig zijn, vergeten worden
péché	*		zondigen
péché		*	
plagiat	*		
sabotage	*		
sacrilège		*	
sournoiserie	*		
trahison		*	
viol	*		
vol	*		