



**UNIVERSITÉ  
DE GENÈVE**

**Archive ouverte UNIGE**

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

Master

2014

Open Access

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

---

Le passage de termes d'une langue de spécialité à la langue générale : Le cas du domaine spatial

---

Humbert-Droz, Julie

**How to cite**

HUMBERT-DROZ, Julie. Le passage de termes d'une langue de spécialité à la langue générale : Le cas du domaine spatial. Master, 2014.

This publication URL: <https://archive-ouverte.unige.ch/unige:41152>

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

Julie Humbert-Droz

# **Le passage de termes d'une langue de spécialité à la langue générale**

Le cas du domaine spatial

Mémoire présenté à la Faculté de Traduction et d'Interprétation pour l'obtention  
de la Maîtrise universitaire en traduction, mention Technologies de la traduction.

Directrice :

Mme Aurélie Picton

Jurée :

Mme Donatella Pulitano

Université de Genève

Année académique 2013-2014

Juillet 2014



# TABLE DES MATIÈRES

<b>1. INTRODUCTION GÉNÉRALE .....</b>	<b>4</b>
<b>2. ÉTAT DE L'ART.....</b>	<b>7</b>
2.1. Langue générale et langues de spécialité .....	7
2.2. Terminologie textuelle .....	9
<b>2.3. La déterminologisation .....</b>	<b>11</b>
2.3.1. Remarques préliminaires .....	11
2.3.2. Déterminologisation, terminologisation et migration .....	11
2.3.3. Points de vue existants .....	13
2.3.3.1. Guilbert et la banalisation lexicale .....	13
2.3.3.2. Galisson et la banalisation lexicale.....	13
2.3.3.3. Gouadec et la dé-spécialisation .....	15
2.3.3.4. Rastier et Valette et la dédomanialisation .....	15
2.3.3.5. Meyer et Mackintosh et la déterminologisation.....	15
2.3.4. Synthèse.....	17
<b>3. MÉTHODOLOGIE .....</b>	<b>19</b>
<b>3.1. Objectifs.....</b>	<b>19</b>
<b>3.2. Corpus des communiqués de presse de l'ESA – premier corpus de travail.....</b>	<b>19</b>
3.2.1. Pourquoi l'ESA.....	19
3.2.2. Pourquoi les communiqués de presse .....	20
3.2.3. Critères de constitution du corpus.....	21
3.2.3.1. Remarques préliminaires .....	21
3.2.3.2. Domaine .....	21
3.2.3.3. Langue.....	22
3.2.3.4. Taille – date.....	22
3.2.4. Caractéristiques du corpus .....	23
3.2.5. Extraction des termes .....	24
3.2.5.1. TermoStat.....	24

3.2.5.2.	Extraction des termes candidats nominaux .....	26
3.2.5.3.	Extraction des termes candidats verbaux .....	27
3.2.6.	Sélection des termes.....	27
3.2.6.1.	Validation des termes candidats.....	27
3.2.6.2.	Liste de termes définitive .....	30
<b>3.3.</b>	<b>Corpus de presse – deuxième et troisième corpus de travail .....</b>	<b>31</b>
3.3.1.	Choix des textes.....	31
3.3.2.	Critères de constitution des corpus .....	31
3.3.2.1.	Considérations générales.....	31
3.3.2.2.	Lingue.....	32
3.3.2.3.	Taille.....	32
3.3.3.	Caractéristiques des corpus.....	33
<b>4.</b>	<b>OBSERVATIONS.....</b>	<b>35</b>
<b>4.1.</b>	<b>Observations quantitatives globales .....</b>	<b>35</b>
4.1.1.	Représentation des termes dans les corpus .....	35
4.1.2.	Comparaison de la fréquence de certaines unités dans les deux corpus.....	38
<b>4.2.</b>	<b>Formes verbales .....</b>	<b>39</b>
4.2.1.	Unités dont la forme verbale est attestée dans les ressources spécialisées .....	42
4.2.1.1.	LAUNCH.....	42
4.2.1.2.	ECLIPSE.....	46
4.2.1.3.	ORBIT.....	49
4.2.2.	Unités dont la forme verbale n'est pas attestée dans les ressources spécialisées	53
4.2.2.1.	PROBE.....	53
4.2.2.2.	ROCKET .....	56
4.2.2.3.	STAR.....	59
4.2.3.	Remarques conclusives.....	61
<b>5.</b>	<b>CONCLUSION .....</b>	<b>65</b>
<b>6.</b>	<b>BIBLIOGRAPHIE .....</b>	<b>68</b>
<b>7.</b>	<b>ANNEXES .....</b>	<b>73</b>

7.1.	Annexe 3.1 : extrait du corpus des communiqués de presse de l'ESA	73
7.2.	Annexe 3.2 : extrait du corpus orienté termes .....	76
7.3.	Annexe 3.2 : extrait du corpus aléatoire .....	79
8.	INDEX DES TABLEAUX .....	84

# 1. INTRODUCTION GÉNÉRALE

La langue générale et les langues de spécialité n'ont pas de frontières nettes, il existe une certaine porosité entre elles, qui s'illustre par des mouvements entre elles. Il peut s'agir de mouvements allant d'une langue de spécialité à la langue générale, le contraire, c'est-à-dire de la langue générale pour pénétrer une langue de spécialité, ou encore de mouvements d'une langue de spécialité à une autre.

Dans ce travail, nous nous intéressons en particulier aux mouvements d'unités appartenant à une langue de spécialité vers la langue générale. Ce phénomène linguistique est appelé ici *déterminologisation*, même si nous verrons par la suite qu'il a adopté plusieurs dénominations et n'a pas toujours été défini de la même façon.

Nous nous intéressons plus particulièrement aux mouvements entre le domaine spatial et la langue générale. Le domaine spatial est un domaine riche et qui intéresse la grand public. La langue générale en est d'ailleurs marquée de plusieurs façons. Citons notamment *étoile*, dans l'expression *danseuse étoile*, ou *star*, en anglais, qui désigne aujourd'hui tout un éventail de personnalités un tant soit peu connues (acteurs, chanteurs, musiciens, etc.). L'intérêt pour le domaine spatial ne se limite pas aux locuteurs d'une seule langue, citons notamment l'expression *lune de miel* en français, qui se dit *honeymoon* en anglais et *luna di miele* en italien. Nous pourrions ainsi multiplier les exemples, mais il vaut mieux retenir que certaines unités appartenant au domaine spatial ont déjà pénétré la langue générale, y sont déjà largement diffusées et qu'il est très probable que d'autres unités suivent le même chemin.

L'objet de notre travail est donc l'étude du passage des termes appartenant au domaine spatial dans la langue générale et plus précisément les changements sémantiques et morphologiques qui s'ensuivent.

Notre étude se situe dans la continuité des travaux de Condamines et Picton (À paraître a), qui étudient la déterminologisation en français, dans des quotidiens français, à partir des communiqués de presse du CNES (Centre national d'études spatiales). Nous avons fait le choix de reprendre leur méthodologie et, ainsi, d'étendre le champ de recherche à l'anglais, l'anglais européen, plus précisément, sur la base des communiqués de presse de l'ESA.

À la suite des travaux de Condamines et Picton (À paraître a), nous émettons l'hypothèse que certains termes du domaine spatial « imprègnent » le grand public, et donc la langue générale, par

la presse généraliste et qu'ils y apparaissent par les communiqués de presse d'un organisme spatial, dans le cas de cette étude, l'ESA (l'Agence spatiale européenne). Notre hypothèse est que le phénomène de la déterminologisation se produirait sous la forme d'un flux, qui pénétrerait petit à petit la langue générale. C'est justement ce flux que nous nous proposons d'étudier.

Nous nous situons dans le champ de la terminologie textuelle, qui est, selon nous, l'approche la plus pertinente pour notre travail, car elle possède les outils nécessaires à la réalisation de notre étude. Nous exploitons ainsi les méthodes de linguistique de corpus, afin d'observer la déterminologisation directement dans les textes.

Notre méthodologie consiste donc, dans un premier temps, à constituer un corpus à partir des communiqués de presse de l'ESA. Ce corpus sert de base à une extraction de termes, laquelle nous permettra d'établir une liste de termes de référence du domaine spatial. Dans un deuxième temps, notre méthodologie consiste à constituer deux corpus de presse généraliste, le premier *orienté termes*, dont les articles doivent contenir au moins un terme figurant sur la liste précédemment établie, et le second *aléatoire*, dont les articles sont sélectionnés aléatoirement. Ces deux corpus servent ainsi de base à l'observation du phénomène de la déterminologisation. Le premier corpus contiendra nécessairement la plupart des termes qui nous intéressent, ce qui nous fournit une certaine garantie quant au phénomène à observer. Toutefois, ce corpus reste un corpus orienté et nous pensons qu'il est important d'observer le phénomène de la déterminologisation dans un corpus qui reflète réellement l'état de la langue, d'où le second corpus.

La deuxième section de ce travail concerne l'ancrage théorique dans lequel nous nous situons. Nous y définissons notamment les concepts de *langue générale* et de *langue de spécialité*, nous expliquons ce qu'est la terminologie textuelle et en quoi elle se différencie de la terminologie traditionnelle et, finalement, nous expliquons ce qu'est la déterminologisation, en procédant à un petit historique du phénomène et de ses dénominations.

La méthodologie que nous avons adoptée fait l'objet de la troisième section. Nous y expliquons en détail la constitution du premier corpus de travail, l'élaboration de la liste de termes de référence et la constitution des deuxième et troisième corpus de travail. Chacun de nos choix méthodologique est détaillé.

La quatrième section concerne l'exploration des deux corpus de presse généraliste. Nous y faisons quelques observations quantitatives générales avant de nous focaliser sur les formes verbales de six unités en particulier, par rapport à leurs formes nominales. Nous procédons à une

analyse systématique de plusieurs aspects de ces unités, qui nous permettra de cerner le phénomène de la déterminologisation sous la forme d'un flux.

La conclusion fait l'objet de la cinquième section, nous y présentons les tendances générales que nous avons essayé de dégager dans ce travail, ainsi que des perspectives en vue d'étendre le champ de recherche et d'améliorer nos observations.

## 2. ÉTAT DE L'ART

### 2.1. Langue générale et langues de spécialité

Puisque notre travail s'intéresse aux mouvements entre la langue générale et les langues de spécialité, il convient en premier lieu de définir ces deux concepts.

Le concept de *langue générale* n'est pas toujours défini de la même manière, selon les auteurs, et est souvent comparé au concept de *langue commune*. Rondeau (1981), par exemple, met en opposition la langue commune et les langues de spécialité. Pour lui, la langue commune est « l'ensemble des mots et expressions qui, dans le contexte où ils sont employés, ne se réfèrent pas à une activité spécialisée » (Rondeau, 1981 : 26). De plus, la langue commune et toutes les langues de spécialité forment ce que Rondeau appelle « l'ensemble lexical général d'une langue » (*Ibid.* : 27). La langue commune est donc un ensemble que partagent tous les locuteurs d'une langue donnée, tout en ne relevant pas d'un domaine spécialisé. La langue générale est cette langue donnée, c'est-à-dire toute la langue.

Kocourek (1991a), quant à lui, parle de la « langue tout entière » (Kocourek, 1991a : 13), c'est-à-dire de toute la langue, concept que nous rapprochons de « l'ensemble lexical général d'une langue » de Rondeau (1981 : 27). Il passe également en revue toutes les dénominations qu'a pu prendre la langue générale dans le sens de langue commune, selon les auteurs qui en ont parlé :

« Le concept de la langue usuelle (Galisson et Coste '76 :583) est très répandu bien que diversement délimité et nommé : langue courante (français courant de Bonnard '81 :5), langue quotidienne (Phal), commune (Rondeau), générale, ordinaire (Mounin) » (Kocourek, 1991a : 14).

Nous pouvons donc déjà constater que les visions de Rondeau et de Kocourek se rejoignent : il existe une partie de la langue qui est commune à tous les locuteurs et qui ne relève pas du discours spécialisé.

Cabré (1998 : 115), finalement, parle d'abord de « la langue dans le sens le plus global », qui comprend « la langue commune ou générale », puis elle fait une distinction entre la langue commune et la langue générale, que nous pouvons rapprocher de la conception de « langue tout entière » de Kocourek (1991a : 13) :

« La langue générale [...] peut être considérée comme un ensemble d'ensembles, imbriqués et reliés entre eux selon de nombreux points de vue. Le lien entre tous ces ensembles est la langue commune » (Cabré, 1998 : 115).

Le terme *langue générale* pourrait alors s'employer soit dans le sens de langue entière soit dans le sens de langue commune. Cette ambiguïté de dénominations est, le plus souvent, ignorée dans la pratique et le terme *langue générale* est généralement le terme retenu pour désigner l'« ensemble de règles, d'unités et de restrictions qui font partie des connaissances de la majorité des locuteurs d'une langue » (*Ibid.* : 115).

En ce qui concerne les langues de spécialité, il semble y avoir un certain consensus concernant leur définition. Ainsi, Rondeau se réfère à une communication liée à une activité spécialisée et parle alors de termes et non plus de mots ou d'expressions (Rondeau, 1981 : 26).

De plus, Rondeau insiste sur le fait que certains termes n'appartiennent pas strictement à un domaine et que les frontières entre langues de spécialité sont « perméables » (*Ibid.* : 26), surtout en ce qui concerne « des domaines connexes comme par exemple ceux des sciences de la santé, ou encore ceux de la chimie-physique » (*Ibid.* : 26). Les frontières entre la langue générale et les langues de spécialités sont également perméables, ce qui peut impliquer qu'une unité se retrouve à la fois dans une langue de spécialité et dans la langue générale (*Ibid.* : 26).

Pour Kocourek, la langue de spécialité est

« une sous-langue de la langue dite naturelle, enrichie d'éléments brachygraphiques, à savoir abrégatifs et idéographiques, qui s'intègrent à elle en se conformant à ses servitudes grammaticales » (Kocourek, 1991a : 12).

La langue de spécialité est donc une « variété de la langue entière » (*Ibid.* : 13). Kocourek la considère comme une sous-langue, car elle possède certaines spécificités plus importantes que juste le registre, le style ou le vocabulaire. Il parle d'ailleurs de langues de spécialité au pluriel, ce qui

« a l'avantage de mettre en évidence les différences qui existent, et d'indiquer une approche pluraliste permettant à chaque spécialité de développer plus librement ses propres démarches linguistiques » (Kocourek, 1991a : 17).

Pluriel que nous reprenons tout au long de ce travail, précisément pour ces raisons.

Cabré (1998 : 115) définit les langues de spécialité comme suit :

« nous parlons de langues de spécialité (ou langues spécialisées) lorsque nous faisons référence à l'ensemble des sous-codes – qui coïncident partiellement avec le sous-code de la langue commune – caractérisés par un certain nombre de traits particuliers ».

Cette idée de « traits particuliers » semble être ce qui fait se recouper ces trois points de vue. Une langue de spécialité est utilisée dans une situation de communication spécialisée, qui peut être conditionnée par « les interlocuteurs [...], les circonstances de la communication, ainsi que les buts et les intentions des intervenants » (*Ibid.* : 115).

Cabré explique également que les frontières entre langues de spécialité sont floues, car elles

« constituent un continuum à l'intérieur duquel on peut distinguer les différents domaines, chacun desquels constitue une langue de spécialité donnée, et l'ensemble, la langue de spécialité en général ou discours spécialisé » (Cabré, 1998 : 128).

Après avoir considéré ces différents points de vue, nous adoptons celui de Cabré selon lequel la langue générale et les langues de spécialités font partie de la langue dans un sens plus global (Cabré, 1998 : 115). Ainsi, dans ce travail, nous faisons le choix d'employer le terme *langue* quand il s'agit de la langue dans un sens global, le terme *langue générale* quand il s'agit de cette partie de la langue qui est commune à tous les locuteurs et le terme *langues de spécialité* lorsqu'il s'agit du discours spécialisé.

## 2.2. Terminologie textuelle

Il convient de rappeler l'ancrage théorique dans lequel nous situons notre travail. Dans cette section, nous expliquons donc ce qu'est la terminologie textuelle et en quoi elle diffère de la théorie traditionnelle de la terminologie.

La théorie traditionnelle de la terminologie nous vient principalement d'Eugen Wüster, considéré comme le fondateur de la discipline. Sa thèse, publiée en 1931, n'étant pas disponible en français, nous n'y avons pas un accès direct, mais seulement à travers d'autres auteurs, qui expliquent ou critiquent sa vision, tels que Lerat (1995), Cabré (1998) ou L'Homme (2004), pour n'en citer que quelques-uns. Le concept est au cœur de la réflexion terminologique ; on l'étudie selon une

approche onomasiologique et d'un point de vue synchronique. Le concept et le terme entretiennent une relation de bi-univocité, ce qui sous-entend qu'idéalement, la synonymie ne devrait pas exister. Rappelons que Wüster a travaillé pour une amélioration de la communication au niveau international (Van Campenhoudt, 2006 : 3) et que c'est de ce point de vue que doit se comprendre l'idée d'une relation de bi-univocité entre le concept et le terme (Cabré, 1998; L'Homme, 2004; Ungureanu, 2003; Van Campenhoudt, 2006).

Depuis les années 1980 environ, un certain nombre de critiques ont été émises quant à la théorie de Wüster. En fait, la vision de la terminologie a évolué et aujourd'hui les pratiques terminologiques sont principalement basées sur des corpus (L'Homme, 2003 : 404). L'exploitation de corpus permet en effet « de traiter de grandes quantités de données textuelles sur support électronique » (Condamines, 2005 : 37). Ainsi, grâce aux moyens techniques d'aujourd'hui, nous pouvons nous appuyer sur des volumes de textes considérables pour décrire ou attester un usage, perspective qui s'éloigne de la terminologie traditionnelle. Le corpus devient alors « l'objet d'étude » (Condamines, 2005 : 41) des terminologues, les textes étant représentatifs d'un aspect particulier de la langue, comme la langue de spécialité d'un domaine ou un type de texte particulier (Bowker & Pearson, 2002 : 12).

De cette manière, la terminologie n'est plus uniquement normative, elle est aussi descriptive (Béjoint & Thoiron, 2000 : 15). De plus, l'exploitation de corpus textuels informatisés permet de porter un nouveau regard sur la dimension diachronique de la terminologie, qui peut se traduire, par exemple, par l'observation de l'évolution des concepts et des dénominations, sur une période plus ou moins longue (Dury, 2006). Les corpus permettent aussi de montrer que la synonymie est une réalité en terminologie et dans les langues de spécialité (Humbley, 2004), et ce, pas nécessairement dans des domaines différents.

Ces observations, faites à partir d'usages réels attestés dans des textes, ont amené les terminologues à reconsidérer la théorie de la terminologie de Wüster, à remettre en question cette « optique conceptuelle »<sup>1</sup>, ce qui a conduit à un renouveau théorique et méthodologique. De nouvelles propositions théoriques ont émergé, comme la théorie communicative de la terminologie de Cabré (1998) ou l'optique lexico-sémantique de L'Homme (2004). La terminologie textuelle (Condamines, Bourigault, Slodzian) est la théorie qui nous intéresse plus particulièrement dans ce travail. En effet, le cadre théorique et méthodologique qu'elle propose correspond aux besoins de notre étude, à savoir, principalement, l'exploitation de corpus textuels pour observer un phénomène linguistique en particulier.

---

<sup>1</sup> Terme utilisé par L'Homme pour désigner la théorie de la terminologie de Wüster.

La terminologie textuelle est une proposition théorique établie par le groupe « Terminologie et Intelligence Artificielle ». Bourigault et Slodzian (1999 : 31) considèrent qu'il faut s'éloigner de la terminologie traditionnelle pour pouvoir prendre en considération le texte : « la terminologie doit sortir d'une sémiotique du signe fondée sur la triade terme/concept/référent qui la rend inapte à aborder le texte ». Ils proposent alors une approche descriptive de la terminologie, qui « ouvre largement les portes à tous les acquis de l'analyse linguistique et textuelle (on dépasse ainsi la vision étroite des LSP) » (*Ibid.* : 32).

Dans cette section, nous avons présenté le cadre théorique dans lequel nous situons notre travail, la terminologie textuelle. Dans la section suivante, nous expliquons en détail le phénomène linguistique qui est au cœur de notre travail, la déterminologisation. Nous expliquons d'abord ce qu'est ce phénomène, nous passons ensuite en revue les différents points de vue existants sur le sujet, nous faisons ensuite la synthèse de ces informations et présentons finalement la définition de la déterminologisation que nous retenons pour ce travail.

## **2.3. La déterminologisation**

### **2.3.1. Remarques préliminaires**

La déterminologisation est un phénomène linguistique peu étudié, mais ce n'est en rien un phénomène nouveau (Guilbert, 1975 : 84; Meyer & Mackintosh, 2000 : 200). Toutefois, si nous nous y intéressons de plus en plus, c'est justement parce qu'il se produit de plus en plus, du fait, notamment, de la « société du savoir » dans laquelle nous vivons (Meyer & Mackintosh, 2000 : 201). La déterminologisation se définit essentiellement par le passage de termes appartenant à une langue de spécialité dans la langue générale. S'ensuivent alors de nombreuses questions sur les mutations, tant sémantiques que morphologiques, que subissent ces termes. Ainsi, la déterminologisation désigne non seulement le phénomène linguistique, mais aussi son résultat. Meyer et Mackintosh (2000 : 203) qualifient les termes passés dans la langue générale, et donc plus uniquement employés par des spécialistes du domaine, de « mots déterminologisés ». Nous reprendrons ce terme tout au long de notre travail.

### **2.3.2. Déterminologisation, terminologisation et migration**

Le phénomène de la déterminologisation, comme la terminologisation et la migration, est un phénomène d'échange, ou d'emprunt, inhérent à la proximité des langues de spécialité et de la langue générale et aux interactions entre elles. En effet, « les langues de spécialité ne sont pas

étanches, une importante circulation de termes existe entre elles et avec la langue commune » (Celotti et Musacchio 2004 : 268).

Ainsi, si la déterminologisation est le passage de termes vers la langue générale, la terminologisation (Cabré, 1998), ou la spécialisation (Calberg-Challot, 2007; Gouadec, 1990), est le contraire, c'est-à-dire l'emprunt à la langue générale pour dénommer un concept d'une langue de spécialité. Calberg-Challot (2007) explique que le sens des unités se spécialise pour répondre à une demande d'une langue de spécialité. L'informatique fournit de nombreux exemples de terminologisation, notamment *fenêtre* et *bureau* ou, pour reprendre les exemples de Meyer et Mackintosh (2000 : 214), *souris* et *autoroute de l'information*.

Certaines unités peuvent être plus susceptibles de passer dans la langue générale que d'autres (Meyer & Mackintosh, 2000 : 214)<sup>1</sup>. Si nous reprenons les exemples que nous venons de citer, nous pouvons constater qu'ils ont d'abord été terminologisés, c'est-à-dire qu'ils sont passés dans une langue de spécialité, celle de l'informatique en l'occurrence. Leur sens spécialisé est ensuite passé dans la langue générale et est maintenant largement diffusé. Nous pouvons expliquer ces mouvements non seulement par la place de plus en plus importante que prend l'informatique dans notre vie quotidienne, mais aussi par la simplicité morphologique des termes et peut-être aussi par le fait qu'ils étaient déjà bien présents dans la langue générale avant d'adopter un sens spécialisé.

Le troisième phénomène dont il est question dans cette section est la migration. Même si ce phénomène peut aussi recouvrir les réalités de la terminologisation et/ou de la déterminologisation (Celotti & Musacchio, 2004; Dury, 2008; Guilbert, 1975; Toma, 2003, 2011), nous l'entendons ici comme le passage d'un terme d'une langue de spécialité à une autre.

Comme dans le cas de la terminologisation, la langue peut faire recours à la métaphore ou à l'analogie, comme dans le cas du terme *virus*, qui est passé du domaine médical au domaine informatique (Humbley, 2006). L'analogie entre le comportement d'un virus en médecine et celui d'un virus en informatique a servi à développer tout un vocabulaire décrivant ce comportement : « l'action du virus qui s'attaque à sa victime, qui se reproduit, qui se répand » (Humbley, 2006).

---

<sup>1</sup> « L'idéal serait que les unités lexicales susceptibles de se déterminologiser soient faciles à comprendre et à employer. Une telle approche s'observe déjà dans le domaine de l'informatique où les métaphores (par exemple, la *souris*, l'*autoroute de l'information*) sont de plus en plus populaires. » (Meyer & Mackintosh, 2000 : 214)

### **2.3.3. Points de vue existants**

La déterminologisation étant un phénomène peu étudié systématiquement, il n'existe pas de consensus quant à sa dénomination. Ainsi, Guilbert (1975) et Galisson (1978) parlent de *banalisation lexicale*, Gouadec (1990), de *dé-spécialisation* et Rastier et Valette (2009), de *dédomanialisation* ; Meyer et Mackintosh (2000), quant à elles, parlent finalement de *déterminologisation*. Chaque dénomination possède des caractéristiques propres à son ou ses auteurs. Nous proposons de les examiner une par une dans les sections suivantes puis de faire un point sur la définition de la déterminologisation que nous avons prise comme base pour ce travail.

#### **2.3.3.1. Guilbert et la banalisation lexicale**

Guilbert (1975 : 84) parle de banalisation lexicale et de métaphorisation des termes techniques :

« Le transfert par métaphore de la réalité au plan de la désignation linguistique s'opère désormais à partir des concepts scientifiques les plus vulgarisés ou des pratiques techniques les plus courantes ».

Pour lui, ce phénomène est le passage de termes techniques ou scientifiques dans la langue générale ; il s'explique par le fait que nous sommes tous en relation étroite avec un domaine et qu'il est ainsi naturel que nous commençons à utiliser des termes y appartenant. Nous avons toujours été plus ou moins proches de domaines plus ou moins techniques, sauf que c'est de plus en plus le cas pour de plus en plus de domaines différents et pour de plus en plus de personnes. Suivant le raisonnement de Guilbert, il serait donc naturel d'emprunter occasionnellement aux langues de spécialité de ces domaines-là.

Guilbert apparente la banalisation lexicale à la vulgarisation, car, pour lui, ce sont justement des domaines techniques et scientifiques déjà vulgarisés qui pénètrent notre vie quotidienne et auxquels nous empruntons certains termes.

#### **2.3.3.2. Galisson et la banalisation lexicale**

Galisson (1978) ne parle pas tant du phénomène linguistique que de son résultat ; il met ainsi l'accent sur le vocabulaire banalisé d'un domaine, qui est, selon lui, le résultat de la banalisation lexicale de ce même domaine. Pour lui, la banalisation lexicale a pour but de diffuser plus largement un domaine, dans le cas de son étude, le football. Son point de vue est plus centré sur la communication et la simplification de cette communication.

Pour lui, le vocabulaire banalisé d'un domaine sert justement à ce que des non-spécialistes puissent comprendre le domaine en question. Il s'agit bien là de simplifier la communication entre experts et non-experts.

Cependant, Galisson préfère parler de langage banalisé et non de langage vulgarisé, parce que, selon lui, cela ne recouvre pas exactement la même réalité. Il définit la banalisation lexicale comme l'« adaptation du vocabulaire d'une langue de spécialité au niveau de connaissance des non-spécialistes » (Galisson, 1978 : 276). Le langage banalisé est un « langage décodable par le plus grand nombre, sinon par tous » (*Ibid.* : 275). On dénote ici une intention de la part des utilisateurs de ce langage banalisé à communiquer avec le plus grand nombre. Nous pensons que c'est là déjà une légère différence de la vision de Guilbert. Galisson insiste plusieurs fois sur cette intention, cette volonté de communiquer. Même si le phénomène de banalisation n'est pas tout à fait conscient, il est imprégné d'une volonté qui n'apparaît pas chez Guilbert.

Selon Galisson, la banalisation lexicale et la vulgarisation sont deux phénomènes qui se recoupent, mais qui ne désignent pas tout à fait la même chose, le premier « fonctionne sur les bases d'un large consensus (souvent grâce à l'utilisation massive et à l'action uniformisante des media) » (Galisson, 1979 : 75), tandis que le second est instable et occasionnel, et « utilise circonstanciellement des moyens qui peuvent varier d'un locuteur à l'autre » (*Ibid.* : 75).

Galisson part du principe qu'il y a différents niveaux de langues de spécialité, plus ou moins spécialisés ; il parle de « stratifications » à l'intérieur des langues de spécialité. Les différents degrés d'expertise de Bowker et Pearson (2002)<sup>1</sup>, à savoir expert, semi-expert et non-expert, et des différents types de communication qui en découlent (expert à expert, expert à semi-expert et expert à non-expert)<sup>2</sup> se rapprochent beaucoup de l'idée de Galisson. En effet, le niveau de spécialisation d'une langue de spécialité peut varier selon le contexte. Pour Galisson, la banalisation lexicale est l'une de ces stratifications. Nous pensons que cette idée de degrés d'expertise et de stratifications à l'intérieur d'une langue de spécialité, et donc de types de communication différents, est au cœur du phénomène de déterminologisation : c'est parce qu'il y a interaction entre les experts, les semi-experts et les non-experts, que certains termes spécialisés passent dans la langue générale.

---

<sup>1</sup> « Because LSP users have different levels of expertise, there are also different levels of LSP communication » (Bowker & Pearson, 2002 : 28).

<sup>2</sup> « When experts communicate (e.g. through publications in research journals), they tend to use a highly specialized language. [...] Another type of communication takes place between experts and semi-experts, such as students or experts from related fields. [...] A third type of communication occurs between experts and non-experts. [...] The expert does not expect the non-expert to achieve the same level of understanding of the terms uses as long as the general idea is understood. » (Bowker & Pearson, 2002 : 28).

### **2.3.3.3. Gouadec et la dé-spécialisation**

Gouadec (1990), quant à lui, parle de dé-spécialisation et de vulgarisation, deux appellations qui décrivent le même phénomène. Il entend par là le passage de termes dans un groupe de non-spécialistes avec, pour conséquence, des changements de sens plus ou moins importants, que ce soient des nouvelles acceptions ou des extensions de sens. Comme les non-spécialistes s'approprient certains termes et en font un usage « profane », lesdits termes se retrouvent dépourvus de certains des sèmes qui définissaient leur sens d'origine. Ainsi, le concept qu'ils décrivent est moins précis, moins détaillé, qu'en langue de spécialité, d'où le terme de « dé-spécialisation ».

### **2.3.3.4. Rastier et Valette et la dédomanialisation**

Rastier et Valette (2009) se sont intéressés à la polysémie et à la néosémie. Dans leur étude, ils ont décrit un phénomène qui se rapproche fortement de la déterminologisation et des définitions qu'en donnent les auteurs que nous avons déjà cités. Il s'agit de la dédomanialisation, qui « autonomise un sémème par rapport à son domaine d'origine » (Rastier & Valette, 2009 : 7). Autrement dit, une unité perdrait, en quelque sorte, son appartenance à un domaine et ferait ainsi partie de la langue générale.

Ils définissent le domaine comme étant « un groupe de classes sémantiques (ou *taxèmes*) lié à une pratique sociale » (*Ibid.* : 7), « les classes sémantiques [regroupant] des unités lexicales (lexies) en fonction de sèmes génériques (i.e. communs à toutes les unités lexicales de la classe) et se [structurant] en fonction de sèmes spécifiques » (Valette et al. 2006 : 4).

### **2.3.3.5. Meyer et Mackintosh et la déterminologisation**

Meyer et Mackintosh (2000) parlent de la déterminologisation comme de l'« étirement du sens » d'un terme (Meyer & Mackintosh, 2000 : 198). Autrement dit, un terme peut voir son sens s'étirer, s'étendre, pour recouvrir une réalité plus vaste qu'à l'origine. Sur ce point, elles rejoignent Gouadec en ce que le terme dont le sens est étiré aura un sens plus général que le terme dans son contexte d'origine, dans son domaine d'origine. Les auteures nuancent toutefois leur propos en distinguant deux types de changement sémantique :

« Soit l'essence du sens terminologique demeure dans le mot déterminologisé, soit on assiste à des changements conceptuels plus considérables, l'extension sémantique du mot s'éloignant alors nettement de celle du terme d'origine » (Meyer & Mackintosh, 2000 : 202).

Pour Guilbert, l'utilisation en langue générale de termes appartenant à des domaines relativement proches de nous s'explique en partie par le fait que nous sommes de plus en plus proches des sciences et des techniques, nous avons la possibilité d'en savoir toujours plus, et notre quotidien en est désormais imprégné. Meyer et Mackintosh énoncent la même idée : comme nous nous tenons de plus en plus informés sur les sciences et les techniques, il est naturel que nous finissions par utiliser la terminologie propre à ces domaines. Toutefois, nous l'adaptions, en quelque sorte, à nos besoins et à nos connaissances, d'où ce phénomène de déterminologisation. Cette adaptation, que nous faisons inconsciemment, s'explique par le fait que, comme nous ne sommes pas des spécialistes des domaines en question, nous n'avons pas besoin du sens exact des termes. Ce sens peut donc s'étendre ou se restreindre, mais, le plus souvent, il s'étend et recouvre une réalité plus large. Pearson (1998), citée par Meyer et Mackintosh (2000 : 201-202) l'explique ainsi :

« ce qui différencie ce dialogisme particulier [expert/non-spécialiste] des deux précédents [expert/expert et expert/étudiant] est qu'il ne requiert pas le même niveau de compréhension des termes employés de la part de l'auteur et du lecteur, pourvu que la portée générale du message soit comprise... Ce dialogisme ne favorise donc pas un emploi rigoureux des termes »<sup>1</sup>.

L'idée qu'il existe des niveaux différents de compréhension est cruciale. C'est la raison pour laquelle il y a un flux de termes qui passent dans la langue générale. En langue générale, nous parlons des mêmes concepts qu'en langue de spécialité, mais nous en avons une connaissance moins approfondie, « l'essence du sens terminologique demeure » alors, pour reprendre le premier type de changement sémantique énoncé par Meyer et Mackintosh (2000 : 202).

Certains mots déterminologisés s'éloignent plus de leur sens et de leur domaine d'origine que d'autres, ce qui est à rapprocher du second type de changement sémantique de Meyer et Mackintosh (2000 : 202). L'explication que donnent les auteures est encore une fois très intéressante : lorsqu'un terme ne subit qu'une extension de sens, il reste le plus souvent très proche de son sens d'origine. Citons notamment *VIH* ou *IRM*, car, bien qu'ils soient très largement utilisés dans la langue générale, la plupart des gens n'ont pas du tout la même

---

<sup>1</sup> « *What distinguishes this particular communicative setting [expert-to-layperson] from the two previous ones [expert-to-expert and expert-to student] is that there is no need for author and reader to achieve the same level of understanding of the terms used as long as the broad thrust of the message is understood... this particular communicative setting is not conducive to terms being used in a rigorous manner or being perceived as such.* » (Pearson, 1998).

connaissance du concept qu'un médecin ou qu'un technicien en radiologie. Le sens fondamental n'a pourtant pas changé, seul le degré de compréhension du concept est différent.

Lorsqu'un terme subit plusieurs extensions de sens, le sens d'origine peut parfois être extrêmement éloigné. Reprenons un exemple de Meyer et Mackintosh :

« D'autre part, *anorexic* peut s'employer au sens plus familier de "extrêmement mince" et, si l'on pousse le concept plus loin, de "sans fond" comme dans *anorexic plot* ou encore de "dévalué", comme dans *anorexic Canadian dollar* » (Meyer & Mackintosh, 2000 : 204).

Il est évident que le sens d'*anorexic* dans *anorexic Canadian dollar* n'est plus du tout le même que son sens d'origine, en médecine.

Les auteurs appellent ces extensions de sens la « *dilution* » du sens des termes (*Ibid.* : 204).

#### **2.3.4. Synthèse**

Nous souhaitons souligner que, puisque l'étude de ce phénomène a généré un certain nombre de dénominations différentes, nous faisons le choix d'utiliser la terminologie de Meyer et Mackintosh (2000) dans notre travail, à savoir l'emploi de *déterminologisation* pour désigner le phénomène linguistique et l'emploi de l'adjectif *déterminologisé* pour désigner une unité largement diffusée dans la langue générale.

Nous avons établi un premier historique de la déterminologisation et il semble que ce phénomène soit encore mal connu et peine à trouver une explication qui satisfasse tous les auteurs. Néanmoins, l'idée fondamentale est la même, il s'agit du passage de termes d'une langue de spécialité à la langue générale, affaiblissant ainsi le degré de compréhension du concept d'origine, ce qui peut entraîner un certain nombre de changements sémantiques plus ou moins importants. En effet, le sens peut parfois s'éloigner considérablement du concept d'origine.

Les auteurs semblent s'accorder sur le fait que la déterminologisation est un phénomène qui se produit sous la forme d'un flux, qui imprègne petit à petit la langue générale. C'est précisément sur ce point que nous nous appuyons dans le cadre de cette étude.

Un deuxième point est d'une importance capitale pour nos recherches, et rejoint l'idée de flux que nous venons d'énoncer : la déterminologisation, au contraire de la vulgarisation, qui est un processus langagier tout à fait conscient<sup>1</sup>, semble être un phénomène inconscient.

Nous nous représentons donc la déterminologisation comme un flux de termes qui pénètrent la langue générale de manière relativement inconsciente. Il n'y aurait pour ainsi dire pas la même volonté de transmettre le savoir que dans le cas de la vulgarisation, qui reste un processus dénotant une certaine volonté de transmettre le savoir. Le Petit Robert en ligne définit d'ailleurs la vulgarisation scientifique comme « le fait d'adapter un ensemble de connaissances techniques, scientifiques, de manière à les rendre accessibles à un lecteur non spécialiste »<sup>2</sup>.

Nous souhaitons donc concentrer nos recherches sur la déterminologisation comme processus inconscient, à travers une étude réalisée en linguistique de corpus d'un point de vue diachronique. Nous espérons ainsi pouvoir étudier le flux de termes arrivant dans la langue générale.

---

<sup>1</sup> Nous nous référons ici à la définition de « vulgarisation scientifique » du Petit Robert, donnée au paragraphe suivant.

<sup>2</sup> Le Petit Robert 2014, <http://pr.bvdep.com/>, consulté le 5 mars 2014.

## **3. MÉTHODOLOGIE**

### **3.1. Objectifs**

L'objectif de ce travail est d'observer le phénomène de la déterminologisation, qui semble se produire sous la forme d'un flux, dans le domaine spatial en vue de pouvoir en dégager des tendances générales. Avant de nous focaliser sur certaines unités en particulier, nous ferons quelques observations quantitatives générales. Nous nous concentrerons ensuite sur la manière dont les termes appartenant au domaine spatial imprègnent le grand public. Pour ce faire, nous observerons les différences de comportement de six unités dans les corpus de presse (voir point 3.3.) par rapport au corpus des communiqués de presse (voir point 3.2.). Nous verrons ensuite à quels changements sémantiques ces différences de comportement pourraient mener. En observant les différences d'emploi de ces unités dans la presse par rapport aux communiqués de presse, nous tenterons de voir le chemin que peut prendre une unité lorsqu'elle passe du discours spécialisé au discours général.

### **3.2. Corpus des communiqués de presse de l'ESA – premier corpus de travail**

#### **3.2.1. Pourquoi l'ESA**

L'ESA est l'agence spatiale européenne (*European Space Agency* en anglais), une organisation internationale, créée en 1975 et comprenant vingt États membres, à savoir l'Autriche, la Belgique, la République Tchèque, le Danemark, la Finlande, la France, l'Allemagne, la Grèce, l'Irlande, l'Italie, le Luxembourg, les Pays Bas, la Norvège, la Pologne, le Portugal, la Roumanie, l'Espagne, la Suède, la Suisse et le Royaume-Uni. Même s'ils ne sont pas membres à part entière, le Canada, la Hongrie, l'Estonie, la Lettonie et la Slovénie ont signé un accord de coopération avec l'ESA. L'ESA est donc véritablement une organisation européenne, dans le sens géographique du terme, puisque la Suisse et la Norvège ne font pas partie de l'Union européenne. L'ESA a pour but d'encourager la coopération de ces États en ce qui concerne la recherche spatiale. Ainsi, elle coordonne de nombreux programmes et missions<sup>1</sup>.

Dans ce travail, nous souhaitons étudier le phénomène de la déterminologisation dans la langue anglaise, dans une perspective européenne. Notre choix d'exploiter les communiqués de presse

---

<sup>1</sup> Agence spatiale européenne (ESA), <http://www.esa.int>, consulté le 5 mars 2014.

de l'ESA nous permet de conserver cette dimension européenne dans la variété de l'anglais que nous étudions.

### **3.2.2. Pourquoi les communiqués de presse**

Tout d'abord, puisque nous reprenons la méthodologie employées par Condamines et Picton (À paraître a), nous faisons le choix d'exploiter les communiqués de presse. Le but de cette étude explique partiellement ce choix. Comme il s'agit d'observer l'emploi de termes appartenant au domaine du spatial dans la langue générale, en l'occurrence dans la presse généraliste, nous devons en premier lieu décider des termes à observer. Les communiqués de presse de l'ESA nous ont paru être une source évidente, puisqu'ils sont destinés aux journalistes. Les termes y apparaissant sont donc susceptibles d'être utilisés dans des articles de presse.

Bien que nous ayons déjà formulé notre hypothèse dans l'introduction, nous souhaitons l'explicitier avant de poursuivre : les journalistes se réfèrent aux communiqués de presse de l'ESA, ils peuvent donc y reprendre des termes appartenant au domaine du spatial et les diffuser au moyen de leurs articles ; ces articles deviennent le média privilégié de la diffusion de ces termes dans le grand public, c'est-à-dire dans la langue générale. Une autre raison découle de notre hypothèse et prend un peu la question à l'envers : puisque deux corpus de presse généraliste font l'objet de cette recherche, il fallait s'assurer notamment de la présence des termes issus du corpus de base dans ces corpus de presse. La question d'un corpus de base construit avec des textes relevant d'une communication expert à expert s'est posée, mais n'a pas été retenue, de peur d'extraire des termes n'apparaissant pas ou trop peu dans la presse généraliste. Encore une fois, le genre du communiqué de presse semblait le plus adapté car c'est « un genre de textes "à l'interface" entre spécialistes et grand public » (Condamines & Picton, À paraître a), c'est, en quelque sorte, une passerelle entre le savoir des spécialistes et les journalistes, en partie responsables de la diffusion de ce savoir. En le rendant accessible, ils participent à l'évolution de notre « société du savoir » (Meyer & Mackintosh, 2000 : 201).

Puisque les communiqués de presse émanent de l'ESA, une organisation officielle, nous considérons leur source comme fiable. Ils sont donc tout à fait appropriés pour la constitution d'un corpus.

### 3.2.3. Critères de constitution du corpus

#### 3.2.3.1. Remarques préliminaires

Bien que ce corpus constitué des communiqués de presse de l'ESA apparaisse comme un choix logique, il répond tout de même à certains critères spécifiques. De plus, même s'il ne sert que de base à une extraction terminologique, il n'en est pas moins construit en respect de certaines contraintes.

S'il existe de nombreuses références quant à la constitution de corpus de langues de spécialité, nous nous sommes principalement appuyée sur les ouvrages de Bowker et Pearson (2002) et de L'Homme (2004). Plusieurs auteurs, notamment Meyer et Mackintosh (1996), Ahmad et Rogers (2001) et Pearson (1998), tout en étant moins spécifiques que les deux ouvrages cités précédemment, s'accordent également sur l'importance à donner au domaine, à la taille et à la représentativité des textes à sélectionner.

#### 3.2.3.2. Domaine

Le domaine spatial est un domaine hétérogène, il contient en effet de nombreux sous-domaines, comme l'astronomie, l'astrophysique, l'astronautique ou les corps célestes. Nous pensons que les communiqués de presse de l'ESA représentent cette hétérogénéité, d'une part, parce que l'ESA traite de tous les sous-domaines du spatial et, d'autre part, parce que l'ESA fait part au média, via les communiqués, de nombreux événements différents. Nous pouvons nous en rendre compte à la simple lecture du titre de certains communiqués : « *Europe's largest spaceship reaches its orbital port* »<sup>1</sup>, « *ESA pays tribute to Mr Erik Quistgaard* »<sup>2</sup> ou « *ESA Director General to meet the press in Paris* »<sup>3</sup>.

Nous pensons qu'il est d'une importance capitale de ne pas se focaliser sur un seul sous-domaine, mais sur le domaine tout entier, aussi hétérogène qu'il soit. Nous nous assurons ainsi d'obtenir un corpus le plus représentatif possible. Les communiqués de presse ayant un contenu assez varié pour recouvrir tous les sous-domaines du domaine spatial, nous considérons qu'ils répondent à cette contrainte, le point 3.2.6.2. confirme d'ailleurs notre propos.

---

<sup>1</sup> Communiqué de presse du 15 juin 2013 (n° 23), Agence spatiale européenne (ESA), <http://www.esa.int>, consulté le 5 mars 2014.

<sup>2</sup> Communiqué de presse du 5 mars 2013 (n° 2), idem.

<sup>3</sup> Communiqué de presse du 5 janvier 2009 (n° 2), idem.

### **3.2.3.3.Langue**

La langue de ce corpus étant l'anglais, la question de savoir si les communiqués de presse le constituant avaient été écrits par des locuteurs natifs ou non s'est évidemment posée. Bien que Bowker et Pearson (2002 : 52) conseillent de n'inclure que des textes originaux et rédigés par des locuteurs natifs, nous pensons que, puisque les communiqués de presse sont écrits par des professionnels de la communication, leur langue est d'une qualité qui se prête tout à fait à la constitution de ce corpus.

De plus, il nous est impossible de savoir si les communiqués de presse de l'ESA en anglais sont des traductions ou non, même s'il est très probable, sans être une certitude pour autant, qu'ils soient d'abord rédigés en anglais, puis traduits en français et en allemand, les deux autres langues des communiqués. En effet, certains communiqués n'existent qu'en anglais, ils n'ont donc pas été traduits en français et en allemand. Néanmoins, nous n'excluons pas la possibilité, même si elle est peu probable, que certains aient été rédigés d'abord en français ou en allemand puis traduits en anglais. Nous considérons toutefois que, le cas échéant, ils sont de bonne qualité et conviennent tout à fait pour la constitution de ce corpus.

### **3.2.3.4.Taille – date**

Le critère de la taille est un critère important lors de la constitution d'un corpus, car il contribue à en garantir la représentativité. Néanmoins, il convient de privilégier la qualité des textes par rapport à leur quantité, comme l'affirment Bowker et Pearson (2002 : 45-46) : « *You may find that you can get more useful information from a corpus that is small but well designed than from one that is larger but is not customized to meet your needs* ».

Il convient de remarquer que les différents auteurs consultés n'ont pas toujours le même point de vue quant à la taille d'un corpus de langue de spécialité. D'où l'importance de nuancer ce critère et, surtout, de prendre en considération le but de l'étude basée sur ce corpus. L'Homme (2004 : 128) semble vouloir trancher en affirmant qu'il « n'existe pas de véritable consensus en ce qui concerne la taille idéale d'un corpus spécialisé ».

Ainsi, nous souhaitons traiter certains des facteurs dont dépend la taille d'un corpus spécialisé, selon Bowker et Pearson (2002 : 45) à savoir les besoins de l'étude et la disponibilité des ressources.

Dans notre cas, le corpus constitué des communiqués de presse de l'ESA sert de base à l'extraction des termes appartenant au domaine spatial, il doit donc être représentatif de ce domaine, comme nous l'avons mentionné au point 3.2.3.2. Nous privilégions ainsi un corpus dense en terminologie, où de nombreux termes appartenant au domaine spatial sont présents.

La question de la disponibilité des ressources a très vite été résolue car, les communiqués de presse de l'ESA étant publiés sur le site Internet de l'ESA, ils sont accessibles au grand public, il nous a donc été très facile de les récupérer pour construire notre corpus. Cette question ayant été déjà abordée au point 3.2.2., nous n'allons pas y revenir plus en détail.

Il convient de rappeler que notre étude du phénomène de la déterminologisation de termes du domaine spatial se situe sur une période de seize ans, de 1998 à 2013<sup>1</sup>. Il semble alors pertinent de considérer la même période pour l'établissement d'une liste de termes à examiner. La question de la taille devient alors une question de dates. Nous avons ainsi fait le choix de sélectionner tous les communiqués de presse de l'ESA du 1<sup>er</sup> janvier 1998 au 31 décembre 2013.

#### **3.2.4. Caractéristiques du corpus**

Après avoir répondu à toutes ces questions, le corpus constitué des communiqués de presse de l'ESA du 1<sup>er</sup> janvier 1998 au 31 décembre 2013 contient seize sous-corpus, comprenant chacun tous les communiqués de presse d'une année, pour un total de 540 974 mots. Il est constitué de 874 communiqués de presse, ce qui représente une moyenne d'environ 54 communiqués par année. Un extrait de ce corpus figure en annexe (annexe 3.1).

Le tableau ci-dessous répertorie le nombre de mots de chaque sous-corpus.

---

<sup>1</sup> Puisque nous nous basons sur les travaux de Condamines et Picton (À paraître a), nous avons décidé de considérer la même période.

<b>Année</b>	<b>Occurrences</b>
1998	22 125
1999	28 079
2000	43 914
2001	46 604
2002	46 029
2003	50 897
2004	38 888
2005	47 765
2006	36 150
2007	22 657
2008	33 932
2009	21 157
2010	16 799
2011	18 208
2012	31 385
2013	36 385
<b>Total</b>	<b>540 974</b>

Tableau 3.1 : Nombre d'occurrences dans les sous-corpus du corpus des communiqués de presse de l'ESA

Le corpus a été nettoyé pour ne contenir que le titre, la date et le texte de chaque communiqué de presse. Des adresses apparaissent toutefois dans le corpus, nous avons fait le choix de ne pas les enlever, notamment pour des raisons de temps. Il nous a semblé préférable de trier la liste des termes obtenue après extraction, si elle devait comporter l'un ou l'autre des éléments d'une adresse, le potentiel bruit occasionné par ces éléments pouvant être éliminé bien plus rapidement que lors de la constitution du corpus.

### **3.2.5. Extraction des termes**

#### **3.2.5.1. TermoStat**

TermoStat est un système hybride d'extraction semi-automatique des termes, mis au point par Patrick Drouin (2003). TermoStat extrait les termes candidats d'un corpus, c'est-à-dire qu'il établit une liste de toutes les unités qui pourraient probablement être des termes. Un terminologue examine ensuite cette liste de termes candidats, afin de ne retenir que les termes qu'il estime pertinents. C'est pourquoi il s'agit d'extraction semi-automatique et non d'extraction automatique. Le rôle du terminologue est en effet absolument nécessaire, car il doit évaluer le

potentiel terminologique d'un terme candidat pour pouvoir le valider en tant que terme. La dimension sémantique est cruciale pour la validation des termes candidats et c'est justement ce que l'ordinateur ne peut pas prendre en charge.

Il existe plusieurs systèmes d'extraction de termes : les systèmes linguistiques, les systèmes statistiques et les systèmes hybrides, qui exploitent les avantages des deux premiers systèmes. TermoStat est un système hybride, car il applique un traitement linguistique puis un traitement statistique au corpus d'analyse, c'est-à-dire le corpus duquel seront extraits les termes candidats. D'ailleurs, selon Cabré, Estopà et Vivaldi (2001 : 29), un système hybride donne de meilleurs résultats qu'un système uniquement linguistique ou uniquement statistique.

Le traitement linguistique appliqué par TermoStat consiste à lemmatiser et à étiqueter le corpus d'analyse, puis à repérer les termes candidats sur la base de leur construction (patron). Un patron représente les catégories grammaticales de chaque mot composant un terme complexe. Les patrons peuvent être très variés, présentant des combinaisons de catégories grammaticales plus ou moins complexes. Considérons, par exemple, *parabolic flight*, *solar wind*, *geostationary orbit* et *deep space*. Nous pouvons remarquer que ces quatre exemples sont construits selon le même patron, un adjectif suivi d'un nom. Puisque ce sont tous les quatre des termes, nous pouvons en déduire qu'une unité lexicale construite selon le même patron (adjectif + nom) peut être un terme. TermoStat recense donc toutes les unités lexicales correspondant à un patron défini à l'avance.

TermoStat filtre ensuite ces résultats en appliquant un calcul statistique basé sur la fréquence des termes candidats dans le corpus d'analyse et sur leur fréquence dans le corpus de référence, c'est-à-dire le corpus de langue générale utilisé comme référence (Drouin, 2003). Ce calcul compare la fréquence relative d'un terme candidat dans les deux corpus et lui attribue un score. Plus le score est élevé, plus la probabilité qu'il soit un terme appartenant au domaine du corpus d'analyse est élevée (*Ibid.*). Pour l'anglais, le corpus de référence contient environ huit millions d'occurrences provenant pour la moitié d'articles du journal montréalais *The Gazette* publiés entre mars 1989 et mai 1989 et, pour l'autre moitié, du British National Corpus<sup>1</sup>.

TermoStat comporte de nombreux avantages, parmi lesquels le fait qu'il repère aussi bien des termes simples que des termes complexes, ce qui n'est pas le cas de tous les systèmes (Lemay, L'Homme, & Drouin, 2005 : 228). Le score, qui prend en compte un calcul basé sur la fréquence des termes candidats et non uniquement leur fréquence, est également un réel avantage de ce système. Ainsi, même si un terme a une fréquence très basse, il peut tout de même être repéré par

---

<sup>1</sup> TermoStat, <http://termostat.ling.umontreal.ca/index.php>, consulté le 17 mars 2014.

TermoStat. Nous avons donc décidé d'extraire les termes du corpus constitué des communiqués de presse de l'ESA avec TermoStat principalement pour ces deux raisons.

Nous avons donc d'abord procédé à l'extraction des termes candidats nominaux, puis des termes candidats verbaux. Si nous avons décidé d'extraire également les termes candidats verbaux, c'est parce que nous pensons que les verbes sont tout aussi intéressants que les noms en terminologie. Il est en effet aujourd'hui reconnu que les verbes peuvent tout à fait véhiculer des connaissances spécialisées (L'Homme, 2012). Dans notre cas, et à partir de nos propres connaissances du domaine spatial, nous supposons que de nombreux verbes pourraient avoir un sens spécialisés dans ce domaine et qu'il serait donc intéressant de les examiner de plus près dans notre travail. En effet, les verbes « voler »/« *to fly* », « atterrir »/« *to land* », « lancer »/« *to launch* » ou « propulser »/« *to propel* », par exemple, véhiculent un sens spécialisé dans le domaine spatial. Ainsi, l'inclusion des verbes à notre étude nous permettrait de représenter au mieux le domaine, d'une part, et, d'autre part, d'obtenir une liste de termes diversifiée, susceptible de contenir de nombreux candidats à la déterminologisation. L'extraction des termes candidats verbaux nous permettra également d'asseoir ou non notre intuition à propos du verbe dans ce domaine.

### **3.2.5.2.Extraction des termes candidats nominaux**

La liste de termes candidats nominaux obtenue avec TermoStat était globalement conforme à nos attentes, même s'il y avait plus de bruit que ce que nous avons pensé au départ. Nous pensons pouvoir expliquer ce bruit par le degré de spécialisation du corpus, qui était plus faible que s'il s'était agi de textes provenant de revues scientifiques ou de manuels, par exemple.

Toutefois, nous avons pu rapidement diminuer le bruit en supprimant immédiatement de nombreux termes candidats, comme *http*, *int*, *media@esa*, ou *rue*. Les communiqués de presse ont un contenu particulier, par rapport à d'autres genre textuels, ce qui peut expliquer ce bruit. En effet, des adresses, des adresses électroniques ou des URL apparaissent bien plus fréquemment dans les communiqués que dans d'autres genres. Certains éléments apparaissent même dans tous les communiqués. Il est alors tout à fait compréhensible que TermoStat ait retenu ces éléments en tant que termes candidats, puisqu'ils ont une fréquence relative bien plus élevée dans le corpus d'analyse que dans le corpus de référence.

Prenons l'exemple de *rue*. Les communiqués de presse contiennent un certain nombre d'adresses (d'autres agences spatiales, de personnes à contacter, etc.), dont certaines se trouvent dans un pays francophone. De plus, même si *rue* existe en anglais, il est bien moins fréquent que le *rue* français utilisé dans des adresses figurant dans les communiqués de presse. Si TermoStat a retenu

*rue* comme terme candidat, c'est donc parce qu'il apparaît significativement plus fréquemment dans le corpus d'analyse que dans le corpus de référence.

Les unités retenues comme termes candidats telles que *information* ou *contact* s'expliquent par le fait que les mêmes phrases types reviennent dans la plupart des communiqués de presse. Leur fréquence est effectivement très haute : 1924 et 885 respectivement. *Information* revient principalement dans les phrases « *For further information ...* » et « *For more information...* », ainsi que dans leurs variantes « *For additional information ...* », « *For more detailed information ...* » et « *For information on ...* ». *Contact* apparaît le plus souvent dans l'expression « *please contact : ...* » ou simplement « *contact : ...* ».

Ces quelques exemples pour montrer que les résultats de l'extraction de termes sont typiques du genre particulier des communiqués de presse. Néanmoins, nous avons rapidement pu éliminer ce bruit.

### **3.2.5.3.Extraction des termes candidats verbaux**

Nous avons pu constater un certain bruit dans la liste de termes candidats verbaux établie par TermoStat. En effet, de nombreux verbes font partie des phrases typiques des communiqués de presse, comme *attend* et *wish* dans « *Media representatives wishing to attend...* », ou du vocabulaire général, comme *schedule* ou *follow*.

Peu de verbes appartiennent clairement au domaine du spatial. Mentionnons tout de même *launch*, *orbit* et *probe*, qui ont également été retenus lors de l'extraction de termes candidats nominaux. Les formes canoniques nominales et verbales de ces unités se confondent en anglais et, si nous les retenions, nous aboutirions à une liste de termes comprenant des doublons, ce qui n'est pas souhaitable dans le cadre de nos recherches.

Au vu du peu de termes candidats verbaux validés en tant que termes et dont la forme canonique est différente du terme nominal déjà retenu, nous avons décidé de continuer cette étude uniquement sur la base des termes nominaux obtenus avec la première extraction.

## **3.2.6. Sélection des termes**

### **3.2.6.1. Validation des termes candidats**

La liste de termes définitive, obtenue à partir de l'extraction de TermoStat, sert un double objectif. D'abord, elle constitue la base sur laquelle le premier corpus de presse (orienté termes)

est construit<sup>1</sup> : au moins un des termes de la liste doit figurer dans chaque article constituant le corpus. Cette liste de termes appartenant au domaine du spatial servira ensuite de liste de base pour les unités à observer dans les deux corpus de presse. Il n'est pas exclu que d'autres termes fassent par la suite l'objet de nos observations, en fonction de ce que nous observerons dans le corpus.

Nous nous sommes alors posé deux questions fondamentales, la première concernant le nombre de termes à retenir pour cette étude et la seconde concernant leur appartenance au domaine du spatial et surtout comment la déterminer.

Le nombre de termes à retenir est arbitraire, a priori, car tous les termes du domaine spatial extraits par TermoStat sont potentiellement intéressants et peuvent potentiellement faire partie de cette étude. Toutefois, en premier lieu, il n'était pas possible d'extraire semi-automatiquement absolument tous les termes appartenant au domaine spatial et, en deuxième lieu, il a fallu déterminer un nombre de termes qui soit suffisant pour un travail de cette ampleur sans pour autant être minimaliste ou trop optimiste. Nous avons donc décidé d'exploiter cent termes pour commencer, afin d'avoir un point de départ réaliste, mais nous souhaiterions intégrer d'autres termes par la suite, en vue d'approfondir nos recherches et d'étudier d'autres unités.

Trois facteurs ont été essentiels pour déterminer l'appartenance au domaine du spatial des termes candidats : le score de spécificité, l'observation en contexte et l'intuition, et la vérification dans les ressources.

Tout d'abord, le score de spécificité de TermoStat a pour but de repérer « le vocabulaire spécifique à un sous-corpus par rapport à l'ensemble d'un corpus »<sup>2</sup>, ce qui convient parfaitement à l'objectif de notre étude.

Nous avons considéré les termes candidats dont le score de spécificité était le plus élevé, car plus leur score est élevé, plus il est probable qu'ils soient véritablement des termes.

L'efficacité du score de spécificité, par rapport à la prise en compte de la fréquence seulement, est indéniable. Nous avons en effet pu constater que certains termes candidats ont un score élevé malgré une fréquence relativement peu élevée. *Astronaut*, par exemple, a une fréquence de 1118 et un score de 109,71, tandis qu'*information* a une fréquence de 1924 et un score de 88,27.

---

<sup>1</sup> Voir point 3.3.

<sup>2</sup> Manuel TermoStat, [http://termostat.ling.umontreal.ca/doc\\_termostat/doc\\_termostat.html](http://termostat.ling.umontreal.ca/doc_termostat/doc_termostat.html), consulté le 17 mars 2014.

En nous fiant à nos propres connaissances du domaine, nous avons rapidement pu valider un certain nombre de termes candidats, comme *spacecraft*, *planet*, *comet*, *asteroid*, *orbit* ou *star*, et en éliminer d'autres, n'appartenant pas au domaine du spatial, comme *information*, *data*, *press*, *e-mail* ou *press conference*.

L'observation en contexte, qui est un autre des atouts de TermoStat, a permis de dissiper nos doutes quant à l'appartenance au domaine du spatial de certains termes candidats. Nous avons donc pu valider des termes tels que *earth observation*, *ground station* ou *rover*. L'observation en contexte a également été très utile pour vérifier qu'un terme candidat composé de plusieurs éléments était complet. Par exemple, en observant les contextes d'*observation satellite*, nous avons pu constater que 30 fois sur 33 il s'agissait en fait de *Earth observation satellite*. Autrement dit, dans 30 cas il s'agit de satellites d'observation de la Terre et dans trois cas seulement il s'agit d'autres satellites d'observation. Nous avons ainsi pu lever l'ambiguïté de plusieurs termes candidats.

Finalement, lorsque l'observation en contexte était insuffisante et que le doute subsistait, nous avons vérifié dans les ressources spécialisées, en particulier dans le grand dictionnaire terminologique (GDT)<sup>1</sup> et dans Termium<sup>2</sup>.

Déterminer l'appartenance d'un terme au domaine du spatial n'a pas été une tâche aussi simple qu'elle paraissait au premier abord, car il a fallu d'abord définir le domaine spatial. Pour ce faire, nous nous sommes basée sur les classifications en domaines du GDT et de Termium. Nous avons donc décidé que, pour appartenir au domaine du spatial, un terme doit obligatoirement appartenir au moins à *l'aéronautique* ou à *l'astronomie* dans le GDT. La classification en domaines de Termium étant plus détaillée que celle du GDT, nous n'avons pas pu être aussi catégorique, mais avons vérifié que tous les termes retenus appartenaient au moins à un domaine en relation avec le spatial (*engins spatiaux*, *aéronautique*, *physique spatiale*, *astronomie* ou *astrophysique et cosmographie*). De cette manière, nous avons vérifié les cent termes de la liste établie, avec toutefois une exception. Le GDT et Termium n'enregistrent pas *universe* sous un domaine lié au spatial, ce qui peut paraître étrange, vu son appartenance indéniable au spatial. Nous avons tout de même pris la décision de le garder, car son absence dans les ressources spécialisées peut être la preuve qu'il a déjà imprégné la langue générale. Il nous est donc apparu comme un candidat intéressant pour cette étude.

---

<sup>1</sup> Le grand dictionnaire terminologique, Office québécois de la langue française, Québec, <http://gdt.oqlf.gouv.qc.ca/>, consulté le 19 mars 2014.

<sup>2</sup> Termium, Bureau de la traduction, Travaux publics et Services gouvernementaux, Canada, <http://www.btb.termiumplus.gc.ca/>, consulté le 19 mars 2014.

### 3.2.6.2. Liste de termes définitive

La liste de termes définitive contient plusieurs types de termes : des termes simples (*satellite, planet, comet*) et des termes composés (*earth observation satellite, microwave background*) ; des termes pouvant appartenir à plusieurs domaines plus ou moins différents les uns des autres<sup>1</sup> (*mission, probe*) et des termes appartenant strictement au domaine du spatial (*weightlessness, geostationary orbit*) ; des termes désignant des organismes (*nasa*), des corps célestes (*moon, neutron star*), des noms de satellites (*metop, envisat*), etc. La liste recouvre en fait très largement les différents sous-domaines du spatial, ce qui garantit, en quelque sorte, la représentativité du corpus que nous construisons sur cette base<sup>2</sup>. C'est un point très important, étant donné que le but de cette étude est de s'appuyer sur des données réelles pour étudier la déterminologisation et son impact sur la langue générale.

Certains termes, comme *mission* ou *programme*, sont polysémiques. Ils peuvent, en effet, non seulement appartenir à de nombreux domaines différents, mais ils sont aussi relativement courants dans la langue générale. Néanmoins, dans le but de construire des corpus aussi représentatifs de la langue générale que possible, et donc de ne pas biaiser, en quelque sorte, les données, nous avons fait le choix de ne pas intervenir sur l'ambiguïté de certains termes.

La question de la polysémie est importante et mérite que nous nous y attardions quelque peu car, bien que notre décision d'inclure tous les cent termes de la liste sans rien exclure paraît évidente, elle est le résultat d'une longue réflexion. Prenons les exemples de *space* et de *flight* afin d'illustrer notre propos. *Space* figure dans 29 fiches dans le GDT<sup>3</sup> et dans 18 fiches dans Termium<sup>4</sup>, dans des domaines aussi nombreux que variés. Citons notamment, dans le GDT, l'aéronautique, le bâtiment, l'imprimerie, la médecine, la publicité et les télécommunications, et, dans Termium, l'astronomie, la géographie mathématique, la presse audiovisuelle et le tricot au crochet ! Quant à *flight*, il apparaît dans 31 fiches dans le GDT<sup>5</sup> et dans 22 fiches dans Termium<sup>6</sup>. Et ce, toujours dans des domaines très variés : l'armée, le génie mécanique, le verre ou la zoologie, dans le GDT et l'athlétisme, le golf, le tir à l'arc ou le transport aérien, dans Termium. Le risque de polysémie est donc inévitable, ces deux termes n'étant bien évidemment pas des cas isolés.

---

<sup>1</sup> Par exemple, *mission* peut appartenir à des domaines aussi différents que l'aéronautique et l'éducation, dans le GDT.

<sup>2</sup> Voir la constitution du corpus orienté termes, au point 3.3.

<sup>3</sup> Consulté le 17 mars 2014.

<sup>4</sup> Idem.

<sup>5</sup> Consulté le 18 mars 2014.

<sup>6</sup> Idem.

### **3.3. Corpus de presse – deuxième et troisième corpus de travail**

#### **3.3.1. Choix des textes**

Tout d'abord, notre choix d'exploiter la presse généraliste est fortement lié à notre choix d'exploiter les communiqués de presse, vu le rapport existant entre ces deux genres (voir point 3.2.2.). Ensuite, il s'agit également d'un choix pratique, car les archives de journaux nous permettent d'accéder facilement et rapidement à de nombreux textes, sur une période de temps plus ou moins longue. Enfin, la presse généraliste est un genre qui relève de la langue générale, ce qui se prête bien à notre étude.

Puisque nous avons décidé de limiter nos recherches à l'anglais du Royaume-Uni, nous avons sélectionné des quotidiens largement diffusés en Angleterre. Nous aurions pu utiliser également des publications provenant d'Écosse ou d'Irlande, mais nous avons préféré limiter la quantité de données, quitte à élargir nos recherches par la suite. Il serait en effet intéressant de voir dans quelle mesure les observations faites dans les corpus constitués d'articles émanant de *The Guardian* et de *The Independent* rejoignent celles que l'on pourrait faire dans des journaux tels que *The Irish News* et *The Herald*. De plus, nous avons préféré une presse nationale plutôt qu'une presse régionale, afin d'éviter le risque de retrouver des régionalismes dans les résultats de nos recherches.

En outre, comme nous nous situons dans un cadre européen, nous pensons que ces deux journaux anglais se prêtent particulièrement bien à nos recherches puisqu'ils sont largement diffusés dans les pays européens également. Ce sont des quotidiens très internationaux et, de ce fait, il est légitime de penser qu'ils peuvent exercer une certaine influence dans toute l'Europe, du moins une influence linguistique. Ainsi, cette décision est cohérente avec le choix que nous avons fait de travailler à partir des communiqués de presse de l'ESA.

#### **3.3.2. Critères de constitution des corpus**

##### **3.3.2.1. Considérations générales**

Nous avons construit deux corpus de langue générale, plus précisément de langue générale telle qu'elle est employée dans la presse généraliste. Nous avons décidé d'exploiter ce genre textuel, car, comme nous l'avons déjà dit, nous reprenons la méthodologie de Condamines et Picton (À paraître a) et, d'autre part, car nous pensons que c'est un genre particulièrement propice à

l'observation du phénomène de la déterminologisation, comme nous l'avons expliqué au point 3.2.2.

Dans le corpus orienté termes comme dans le corpus aléatoire, nous avons essayé de sélectionner des articles traitant de sujets différents, toujours dans l'optique d'obtenir la meilleure représentation possible de la langue générale.

### **3.3.2.2. Langue**

Étant donné que nous souhaitons observer l'utilisation dans la langue générale de termes appartenant au domaine du spatial en anglais, il semble préférable, comme le conseillent Bowker et Pearson (2002 : 52), de n'inclure que des textes rédigés par des locuteurs natifs :

*« Texts that have been translated or that have been written by non-native speakers may contain some non-idiomatic expressions. If you want to learn how things are expressed naturally in the LSP in question, it may be best to try to exclude non-original or non-native texts from your corpus ».*

### **3.3.2.3. Taille**

Dans cette section, nous allons reprendre deux des trois critères de Bowker et Pearson (2002 : 45) concernant la taille, à savoir les besoins du projet et la disponibilité des ressources.

Puisqu'il s'agit d'étudier un phénomène linguistique dans la langue générale, nous pensons qu'il vaut mieux privilégier des corpus assez gros, afin de maximiser les probabilités d'apparition de ce phénomène dans cet échantillon de la langue. Comme le dit L'Homme (2004 : 124),

*« L'ensemble des textes est représentatif de ce qu'on souhaite observer. Le corpus est assemblé en fonction de l'élément à étudier et doit comporter un nombre suffisamment élevé d'occurrences de cet élément ».*

Ainsi, dans notre cas, puisque l'élément à étudier est la déterminologisation et que nous ne savons pas dans quelle mesure elle se produit dans la langue générale, il vaut mieux construire de gros corpus pour augmenter les chances d'y voir se produire ce phénomène. Si les corpus sont trop petits, en effet, nous risquons de ne pas pouvoir observer la déterminologisation ou d'y arriver mais dans une mesure trop faible pour pouvoir essayer de dégager des tendances générales, qui est le but que nous nous étions fixé.

Puisqu'un corpus est un échantillon représentatif de la langue<sup>1</sup>, nous avons soigneusement sélectionné des articles provenant de rubriques différentes, rédigés par des auteurs différents, à des dates différentes et qui sont de longueur différente. En somme, nous avons essayé de respecter ce que préconise L'Homme (2004 : 129) :

« il convient, dans un premier temps, de sélectionner des textes différents pour assurer une certaine représentativité. Par exemple, un corpus de 500 000 mots contenant 20 textes (écrits par des auteurs différents) est plus équilibré, donc plus représentatif, qu'un corpus de même taille contenant trois documents volumineux ».

De cette manière, les observations faites dans un corpus équilibré sont plus fiables que dans un corpus non équilibré (L'Homme, 2004 : 129).

Comme pour le corpus des communiqués de presse de l'ESA, les ressources choisies sont très faciles d'accès, puisqu'il s'agit d'archives de journaux. Pour récupérer les articles des deux journaux sélectionnés, nous avons utilisé la plateforme LexisNexis®, qui permet de retrouver les archives de très nombreux journaux à travers le monde. Nous avons donc choisi *The Guardian* et *The Independent* non seulement pour les raisons évoquées plus haut, mais aussi pour leur accessibilité sur LexisNexis®. Ils sont en effet accessibles intégralement dans la période qui nous intéresse, c'est-à-dire du 1<sup>er</sup> janvier 1998 au 31 décembre 2013.

### 3.3.3. Caractéristiques des corpus

Les deux corpus contiennent chacun un total de 19 200 articles, dont 9 600 proviennent de *The Guardian* et 9 600, de *The Independent*. Le corpus orienté termes contient un total de 13 359 485 occurrences et le corpus aléatoire, un total de 10 617 921.

Pour le corpus orienté termes, nous avons sélectionné les 50 premiers articles de chaque mois pour chaque journal, ce qui donne un total de 1 200 articles par année. Pour le corpus aléatoire, nous avons sélectionné 50 articles parus les derniers jours de chaque mois dans des rubriques différentes, pour chaque journal, ce qui donne également un total de 1 200 articles par année.

Un extrait de chaque corpus figure en annexe (annexes 3.2 et 3.3).

Chaque corpus est composé de seize sous-corpus, chacun correspondant à une année. Le détail de la répartition des occurrences est illustré par les tableaux ci-dessous.

---

<sup>1</sup> « *representative sample* » (Bowker & Pearson, 2002 : 11) ; « *representative sample of a defined language* » (Atkins, Clear, & Ostler, 1992 : 2).

<b>Année</b>	<b>Occurrences</b>
1998	795 504
1999	806 046
2000	840 172
2001	797 667
2002	864 738
2003	853 902
2004	862 993
2005	888 785
2006	843 413
2007	840 158
2008	822 689
2009	830 123
2010	870 809
2011	835 683
2012	791 797
2013	815 006
<b>Total</b>	<b>13 359 485</b>

Tableau 3.2 : Nombre d'occurrences dans les sous-corpus du corpus de presse généraliste orienté termes

<b>Année</b>	<b>Occurrences</b>
1998	547 639
1999	594 251
2000	531 204
2001	633 579
2002	650 560
2003	645 090
2004	703 161
2005	698 159
2006	671 081
2007	641 584
2008	667 098
2009	693 989
2010	721 232
2011	747 886
2012	723 883
2013	747 525
<b>Total</b>	<b>10 617 921</b>

Tableau 3.3 : Nombre d'occurrences dans les sous-corpus du corpus de presse généraliste aléatoire

## 4. OBSERVATIONS

Dans la section précédente, nous avons expliqué notre méthodologie. Nous avons donc construit trois corpus, sur lesquels nous basons nos observations. Le corpus des communiqués de presse de l'ESA nous sert principalement de référence pour le discours spécialisé dans le domaine spatial<sup>1</sup>. Les deux corpus de presse nous servent de référence pour le discours général.

Au point 4.1., nous présentons quelques observations quantitatives générales sur la présence ou l'absence de termes appartenant au domaine spatial dans les deux corpus de presse, ce qui a pour but de donner une idée globale de la représentation de ce domaine dans la langue générale. Au point 4.2., nous nous concentrons en particulier sur l'observation détaillée de six unités qui présentent une ambiguïté de forme nom/verbe.

### 4.1. Observations quantitatives globales

#### 4.1.1. Représentation des termes dans les corpus

Avant de nous concentrer sur quelques cas particuliers, nous aimerions présenter quelques données quantitatives concernant la présence ou l'absence de termes dans les deux corpus de presse généraliste. Tout d'abord, seulement 82 % des termes retenus dans la liste obtenue à partir de l'extraction (voir point 3.2.7.2.) apparaissent dans ces deux corpus. Il est étonnant de remarquer que ce pourcentage est le même pour les deux corpus, car, au vu de la méthode utilisée pour la sélection des articles constitutifs du corpus orienté termes (voir point 3.3.), nous aurions pu nous attendre à une plus grande représentation des termes dans ce corpus par rapport au corpus aléatoire.

Les dix termes les plus fréquents dans les deux corpus sont les mêmes, à savoir *programme*, *star*, *launch*, *space*, *flight*, *test*, *Earth*, *mission*, *planet* et *location* (voir tableau ci-dessous), ce qui nous amène à un premier constat : même si les deux corpus ont été construits selon une méthode différente, une même tendance peut se dégager dans les deux. Nous supposons en effet que, puisque les dix termes les plus fréquents sont les mêmes dans les deux corpus, ce sont probablement les termes du domaine spatial les plus répandus dans la langue générale.

---

<sup>1</sup> Nous discutons d'ailleurs ce choix dans notre conclusion.

Termes	Orienté termes		Aléatoire	
	Fréquence	Pourcentage par rapport aux cent termes	Fréquence	Pourcentage par rapport aux cent termes
programme	8 335	17,8	2 790	14,1
star	7 407	15,8	2 838	14,3
launch	4 632	9,9	2 241	11,3
nospace	3 579	7,7	1 457	7,3
flight	2 941	6,3	1 045	5,3
test	2 882	6,2	2 362	11,9
Earth	2 159	4,6	802	4
mission	1 436	3,1	520	2,6
planet	1 277	2,7	508	2,6
location	1 237	2,6	497	2,5
<b>Total</b>	<b>35 885</b>	<b>76,7</b>	<b>15 060</b>	<b>75,9</b>

Tableau 4.1 : Les dix unités les plus fréquentes dans les deux corpus de presse

Néanmoins, nous pouvons nuancer ce propos en considérant le caractère polysémique de certains termes, que nous n'avons pas pris en compte lors de la constitution du corpus orienté termes. Nous ne pouvons alors pas savoir dans quelle mesure leur passage dans la langue générale s'est effectué à partir des termes appartenant au domaine spatial, vu qu'ils peuvent appartenir à de nombreux domaines différents. À titre d'exemple, *programme* peut appartenir à 12 domaines selon les GDT<sup>1</sup> et à 13 domaines selon Termium<sup>2</sup>.

Dix-huit termes sont absents de chaque corpus, dont douze sont absents des deux corpus. Il s'agit de *cargo bay*, *geostationary orbit*, *geostationary satellite*, *ground segment*, *ion engine*, *meteorological satellite*, *metop*, *navigation satellite*, *parabolic flight*, *parking orbit*, *telecommunication satellite* et *upper stage*. Les termes absents du corpus aléatoire mais présents dans le corpus orienté termes sont *atmospheric drag*, *exobiology*, *polar orbit*, *space system*, *space weather* et *weightlessness*; les termes absents du corpus orienté termes mais présents dans le corpus aléatoire sont *Earth observation*, *Earth observation satellite*, *elliptical orbit*, *envisat*, *gravity field* et *perigee*.

En reprenant l'idée de Meyer et Mackintosh (2000 : 214) selon laquelle certains termes seraient plus susceptibles que d'autres de passer dans la langue générale, nous pourrions supposer que les douze termes absents des deux corpus sont les moins susceptibles de pénétrer la langue générale.

<sup>1</sup> Le grand dictionnaire terminologique, Office québécois de la langue française, Québec, <http://gdt.oqlf.gouv.qc.ca/>, consulté le 15 avril 2014.

<sup>2</sup> Termium, Bureau de la traduction, Travaux publics et Services gouvernementaux, Canada, <http://www.btb.termiumplus.gc.ca/>, consulté le 15 avril 2014.

Ces 24 termes sont rares dans la langue générale, vu leur faible fréquence dans les deux corpus de presse. Leur absence d'un des deux corpus, voire des deux, pourrait être due à la manière dont nous avons sélectionné les articles et aussi en partie au hasard. *Envisat*, par exemple, apparaît trois fois dans le corpus aléatoire, mais trois fois dans le même article. Nous aurions pu facilement passer à côté de cet article et nous sommes ainsi probablement passée à côté d'autres unités rares.

Le pourcentage de termes présents dans le corpus aléatoire (82 %) nous donne un élément de preuve pouvant confirmer que le domaine spatial est un domaine qui intéresse le grand public. Les exemples de *microgravity*, *cosmodrome* ou *envisat* prouvent que la presse généraliste parle de sujets touchant au domaine spatial. En voici quelques exemples :

- *Initially, however, it will be a laboratory for **microgravity** experiments which would be hard (though not impossible) to carry out on Earth<sup>1</sup>.*
- *The first tenants, who moved in late last year, brought a treadmill, and ran on it before lunch and before dinner to offset bone loss and muscle wastage that goes with life in **microgravity**<sup>2</sup>.*
- *Guy Laliberte joined two astronauts aboard a Russian Soyuz rocket which lifted off at 8.14am UK time from the Baikonur **cosmodrome** in Kazakhstan<sup>3</sup>.*
- *There was no Nasa-style countdown at the **cosmodrome**, only occasional and curt reports blaring from loudspeakers<sup>4</sup>.*
- *The first pictures from Europe's flagship earth observation satellite, **Envisat**, show how coastlines shaped over millions of years are now changing in a generation<sup>5</sup>.*
- ***Envisat**, the biggest and most expensive satellite put into orbit by Europe, was launched on an Ariane 5 rocket from Kourou in French Guiana on 28 February<sup>6</sup>.*

La presse éprouve un réel intérêt pour ce domaine. Même si de nombreux termes n'apparaissent que quelques fois dans le corpus, voire qu'une seule fois, comme nous venons de le mentionner dans le paragraphe précédent, ils apparaissent tout de même et c'est sur ce point que nous souhaitons mettre l'accent. De ce fait, on pourrait penser qu'ils seraient susceptibles de se déterminologiser dans l'avenir et de se diffuser plus largement dans la langue générale. Nous pourrions ainsi supposer que le domaine spatial n'a pas fini de perfuser la langue générale.

---

<sup>1</sup> Corpus orienté termes, 1<sup>er</sup> novembre 2000.

<sup>2</sup> Corpus orienté termes, 2 avril 2001.

<sup>3</sup> Idem, 1<sup>er</sup> octobre 2009.

<sup>4</sup> Idem, 1<sup>er</sup> novembre 2000.

<sup>5</sup> Corpus aléatoire, 30 mars 2002.

<sup>6</sup> Idem.

Vu l'intérêt pour le domaine spatial dans la presse généraliste, nous pouvons faire un parallèle avec Meyer et Mackintosh (2000 : 201) qui soulignent que notre société devient de plus en plus une « société du savoir ».

#### 4.1.2. Comparaison de la fréquence de certaines unités dans les deux corpus

Nous avons remarqué que la plupart des unités apparaissent plus souvent dans le corpus orienté termes que dans le corpus aléatoire, ce qui est très probablement dû à la méthode de sélection des articles constitutifs du corpus orienté termes. En effet, pour être sélectionné, un article devait contenir au moins un terme de la liste établie au point 3.2.7.2. À l'inverse, certaines unités apparaissent plus souvent dans le corpus aléatoire que dans le corpus orienté termes. Si l'on exclut les unités absentes du corpus orienté termes et présentes dans le corpus aléatoire, seules trois unités sont plus fréquentes dans le corpus aléatoire que dans le corpus orienté termes. Elles figurent dans le tableau ci-dessous.

Termes	Orienté termes	Aléatoire
microwave background	1	5
solar wind	1	6
space debris	1	2

Tableau 4.2 : Unités dont la fréquence est plus élevée dans le corpus aléatoire que dans le corpus orienté termes

En observant la fréquence d'une unité par rapport à la fréquence des cent unités retenues pour nos observations, nous avons pu constater que la plupart avaient une fréquence tout à fait comparable dans les deux corpus de presse, leur pourcentage montrant une différence de moins de 0,4 %. La fréquence de neuf unités seulement a une différence de pourcentage de plus de 0,4 %, la plus grande différence étant de 5,7 %. Ces informations sont illustrées dans les deux tableaux suivants.

Termes	Orienté termes		Aléatoire	
	Fréquence	Pourcentage par rapport aux cent termes	Fréquence	Pourcentage par rapport aux cent termes
coverage	769	1,65	452	2,28
Earth	2159	4,62	802	4,04
flight	2 941	6,3	1 045	5,3
launch	4 632	9,9	2 241	11,3
mission	1 436	3,1	520	2,6
programme	8 335	17,8	2 790	14,1
rover	472	1,01	436	2,2
star	7 407	15,8	2 838	14,3
test	2 882	6,2	2 362	11,9

Tableau 4.3 : Unités dont la différence de fréquence est de plus de 0,4 %

Termes	Orienté termes		Aléatoire	
	Fréquence	Pourcentage par rapport aux cent termes	Fréquence	Pourcentage par rapport aux cent termes
atmosphere	1218	2,6	448	2,26
eclipse	266	0,57	118	0,59
moon	802	1,72	338	1,7
observatory	62	0,13	32	0,16
satellite	847	1,81	343	1,73
spacecraft	73	0,16	31	0,16

Tableau 4.4 : Exemples d'unités dont la fréquence par rapport aux cent termes est comparable dans les deux corpus de presse

## 4.2. Formes verbales

Nous venons de présenter quelques observations quantitatives globales et passons maintenant à l'observation de certains cas spécifiques. Nous nous sommes intéressée de plus près aux unités de notre liste de référence (voir point 3.2.7.2.) dont le nom et le verbe partagent la même forme canonique, phénomène caractéristique de l'anglais. Nous souhaitons en effet observer, s'il y en a, les différences d'emploi du verbe et du nom dans les corpus de presse par rapport au corpus des communiqués de presse, le nom étant généralement propre au discours spécialisé et le verbe, au discours général (Condamines & Bourigault, 1999; L'Homme, 2012).

Notre but est de voir dans quelle mesure une unité est diffusée dans la langue générale et nous pensons que certaines de ces différences d'emploi peuvent être de bons indicateurs de la diffusion d'une unité dans la langue générale. De plus, l'existence d'un verbe uniquement en langue générale (dont la forme canonique est la même que celle d'un terme attesté dans le discours spécialisé) peut être un autre indicateur de la diffusion de ce terme dans la langue générale. Ces deux indicateurs nous permettent de supposer que, si l'emploi d'un verbe correspondant à un terme appartenant au domaine spatial est propre au discours général, cela serait un bon indice de la diffusion de ce terme dans la langue générale.

Nous confronterons ces observations à ce qui est attesté dans les ressources générales et spécialisées et mettrons en évidence les différences existant entre la réalité et ce qui est attesté. Pour ce faire, nous consultons le *Collins*, le *Cambridge* et l'*Oxford* pour les ressources générales, et le Grand dictionnaire terminologique (GDT) et Termium pour les ressources spécialisées. Nous avons choisi ces trois dictionnaires généraux, parce qu'ils décrivent l'anglais britannique, ce qui est tout à fait approprié pour notre étude, vu que nos corpus de presse sont constitués d'articles provenant de journaux anglais. Cela n'aurait en effet pas eu de sens de nous appuyer sur des dictionnaires américains, vu les nombreuses nuances et subtilités entre l'anglais américain et l'anglais britannique. Pour ce qui est des ressources spécialisées, nous avons choisi ces deux grandes banques de terminologie, non seulement car elles sont accessibles intégralement sur Internet, mais aussi car elles ne se limitent pas à un seul domaine, loin de là. Cette caractéristique nous a été très utile, notamment dans le cas d'unités polysémiques. Nous avons pu nous rendre compte des nombreux domaines auxquels pouvaient appartenir certaines unités.

Nous nous sommes donc intéressée de plus près à six unités lexicales : *eclipse*, *orbit*, *launch*, *probe*, *rocket* et *star*. Il convient de souligner que nous n'avons pas désambiguïsé l'emploi de ces six unités dans les corpus. Si cela ne semble pas avoir de conséquences pour *eclipse*, *orbit*, *rocket* et *star*, *probe* et *launch* pourraient par contre être différents. Les quatre premiers termes, en effet, n'appartiennent pas à autant de domaines et sous-domaines, ni à des domaines aussi différents les uns des autres, que *probe* et *launch*. Dans le GDT<sup>1</sup>, *probe* peut appartenir à 20 domaines et sous-domaines et, dans Termium<sup>2</sup>, à 31 domaines et sous-domaines. Quant à *launch*, il peut appartenir à 11 domaines et sous-domaines dans le GDT<sup>3</sup> et à 15 domaines et sous-domaines dans Termium<sup>4</sup>.

---

<sup>1</sup> Le grand dictionnaire terminologique, Office québécois de la langue française, Québec, <http://gdt.oqlf.gouv.qc.ca/>, consulté le 15 mai 2014.

<sup>2</sup> Termium, Bureau de la traduction, Travaux publics et Services gouvernementaux, Canada, <http://www.btb.termiumplus.gc.ca/>, idem.

<sup>3</sup> Consulté le 27 mars 2014.

<sup>4</sup> Idem.

En comparaison, *eclipse* peut appartenir à trois domaines et sous-domaines dans le GDT<sup>1</sup> et dans Termium<sup>2</sup>. L'appartenance à autant de domaines différents, parfois assez éloignés, comme l'*astronautique* et la *psychologie* dans le cas de *probe* dans le GDT, illustre la polysémie de ces deux unités.

Il convient donc de faire preuve d'une certaine prudence lors de l'observation de ces unités dans les deux corpus, au vu de leur caractère polysémique et de la non-désambiguïsation de leurs occurrences. En outre, puisque nous concentrons nos observations sur six unités, nous espérons pouvoir repérer certains comportements, similaires ou divergents, et ainsi tenter de dégager une tendance générale.

Puisque le phénomène que nous étudions semble se produire sous la forme d'un flux (Condamines & Picton, À paraître a; Meyer & Mackintosh, 2000), ce que nous avons relevé pourrait illustrer ce flux. Grâce à l'observation de leur comportement dans le discours général, par rapport au discours spécialisé, nous espérons pouvoir observer le « chemin » que prennent ces termes en passant dans la langue générale. Cela nous permettra également de voir dans quelle mesure ils sont diffusés dans la langue générale, c'est-à-dire de voir si « l'essence du sens terminologique demeure dans le mot déterminologisé » ou si « on assiste à des changements conceptuels plus considérables, l'extension sémantique du mot s'éloignant alors nettement de celle du terme d'origine » (Meyer & Mackintosh, 2000 : 202). Autrement dit, un terme peut soit passer dans la langue générale et conserver son sens spécialisé, le degré de compréhension du concept étant alors plus faible, soit s'éloigner beaucoup plus de son sens spécialisé et recouvrir une autre réalité (*Ibid.*).

Nous avons donc étudié chaque unité une à une, sous différents aspects. Nous avons comparé chacun de ces aspects entre les deux corpus de presse et le corpus des communiqués de presse de l'ESA, afin de repérer les différences de comportement des unités. Pour chaque unité, nous avons observé :

- la fréquence des formes verbales par rapport aux formes nominales ;
- la transitivité des verbes, que nous avons comparée aux indications des ressources générales<sup>3</sup> ;
- les objets des verbes transitifs<sup>1</sup> ;

---

<sup>1</sup> Consulté le 15 mai 2014.

<sup>2</sup> Idem.

<sup>3</sup> Sur ce point, nous ne nous sommes pas référée aux ressources spécialisées, car le GDT et Termium ne donnent aucune indication concernant la transitivité des verbes.

- les différentes formes de l'unité ;
- la répartition des formes dans les sous-corpus de chaque corpus.

Après avoir présenté nos observations pour chaque unité, nous tenterons de dégager une tendance et de formuler une conclusion, en nous concentrant principalement sur des phénomènes qui semblent se produire à chaque fois et qui pourraient caractériser le flux que serait la déterminologisation.

Nous faisons la distinction entre les unités dont la forme verbale est attestée par les ressources spécialisées et celles dont la forme verbale ne l'est pas, car nous voyons une légère nuance dans le phénomène à observer : dans un premier temps, nous observons les unités dont la forme verbale est attestée dans les ressources spécialisées, à savoir *launch*, *eclipse* et *orbit*, afin de voir si ces verbes se comportent de façon différente dans les corpus de presse par rapport au corpus des communiqués de presse, puis, dans un deuxième temps, nous observons les unités dont la forme verbale n'est pas attestée par les ressources spécialisées, à savoir *probe*, *rocket* et *star*, mais qui existe dans les corpus de presse. Dans le premier cas, ce sont précisément les différences de comportement qui peuvent indiquer la diffusion d'une unité dans la langue générale et, dans le second cas, c'est justement le fait que le verbe existe dans le discours général mais pas dans le discours spécialisé qui est indicateur de la diffusion de cette unité dans la langue générale.

#### **4.2.1. Unités dont la forme verbale est attestée dans les ressources spécialisées**

##### **4.2.1.1. LAUNCH**

Le verbe *to launch* est attesté dans Termium, dans le domaine *lancement et manœuvres dans l'espace*, depuis au moins 2005<sup>2</sup>. Le pourcentage de formes verbales relevées par rapport aux formes nominales est de 77,2 % dans le corpus aléatoire, de 68,8% dans le corpus orienté termes et de 60,6 % dans le corpus des communiqués de presse de l'ESA (voir tableau ci-dessous). Nous pouvons déjà constater qu'il y a un pourcentage non négligeable de formes verbales par rapport aux formes nominales du terme *launch* dans les trois corpus, ainsi qu'une augmentation importante en fonction du type de corpus.

---

<sup>1</sup> Puisque nous nous intéressons à la transitivité des verbes, l'observation des objets de ces verbes vient compléter le point sur la transitivité.

<sup>2</sup> Puisque l'indication de date dans les fiches du GDT et de Termium ne précise pas s'il s'agit de la date de création ou de modification de la fiche, nous ne pouvons pas savoir quand exactement une fiche a fait son apparition dans ces ressources.

LAUNCH	Aléatoire	Orienté termes	Communiqués de presse
Fréquence des verbes	1731	3188	1335
Pourcentage de verbes par rapport aux noms	77,2	68,8	60,6

Tableau 4.5 : Proportion des formes verbales de LAUNCH par rapport aux formes nominales

Nous avons ensuite relevé, dans le corpus des communiqués de presse, que les formes du verbe *to launch* sont toujours transitives et le plus souvent à la forme passive. La forme *launched* comporte 335 occurrences à la forme passive contre 149 à la forme active (voir tableau ci-dessous).

LAUNCHED	Fréquence	Pourcentage par rapport au nombre d'occurrences total de LAUNCH
launched	62	4,6
launched aboard	5	0,4
launched as	2	0,15
launched at	6	0,5
launched atop	2	0,15
launched between	2	0,15
launched by	41	3,1
launched for	2	0,15
launched from	24	1,8
launched in	83	6,2
launched into	22	1,6
launched on	59	4,4
launched to	25	1,9
<b>Total formes passives</b>	335	25,1
<b>Total <i>launched</i></b>	484	36,3

Tableau 4.6 : Répartition de la forme passive *launched* dans le corpus des communiqués de presse de PESA

En revanche, dans les deux corpus de presse, nous avons relevé un emploi transitif et un emploi intransitif de ce verbe. Voici quelques-uns des exemples de l'emploi intransitif :

- *The latest service is Sky Movies Active which **will launch** next month and which has a far clearer commercial application.*<sup>1</sup>

<sup>1</sup> Corpus orienté termes, 2 avril 2001.

- *The Now TV "day pass" sports offering **will launch** in spring and aims to attract fans looking to watch live coverage on devices including PCs and tablets.<sup>1</sup>*
- *A new campaign **will launch** in the form of national press and radio adverts from Saturday, emphasising the importance of hygiene in preventing the spread of the virus.<sup>2</sup>*
- *BT's Freeview/ broadband hybrid set-top box **will launch** in the summer.<sup>3</sup>*
- *The new model has been tested overseas and **will launch** worldwide mid-October.<sup>4</sup>*
- *That was until Tuesday, when the site finally **launched** here.<sup>5</sup>*

Ces exemples semblent tous relever du domaine commercial, ce qui peut déjà être une indication de certains changements sémantiques survenus lors du passage de *launch* dans la langue générale. D'ailleurs, on ne peut pas affirmer avec certitude que *launch* est passé du domaine spatial au discours général puis au domaine commercial. Ce cas nous montre bien que les mouvements entre les langues de spécialité et avec la langue générale ne sont pas faciles à décrire. Peut-être *launch* est-il passé dans la langue générale après être passé dans le domaine commercial, ou bien il est passé dans le domaine spatial à partir du domaine commercial. Nous manquons de données pour le savoir, mais il est intéressant de constater que le chemin que peut prendre une unité telle que *launch* n'est pas nécessairement clair ou direct.

Nous avons ensuite observé les objets de *launch*, lorsqu'il est employé transitivement. Nous avons d'abord relevé des unités telles que *satellite*, *observatory*, *campaign*, *project* ou *programme* dans le corpus des communiqués de presse, puis des unités telles que *campaign*, *bid*, *inquiry*, *attack* et *war*, dans les deux corpus de presse. Nous en avons reproduit quelques exemples ci-dessous.

- *This logo will be displayed for the first time on the fairing of the Ariane-5 rocket on the day it **launches** the XMM spacecraft.<sup>6</sup>*
- *The Independent **is launching** a daring challenge to see who really is best at finding those hidden equity gems.<sup>7</sup>*

---

<sup>1</sup> Corpus orienté termes, 1<sup>er</sup> février 2013.

<sup>2</sup> Corpus aléatoire, 31 décembre 2012.

<sup>3</sup> Corpus orienté termes, 2 janvier 2006.

<sup>4</sup> Idem, 1<sup>er</sup> octobre 2007.

<sup>5</sup> Corpus aléatoire, 29 avril 2011.

<sup>6</sup> Corpus des communiqués de presse de l'ESA, 29 octobre 1999.

<sup>7</sup> Corpus aléatoire, 31 janvier 2004.

- *In his e-mail, he denied any intention to mislead the authority, and expressed disappointment that private remarks had been leaked to journalists and indicated he **would launch an inquiry** to establish who was to blame.*<sup>1</sup>
- *But this is changing, too - Jagonari, a women's educational centre in London's East End, **has launched a campaign** against domestic violence in Tower Hamlets.*<sup>2</sup>

Nous avons également relevé, dans les corpus de presse, l'emploi du verbe *to launch oneself*, qui semble être propre à la langue générale, car nous ne l'avons pas relevé dans le corpus des communiqués de presse. De plus, il est attesté dans le *Cambridge*<sup>3</sup>, l'*Oxford*<sup>4</sup> et le *Collins*<sup>5</sup>, ce qui vient renforcer cette idée. Cet emploi apparaît 9 fois dans le corpus aléatoire et 28 fois dans le corpus orienté termes. En voici quelques exemples :

- *The Oscar-winning actress, who is famed for her exotic exercise regime and macrobiotic lifestyle, is facing a fierce critical backlash after attempting to **re-launch herself** as a New Age version of America's domestic guru, Martha Stewart, with a new lifestyle website called Goop.*<sup>6</sup>
- *I **launched myself** down Via Nazionale toward our rendezvous.*<sup>7</sup>
- *From this stage, by way of showing up the idleness of the recumbent Arethusa, he **launches himself** to swim in the clean but chilly waters, kept cool by the constant currents of flowing spring water that give Courances its name.*<sup>8</sup>
- *"It is quite daunting to **launch yourself** overseas when you can't be sure what will be provided."*<sup>9</sup>
- *Contestants dressed as old women, sumo wrestlers and suicidal businessmen **launched themselves** into a fairly accurate recreation of a trench scene in the Somme while a gaggle of wet spectators watched in mild curiosity.*<sup>10</sup>

En ce qui concerne les différentes formes de *launch*, nous n'avons pas relevé de différence majeure. Les quatre formes que nous avons observées (*launch*, *launches*, *launched* et *launching*) apparaissent dans les trois corpus. Nous pouvons toutefois relever que *launch* et *launches* apparaissent sensiblement plus souvent dans le corpus des communiqués de presse que dans les

<sup>1</sup> Corpus orienté termes, 1<sup>er</sup> juin 2006.

<sup>2</sup> Idem, 1<sup>er</sup> janvier 2013.

<sup>3</sup> Cambridge Dictionaries Online, <http://dictionary.cambridge.org>, consulté le 13 juin 2014.

<sup>4</sup> Oxford Dictionaries, <http://www.oxforddictionaries.com>, idem.

<sup>5</sup> Collins Dictionaries, <http://www.collinsdictionary.com>, idem.

<sup>6</sup> Corpus aléatoire, 30 septembre 2008.

<sup>7</sup> Idem, 31 mars 2003.

<sup>8</sup> Idem, 30 juin 2004.

<sup>9</sup> Corpus orienté termes, 1<sup>er</sup> novembre 2000.

<sup>10</sup> Idem, 1<sup>er</sup> décembre 2008.

corpus de presse, tandis que la tendance s'inverse en ce qui concerne *launched* et *launching* (voir tableau ci-dessous). Cette observation n'est que le reflet de la toute première observation que nous avons présentée concernant ce terme, à savoir que le verbe *to launch* apparaît plus fréquemment dans les corpus de presse que dans le corpus des communiqués de presse de l'ESA.

LAUNCH	Aléatoire		Orienté termes		Communiqués de presse	
	Fréquence	%	Fréquence	%	Fréquence	%
<b>Total</b>	2241	100	4632	100	2204	100
<b>launch</b>	862	38,5	2397	51,7	1541	69,9
<b>launched</b>	1036	46,2	1531	33,1	158	7,2
<b>launches</b>	118	5,3	330	7,1	486	22,1
<b>launching</b>	225	10	374	8,1	19	0,8

Tableau 4.7 : Répartition des différentes formes de LAUNCH dans les trois corpus de travail

#### 4.2.1.2. ECLIPSE

Le verbe *to eclipse* est attesté dans les ressources spécialisées depuis au moins 1985<sup>1</sup>, dans le domaine de l'*astronomie*, dans le GDT. Le pourcentage de formes verbales par rapport aux formes nominales est de 55,9 % dans le corpus aléatoire, de 44,4 % dans le corpus orienté termes et de 3,7 % dans le corpus de communiqués de presse de l'ESA (voir tableau ci-dessous). Ces premières observations nous indiquent que le verbe *to eclipse* est bien plus souvent utilisé dans les corpus de presse que dans le corpus des communiqués de presse.

ECLIPSE	Aléatoire	Orienté termes	Communiqués de presse
<b>Fréquence des verbes</b>	66	118	1
<b>Pourcentage de verbes par rapport aux noms</b>	55,9	44,4	3,7

Tableau 4.8 : Proportion des formes verbales de ECLIPSE par rapport aux formes nominales

Nous avons relevé que toutes les occurrences du verbe *to eclipse* sont transitives, aussi bien dans les corpus de presse que dans le corpus des communiqués de presse. La transitivité du verbe est également attestée dans le *Cambridge*, l'*Oxford* et le *Collins*.

Nous avons ensuite observé les objets du verbe dans les trois corpus. Dans les deux corpus de presse, ils sont très variés. En voici quelques exemples :

<sup>1</sup> Puisque l'indication de date dans les fiches du GDT et de Termium ne précise pas s'il s'agit de la date de création ou de modification de la fiche, nous ne pouvons pas savoir quand exactement une fiche a fait son apparition dans ces ressources.

- *There are jitters among some senior Tories that the Prime Minister's high profile during the global financial crisis **will eclipse** the inexperienced Mr Cameron.*<sup>1</sup>
- *Sinn Fein **has eclipsed** the Social Democratic and Labour party in an even more emphatic way than the DUP has challenged the UUP.*<sup>2</sup>
- *Others **eclipsed** the careers of the main actors (Dakota Fanning played a miniature Reese Witherspoon in her first movie).*<sup>3</sup>
- *Being 17 years and 119 days old, he **eclipsed** the record set by James Prinsep 125 years earlier when playing for Clapham Rovers against Old Etonians at the Kennington Oval.*<sup>4</sup>
- *Australian architects Denton Corker Marshall have almost buried the building in the ground in their anxiety not to **eclipse** the monument.*<sup>5</sup>
- *He passed it up - because he felt the new manager's hands-on approach to training **would eclipse** him.*<sup>6</sup>

Ces exemples montrent que les objets d'*eclipse* dans les corpus de presse peuvent être aussi bien animés qu'inanimés et concrets qu'abstraites.

Nous manquons toutefois de données pour le corpus des communiqués de presse, où il n'y a qu'une seule occurrence de la forme verbale d'*eclipse*, son objet est « *it* » et renvoie à une étoile (« *In 1999, this star suddenly entered the astronomical Hall of Fame when the extrasolar planet HD 209458b passed in front of it and partly **eclipsed** it.* »<sup>7</sup>).

Malgré ce manque de données, nous pouvons tout de même constater que les objets du verbe *to eclipse* dans les corpus de presse n'ont plus de lien avec le domaine spatial, ce qui peut indiquer que ce verbe est déjà bien diffusé dans la langue générale.

Nous avons ensuite relevé une forme qui n'apparaît pas dans le corpus des communiqués de presse de l'ESA. Il s'agit de la forme verbale *eclipsing*. Il semble alors qu'elle soit ou bien très rare dans le discours spécialisé, ou bien propre à la langue générale. En voici quelques exemples :

---

<sup>1</sup> Idem, 30 septembre 2008.

<sup>2</sup> Idem, 29 novembre 2003.

<sup>3</sup> Corpus aléatoire, 30 janvier 2010.

<sup>4</sup> Corpus orienté termes, 2 janvier 2009.

<sup>5</sup> Idem, 1<sup>er</sup> août 2002.

<sup>6</sup> Idem, 1<sup>er</sup> octobre 2013.

<sup>7</sup> Corpus des communiqués de presse de l'ESA, 12 mars 2003.

- *Results last week from an unprecedented poll in 46 countries by the US-based Pew Research Centre showed environmental degradation is the number one concern of people around the world, **eclipsing** worry even about nuclear attacks, ethnic rivalries or Aids.*<sup>1</sup>
- *The youngest of Mick's daughters, Georgia is already **eclipsing** big sister Lizzie as the insider's Next Top Model.*<sup>2</sup>
- *The 420-mile round trip to the game will be his longest so far, **eclipsing** the 408 miles he pedalled to Bradford Park Avenue and back for the 5-2 FA Cup win earlier this month.*<sup>3</sup>

Nous avons relevé un autre fait intéressant : la forme verbale *eclipsed* apparaît sensiblement plus souvent dans les deux corpus de presse (32,3 % dans le corpus aléatoire et 22,6 % dans le corpus orienté termes) que dans le corpus des communiqués de presse (3,7 %). Le tableau ci-dessous illustre la répartition des formes de cette unité dans les trois corpus.

ECLIPSE	Aléatoire		Orienté termes		Communiqués de presse	
	Fréquence	%	Fréquence	%	Fréquence	%
<b>Total</b>	118	100	266	100	27	100
<b>eclipse</b>	64	54,2	170	63,9	23	85,2
<b>eclipsed</b>	39	33,1	60	22,6	1	3,7
<b>eclipses</b>	8	6,8	21	7,9	3	11,1
<b>eclipsing</b>	7	5,9	15	5,6	0	0

Tableau 4.9 : Répartition des différentes formes de ECLIPSE dans les trois corpus de travail

Si nous n'avons pas remarqué d'évolution particulière de l'emploi de cette unité, nous avons tout de même remarqué un pic de son utilisation en 1999, avec 41 occurrences dans le corpus orienté termes, soit 15,4 % du nombre total d'occurrences d'*eclipse* dans ce corpus, 19 occurrences dans le corpus aléatoire, soit 16,1 % du nombre total d'occurrences d'*eclipse* dans ce corpus et 16 occurrences dans le corpus des communiqués de presse, soit 59,3 % du nombre total d'occurrences de cette unité dans ce corpus (voir tableau ci-dessous).

<sup>1</sup> Corpus orienté termes, 2 juillet 2007.

<sup>2</sup> Idem, 1<sup>er</sup> janvier 2010.

<sup>3</sup> Corpus aléatoire, 29 novembre 2003.

	Eclipse (eclipse, eclipsed, eclipses, eclipsing)		
	Aléatoire	Orienté termes	Communiqués de presse de l'ESA
1998	3,4	4,1	0
1999	16,1	15,4	59,3
2000	3,4	6	0
2001	1,7	4,1	11,1
2002	4,2	3	0
2003	5,9	4,5	3,7
2004	1,7	6,8	3,7
2005	6,8	4,5	7,4
2006	4,2	1,9	0
2007	5,9	6,4	14,8
2008	6,8	12,8	0
2009	8,5	8,7	0
2010	7,6	6,8	0
2011	7,6	4,5	0
2012	10,3	6	0
2013	5,9	4,5	0

Tableau 4.10 : Répartition en pourcents des occurrences de l'unité *eclipse* dans les seize sous-corpus de chaque corpus de travail

Nous pensons pouvoir expliquer ce pic par le fait qu'une éclipse totale du soleil a eu lieu le 11 août 1999. La presse a très largement couvert cet événement, il est donc tout à fait normal de le retrouver dans nos corpus de presse. En observant les contextes, nous avons pu constater que la plupart des occurrences d'*eclipse* concernaient directement cet événement.

#### 4.2.1.3. ORBIT

La forme verbale d'*orbit* est attestée dans les ressources spécialisées depuis au moins 1982<sup>1</sup>, dans Termium et dans le GDT. Le GDT comprend d'ailleurs trois fiches avec le verbe *to orbit* pour vedette. Le pourcentage de formes verbales par rapport aux formes nominales est de 34,1 % dans le corpus aléatoire, de 26,8 % dans le corpus orienté termes et de 26,4 % dans le corpus des communiqués de presse de l'ESA (voir tableau ci-dessous).

<sup>1</sup> Puisque l'indication de date dans les fiches du GDT et de Termium ne précise pas s'il s'agit de la date de création ou de modification de la fiche, nous ne pouvons pas savoir quand exactement une fiche a fait son apparition dans ces ressources.

ORBIT	Aléatoire	Orienté termes	Communiqués de presse
Fréquence des verbes	31	53	296
Pourcentage de verbes par rapport aux noms	34,1	26,8	26,4

Tableau 4.11 : Proportion des formes verbales de ORBIT par rapport aux formes nominales

Nous pouvons constater que le pourcentage de verbes par rapport aux noms est plus élevé de plus de 7 % dans le corpus aléatoire que dans les deux autres corpus.

Nous avons ensuite observé la transitivité du verbe et avons pu constater qu'il est intransitif et transitif aussi bien dans le corpus des communiqués de presse que dans les corpus de presse. Les emplois transitif et intransitif sont tous les deux attestés dans l'*Oxford*, le *Cambridge* et le *Collins*. En outre, ces trois dictionnaires donnent une indication du domaine spatial dans leurs définitions d'*orbit*. Nous ne sommes donc pas en mesure de déterminer si l'un des emplois est propre à la langue générale, même si nous pensons que cela est peu probable, du fait que les deux emplois apparaissent justement dans le corpus des communiqués de presse.

Nous nous sommes ensuite concentrée sur les différents objets de ce verbe et nous n'avons pas relevé la même diversification des objets que pour *launch* et *eclipse*. Dans le corpus des communiqués de presse et dans les corpus de presse, les objets sont des corps célestes :

- *Scheduled for launch in 2000 as a secondary payload on a European Ariane 5 rocket, it **will** then **orbit** the Moon, mapping the planned landing area in greater detail in preparation of the EuroMoon Lander in 2001.*<sup>1</sup>
- *Rosetta **will orbit** the comet for about a year as it heads towards the Sun, remaining in orbit for another half-year past perihelion (closest approach to the Sun).*<sup>2</sup>
- *They said the planet **is orbiting** a star called mu Arae, which is about 50 light years away in the southern constellation of Ara - the Altar.*<sup>3</sup>
- *This book was written by people who thought that the earth was flat, that the sun **orbited** the earth and that extreme weather phenomena were manifestations of the displeasure of a supreme being.*<sup>4</sup>
- *Over the four years of its mission, Cassini will show how the intricate ringlets keep changing, as they **orbit** the planet.*<sup>1</sup>

<sup>1</sup> Corpus des communiqués de presse de l'ESA, 5 mars 1998.

<sup>2</sup> Idem, 5 mars 2005.

<sup>3</sup> Corpus orienté termes, 1<sup>er</sup> septembre 2004.

<sup>4</sup> Idem, 1<sup>er</sup> juillet 2007.

- *The heat may be caused by radioactive rocks beneath the moon's surface or by the tidal forces created as the moon **orbits** Saturn although neither could generate enough heat to account for the atmosphere, Dr Johnson said.*<sup>2</sup>

Nous avons néanmoins remarqué deux contextes intéressants, où l'objet du verbe *to orbit* n'est pas lié au domaine spatial :

- *It's a little game that's been **orbiting** the Internet for a couple of weeks.*<sup>3</sup>
- *Ziggy You're stuck in a timewarp, eternally **orbiting** the summer of 1972 with David Bowie's doomed alter ego.*<sup>4</sup>

Nous avons observé les différentes formes de cette unité et avons remarqué que la forme *orbiting* avait une fréquence plus élevée dans les corpus de presse que dans le corpus des communiqués de presse. Inversement, la forme *orbit* a une fréquence plus élevée dans le corpus des communiqués de presse que dans les corpus de presse (voir tableau ci-dessous).

ORBIT	Aléatoire		Orienté termes		Communiqués de presse	
	Fréquence	%	Fréquence	%	Fréquence	%
<b>Total</b>	91	100	198	100	1123	100
<b>orbit</b>	60	65,9	138	69,7	944	84
<b>orbited</b>	3	3,3	1	0,5	3	0,3
<b>orbits</b>	8	8,8	24	12,1	83	7,4
<b>orbiting</b>	20	22	35	17,7	93	8,3

Tableau 4.12 : Répartition des différentes formes de ORBIT dans les trois corpus de travail

Nous avons remarqué le même type de phénomène pour *orbit* en 2004 que pour *eclipse* en 1999, avec 27 occurrences dans le corpus orienté termes, soit 13,6 % du nombre total d'occurrences d'*orbit* dans ce corpus, et 15 occurrences dans le corpus aléatoire, soit 16,5 % du nombre total d'occurrences de cette unité dans ce corpus.

<sup>1</sup> Corpus aléatoire, 30 juin 2004.

<sup>2</sup> Idem, 31 août 2005.

<sup>3</sup> Idem, 31 janvier 2000.

<sup>4</sup> Corpus orienté termes, 1<sup>er</sup> août 2005.

	Orbit (orbit, orbited, orbiting, orbits)		
	Corpus aléatoire	Corpus orienté termes	Corpus des communiqués de presse de l'ESA
1998	5,49	7,58	3,29
1999	2,2	6,57	4,63
2000	6,59	7,07	6,59
2001	7,69	7,07	8,46
2002	5,49	4,04	8,1
2003	2,2	3,03	<b>10,1</b>
2004	<b>16,5</b>	<b>13,6</b>	8,28
2005	8,79	6,06	8,99
2006	8,79	4,04	8,73
2007	7,69	5,56	2,94
2008	0	6,57	5,34
2009	2,2	5,05	3,29
2010	8,79	7,07	2,23
2011	6,59	5,05	3,29
2012	6,59	6,06	8,64
2013	4,4	5,56	7,12

Tableau 4.13 : Répartition en pourcents des occurrences de ORBIT dans les seize sous-corpus de chaque corpus de travail

Dans ce cas, le pic n'est pas aussi clair que dans le cas d'*eclipse* et ne se produit d'ailleurs pas en 2004 dans le corpus des communiqués de presse, mais en 2003. De plus, le pic de 2003 dans le corpus des communiqués de presse est moins net que celui de 2004 dans les corpus de presse, les différences de pourcentage avec les autres sous-corpus étant moins grandes.

Il est possible que ces deux pics ne soient pas liés directement, ils ne sont probablement pas causés par le même évènement. En effet, en 2003, l'ESA a lancé la sonde Mars Express, dans le cadre de sa première mission d'exploration d'une autre planète du système solaire, ce qui pourrait expliquer le pic dans le corpus des communiqués de presse. Par contre, en 2004, la NASA a lancé la sonde MESSENGER, qui est la première à être mise en orbite autour de Mercure, ce qui pourrait expliquer en partie le pic dans les corpus de presse. Néanmoins, il est probable que ce pic soit en lien avec le lancement de Mars Express et celui de MESSENGER, et même certainement avec d'autres évènements.

Les trois unités que nous avons observées se comportent chacune d'une façon particulière et présentent certaines spécificités, ainsi que certaines similitudes. Avant de poursuivre avec les

unités dont la forme verbale n'est pas attestée dans les ressources spécialisées, nous décidons de faire un premier point sur ce que nous avons pu observer jusque-là, notamment à propos de la complétude des trois dictionnaires généraux que nous avons consultés. S'il existe des différences non négligeables entre ce qui y est attesté et ce que nous avons observé dans les corpus de presse pour le cas de *launch*, ils nous paraissent beaucoup plus complets en ce qui concerne *eclipse* et *orbit*. Nous en déduisons deux choses. D'abord, nous pensons que les différences relevées dans le *Collins*, le *Cambridge* et l'*Oxford* quant à la transitivité du verbe *to launch*, l'utilisation de la forme passive et les prépositions qui suivent le plus souvent ce verbe lorsqu'il est intransitif peuvent s'expliquer par le fait que cette unité n'est peut-être pas encore stabilisée dans la langue générale. Ensuite, nous pensons que le fait qu'*eclipse* et *orbit* sont bien décrits dans ces trois dictionnaires peut être révélateur de leur plus large diffusion dans la langue générale, ou du moins de leur stabilité. Ils y seraient ainsi passés il y a un certain temps déjà. De plus, en observant la répartition de ces trois unités dans les seize sous-corpus des corpus de presse, nous n'avons pas relevé d'évolution particulière, mis à part les deux pics dont nous avons parlé plus haut, ce qui pourrait venir appuyer le fait que ces unités sont employées dans le discours général depuis un certain temps.

Nous pouvons néanmoins remarquer une différence intéressante. Grâce aux différences de comportement que nous avons relevées entre le discours spécialisé et le discours général, nous avons pu constater que le sens du verbe *to orbit* ne semble pas s'être étendu de la même manière que le sens des verbe *to launch* et *to eclipse*. En effet, le verbe *to orbit* est principalement utilisé dans un contexte spatial, comme nous l'avons montré dans les exemples d'emploi transitif, contrairement aux verbes *to launch* et *to eclipse*, qui sont plus souvent utilisés dans des contextes d'où le domaine spatial est, sinon absent, du moins effacé (voir points 4.2.1.1. et 4.2.1.2.).

## **4.2.2. Unités dont la forme verbale n'est pas attestée dans les ressources spécialisées**

### **4.2.2.1. PROBE**

Le pourcentage de formes verbales relevées par rapport aux formes nominales est de 43,5 % dans le corpus aléatoire, de 46 % dans le corpus orienté termes et de 12 % dans le corpus des communiqués de presse de l'ESA (voir tableau ci-dessous).

PROBE	Aléatoire	Orienté termes	Communiqués de presse
Fréquence des verbes	57	126	32
Pourcentage de verbes par rapport aux noms	43,5	46	12

Tableau 4.14 : Proportion des formes verbales de PROBE par rapport aux formes nominales

Le premier fait intéressant que nous pouvons relever est le troisième pourcentage. Nous pouvons en effet constater une différence plutôt conséquente entre ce que nous avons pu observer dans le corpus des communiqués de presse et les ressources spécialisées que nous avons consultées (le GDT et Termium). En effet, alors que le verbe *to probe* n'est pas attesté dans ces deux ressources, il apparaît tout de même 32 fois dans le corpus des communiqués de presse dans un sens spécialisé.

Nous avons ensuite observé la transitivité de ce verbe. Dans le corpus des communiqués de presse, il est employé uniquement transitivement, alors que dans les deux corpus de presse, il est employé transitivement et intransitivement :

- *I think we need to **probe** a bit further.*<sup>1</sup>
- *If he **probes** still further he will learn that the old "appearance versus reality" conundrum has possessed the most jaw-dropping, dramatic effect since before Shakespeare's day.*<sup>2</sup>
- *I suspect it will simply come to be seen as a rogue decision rather than a real basis for press freedom to **probe** into the private lives of celebrities.*<sup>3</sup>
- *It was Donald who kept the SHA on course whilst it **probed** the new Conservative government, and the 1980 legislation which threatened to tear public and private sectors apart.*<sup>4</sup>
- *For journalists trying to **probe** these murky waters, the problems are so substantial that few media organisations attempt it.*<sup>5</sup>
- *There was no attempt at all to **probe** the killer's psyche, which was a great relief.*<sup>6</sup>

Nous pouvons donc penser que l'emploi intransitif du verbe *to probe* est propre à la langue générale. Afin d'appuyer cette hypothèse, nous avons consulté le *Collins*, le *Cambridge* et l'*Oxford* ;

<sup>1</sup> Corpus aléatoire, 31 mars 2008.

<sup>2</sup> Idem, 31 janvier 2010.

<sup>3</sup> Corpus orienté termes, 2 mai 2005.

<sup>4</sup> Idem, 1<sup>er</sup> mai 2000.

<sup>5</sup> Idem, 2 février 2009.

<sup>6</sup> Idem, 1<sup>er</sup> janvier 2004.

le *Collins* est le seul à indiquer un emploi transitif uniquement. Le fait que deux des trois dictionnaires généraux consultés attestent l'emploi intransitif de ce verbe confirme notre première idée.

Nous avons ensuite observé les objets du verbe *to probe*. Voici quelques objets de ce verbe issus du corpus des communiqués de presse : *subsurface, Earth, atmosphere, Mars*. Nous les mettons en opposition avec quelques objets issus des corpus de presse : *minds, activities, ground, paradox*. En voici des exemples :

- *ExoMars will probe Mars, search for traces of past and present life, and demonstrate new technologies to pave the way for a future Mars sample-return mission in the 2020s.*<sup>1</sup>
- *Its instruments will be able to **probe** the planet from its upper atmosphere down to a few kilometres below the surface, where we hope to find critical clues concerning the conditions for life, in particular traces of water.*<sup>2</sup>
- *Recently, in a high-profile instance in 2009, investigators looking into the Galleon insider-trading scandal reportedly **probed** the activities of an unnamed ex-SAC analyst - though neither SAC or Mr Cohen was accused of any wrongdoing.*<sup>3</sup>
- *Professor Anthony Clare; Psychiatrist who **probed** the minds of public figures on the BBC Radio 4 show 'In the Psychiatrist's Chair'.*<sup>4</sup>

Nous pouvons ainsi remarquer que les exemples provenant du corpus des communiqués de presse de l'ESA sont tous concrets alors que ceux provenant des corpus de presse sont ou concrets ou abstraits. Nous pouvons donc supposer que la diversification des objets ainsi que le passage du concret à l'abstrait peuvent être de bons indices du passage du terme de la langue de spécialité à la langue générale et pourraient être le signe d'un emploi propre à la langue générale.

Nous pouvons également remarquer que des huit contextes provenant des corpus de presse, aucun n'est lié au domaine spatial. Nous pouvons y voir plusieurs explications. Soit le verbe *to probe* ne provient pas du tout du domaine spatial, ce qui peut être une possibilité si l'on considère la polysémie de cette unité, soit ce verbe est passé dans la langue générale il y a longtemps et correspond plutôt au deuxième type de changement sémantique expliqué par Meyer et

---

<sup>1</sup> Corpus des communiqués de presse de l'ESA, 10 juin 2013.

<sup>2</sup> Idem, 25 décembre 2003.

<sup>3</sup> Corpus aléatoire, 20 novembre 2012.

<sup>4</sup> Idem, 31 octobre 2007.

Mackintosh (2000 : 202) (voir point 4.2.). Nous manquons toutefois de données pour pencher vers l'une ou l'autre des explications.

Nous avons ensuite observé la répartition des différentes formes de cette unité et nous avons remarqué que *probed* et *probing* sont plus utilisés dans les corpus de presse que dans le corpus des communiqués de presse<sup>1</sup> (voir tableau ci-dessous).

PROBE	Aléatoire		Orienté termes		Communiqués de presse	
	Fréquence	%	Fréquence	%	Fréquence	%
<b>Total</b>	131	100	274	100	268	100
<b>probe</b>	84	64,1	170	62,1	176	65,7
<b>probed</b>	13	9,9	11	4	2	0,7
<b>probes</b>	10	7,7	36	13,1	81	30,2
<b>probing</b>	24	18,3	57	20,8	9	3,4

Tableau 4.15 : Répartition des différentes formes de PROBE dans les trois corpus de travail

Il est intéressant de remarquer ces différences, même si elles ne sont pas aussi flagrantes que dans le cas de *launch*, *eclipse* ou même *orbit*. Cette constatation vient tout de même appuyer notre hypothèse : le pourcentage de verbes, plus élevé dans les corpus de presse que dans le corpus des communiqués de presse, peut être la preuve que l'emploi de cette unité n'est plus exclusivement réservé au discours spécialisé et est répandu dans la langue générale ; l'emploi du verbe est d'ailleurs plus répandu dans le discours général que dans le discours spécialisé.

#### 4.2.2.2. ROCKET

Nous avons observé le pourcentage de formes verbales relevées par rapport aux formes nominales, qui est de 17,6 % dans le corpus aléatoire, de 10 % dans le corpus orienté termes et de 0 % dans le corpus des communiqués de presse de l'ESA (voir tableau ci-dessous). En effet, les 132 occurrences de *rocket* apparaissant dans ce corpus sont toutes nominales. Nous pouvons en déduire deux choses : d'abord, les ressources spécialisées sont en adéquation avec nos observations en corpus, puis, ce verbe n'est pas très répandu dans la langue générale, vu les pourcentages qui sont relativement faibles, par rapport aux quatre unités étudiées précédemment.

<sup>1</sup> Il convient néanmoins de souligner que *probing* a parfois une fonction d'adjectif.

ROCKET	Aléatoire	Orienté termes	Communiqués de presse
Fréquence des verbes	61	69	0
Pourcentage de verbes par rapport aux noms	17,6	10	0

Tableau 4.16 : Proportion des formes verbales de PROBE par rapport aux formes nominales

En ce qui concerne la transitivité du verbe dans les deux corpus de presse, nous avons pu relever presque uniquement des emplois intransitifs et seulement quelques emplois transitifs. En voici deux exemples :

- *But unless the opinion polls were to **rocket him** into overtaking Cameron by miles, the brutal electoral arithmetic remains the same as ever: voting for your first choice will often let in your worst nightmare.*<sup>1</sup>
- *JOSE MARIA OLAZABAL fired a third-round 66 to **rocket himself** into a challenging position in the Memorial Tournament here on Saturday.*<sup>2</sup>

Nous avons ensuite consulté les trois dictionnaires généraux et seul le *Cambridge* n'atteste que l'emploi intransitif. L'*Oxford* est le seul à attester cet emploi transitif. Comme nous n'en avons trouvé que deux occurrences, nous pensons que ce n'est pas un emploi encore très répandu, ce qui pourrait expliquer qu'un seul dictionnaire sur trois en fait mention.

L'*Oxford* et le *Collins* attestent un emploi transitif particulier : « *Attack with rocket-propelled missiles* »<sup>3</sup> ; « *to propel (a missile, spacecraft, etc.) by means of a rocket* »<sup>4</sup>. Même si nous pouvons rapprocher ces deux définitions, il convient de souligner une nuance de sens. Alors que la définition du *Collins* peut s'appliquer à tout objet propulsé par une fusée ou une roquette, la définition de l'*Oxford* se restreint à un contexte de conflit armé. Il est intéressant de remarquer que la définition du *Collins* inclut un sens appartenant au domaine spatial, alors que ni le GDT ni Termium n'attestent le verbe *to rocket*. Nous ne l'avons pas non plus observé dans le corpus des communiqués de presse. Il vaut la peine d'insister sur cette contradiction. Alors qu'il semble ne pas exister de verbe *to rocket* dans le discours spécialisé dans le domaine spatial ou dans les ressources spécialisées, celui-ci est défini dans une ressource générale. Il serait intéressant de vérifier si ce verbe existe dans un corpus du domaine spatial relevant d'une communication expert à expert, ainsi que dans d'autres ressources spécialisées, peut-être plus axées sur le domaine spatial.

<sup>1</sup> Corpus orienté termes, 1<sup>er</sup> mai 2010.

<sup>2</sup> Corpus orienté termes, 2 juin 2003.

<sup>3</sup> Oxford dictionaries, <http://www.oxforddictionaries.com/>, consulté le 8 avril 2014.

<sup>4</sup> Collins English Dictionary, <http://www.collinsdictionary.com/>, consulté le 18 avril 2014.

Nous avons relevé trois occurrences de ce verbe au sens attesté par l'Oxford. Les trois exemples ci-dessous relèvent d'un contexte de guerre, de terrorisme même, ce qui correspond tout à fait à la définition que donne ce dictionnaire :

- *In January a US drone aircraft **rocketed** a house in a village two miles from yesterday's target, where al-Qaida's number two, Ayman al-Zawabiri, had been expected.*<sup>1</sup>
- *Israel's Prime Minister, Ehud Barak, may have wanted his own worried citizens to believe his forces **rocketed** Mr Barghouti's headquarters as retaliation for Monday's killing of two Israelis.*<sup>2</sup>
- *This is a West Bank village regularly machine-gunned, **rocketed** and shelled by Israeli forces, who say they are retaliating for attacks on a nearby Jewish settlement.*<sup>3</sup>

Nous pouvons donc en déduire que les deux emplois transitifs attestés par les dictionnaires sont en réalité encore rares. Nous pouvons également supposer qu'ils sont en train de se diffuser dans la langue générale.

Il est intéressant de remarquer que ces trois exemples d'emploi transitif du verbe *to rocket* sont éloignés du domaine spatial et sont plutôt, en effet, à rapprocher du domaine militaire. Néanmoins, les deux exemples d'emploi transitif donnés à la page précédente, ainsi que les exemples d'emploi intransitif ci-dessous, sont plutôt à rapprocher du domaine spatial.

- *French Connection's profits have soared from pounds 6.4m to pounds 19m and its share price **has rocketed** from around 300p to 775p.*<sup>4</sup>
- *The 37-year-old, who leases a small plot of land in Samblong, central Thailand, knows the price of rice **has rocketed** - in some cases nearly doubling in three months - and that she is about to reap the benefit when she sells what her family does not eat.*<sup>5</sup>
- *But in 2004, the stock market was fine and house prices **were rocketing**.*<sup>6</sup>
- *Incorporation gave them a national voice and student numbers **rocketed** by 40% to 4.5 million while the money got ever tighter.*<sup>7</sup>

---

<sup>1</sup> Corpus aléatoire, 31 octobre 2006.

<sup>2</sup> Corpus orienté termes, 1<sup>er</sup> novembre 2000.

<sup>3</sup> Idem.

<sup>4</sup> Corpus aléatoire, 31 mars 2001.

<sup>5</sup> Corpus aléatoire, 31 mars 2008.

<sup>6</sup> Corpus orienté termes, 1<sup>er</sup> novembre 2008.

<sup>7</sup> Idem, 1<sup>er</sup> avril 2004.

Il semble que les différents sens de ces verbes dans le discours général soient influencés à la fois par le domaine spatial et par le domaine militaire. Il serait alors intéressant d’approfondir nos observations quant à cette unité, afin de mieux cerner les différents sens présents dans la langue générale.

Même si la forme *rocketing* est parfois utilisée comme un adjectif, les formes *rocketed* et *rocketing* sont caractéristiques de la langue générale, puisqu’elles n’apparaissent que dans les corpus de presse (voir tableau ci-dessous).

ROCKET	Aléatoire		Orienté termes		Communiqués de presse	
	Fréquence	%	Fréquence	%	Fréquence	%
<b>Total</b>	346	100	695	100	131	100
<b>rocket</b>	214	61,8	444	63,9	117	89,3
<b>rocketed</b>	37	10,7	31	4,5	0	0
<b>rockets</b>	83	24	204	29,3	14	10,7
<b>rocketing</b>	12	3,5	16	2,3	0	0

Tableau 4.17 : Répartition des différentes formes de ROCKET dans les trois corpus de travail

#### 4.2.2.3.STAR

Nous avons observé le pourcentage de formes verbales par rapport aux formes nominales, qui est de 18,4 % dans le corpus aléatoire, de 13,8 % dans le corpus orienté termes, et de 0 % dans le corpus des communiqués de presse de l’ESA (voir tableau ci-dessous).

STAR	Aléatoire	Orienté termes	Communiqués de presse
<b>Fréquence des verbes</b>	522	1019	0
<b>Pourcentage de verbes par rapport aux noms</b>	18,4	13,8	0

Tableau 4.18 : Proportion des formes verbales de STAR par rapport aux formes nominales

Comme pour le cas de *rocket*, les ressources spécialisées décrivent bien la réalité, puisque ni le GDT ni Termium n’attestent de verbe et que nous n’en avons pas relevé dans le corpus des communiqués de presse. Remarquons également que le pourcentage d’apparition des formes verbales dans les corpus de presse est plus bas que ce que nous aurions pu imaginer.

Nous avons ensuite relevé aussi bien des emplois intransitifs que des emplois transitifs :

- He **starred** in films, made 18 record albums, and returned to coach the national team to world championships in the 1970s.<sup>1</sup>
- She **starred** alongside Richard Gere and Edward Woodward in the film 'King David', and her television credits include sitcom 'Coupling' and 'Waking the Dead'.<sup>2</sup>
- Thirteen-year-old Irish actress Saoirse Ronan is **starring** in a run of major films: due first is *I Could Never Be Your Woman* with Michelle Pfeiffer, followed in September by *Atonement*, director Joe Wright's much-anticipated adaptation of Ian McEwan's novel. And production begins in October on another film adaptation, Peter Jackson's version of Alice Sebold's best-selling novel *The Lovely Bones*.<sup>3</sup>
- Patricia Hodge **stars** as the woman who has it all, almost.<sup>4</sup>
- The archetypal "frat pack" movie, in that it **stars** Luke Wilson, Vince Vaughn and Will Ferrell as three thirtysomething friends who set up a college fraternity, is actually one of the weaker efforts from that talented and prolific group of comic actors who appear in one another's films, lacking either the zaniness or the heart that their best works have.<sup>5</sup>
- To describe anything as beyond comparison is a bold step but Cammell and Nic Roeg's movie - which **starred** Mick Jagger and James Fox - is precisely that.<sup>6</sup>

La transitivité et l'intransitivité de ce verbe sont attestées dans les trois dictionnaires généraux que nous avons consultés. Nous pouvons donc constater que les différents emplois de *star* que nous avons relevés dans les corpus de presse sont déjà bien décrits par les ressources générales et nous pensons que cela peut être un indice de la large diffusion de ce verbe dans la langue générale.

Nous avons ensuite observé les objets de *star* et nous avons remarqué une très grande majorité d'objets désignant un être animé. Nous avons également relevé les emplois de *star* suivi soit de *as* soit de *in*, pour désigner respectivement le rôle joué par un acteur dans un film et le film en question. Nous avons ainsi pu nous faire une première idée de certains changements sémantiques survenus lors du passage de cette unité dans la langue générale.

En ce qui concerne les différentes formes de cette unité, nous avons pu faire le même constat que pour *rocket* : les formes *starred* et *starring* n'apparaissent que dans les corpus de presse (voir tableau

<sup>1</sup> Corpus aléatoire, 8 août 2009.

<sup>2</sup> Idem, 31 juillet 2007.

<sup>3</sup> Corpus orienté termes, 31 juillet 2007.

<sup>4</sup> Idem.

<sup>5</sup> Corpus aléatoire, 29 septembre 2007.

<sup>6</sup> Corpus orienté termes, 1<sup>er</sup> mai 2004.

ci-dessous), ce qui ne vient que confirmer ce que nous avons observé en tout premier lieu. Le verbe *to star* est uniquement utilisé dans la langue générale.

STAR	Aléatoire		Orienté termes		Communiqués de presse	
	Fréquence	%	Fréquence	%	Fréquence	%
<b>Total</b>	2838	100	7407	100	369	100
<b>star</b>	1611	56,8	4354	58,8	180	48,8
<b>starred</b>	138	4,9	209	2,8	0	0
<b>stars</b>	818	28,8	2290	30,9	189	51,2
<b>starring</b>	271	9,5	554	7,5	0	0

Tableau 4.19 : Répartition des différentes formes de STAR dans les trois corpus de travail

Nous avons remarqué une ressemblance tout à fait intéressante entre *star* et *rocket*. Ces unités ont très probablement pénétré la langue générale il y a longtemps, du moins avant 1998, et y sont déjà bien diffusées. Nous pensons en effet que la création d'un verbe par conversion, un verbe qui n'existe apparemment pas dans le discours spécialisé, est un indice de la large diffusion d'une unité dans la langue générale et de son adoption, en quelque sorte, par ses locuteurs. Le passage du nom au verbe pourrait donc refléter l'« étirement du sens » (Meyer & Mackintosh, 2000) de ces deux unités. Ce changement de catégorie grammaticale pourrait être, en quelque sorte, une nouvelle étape du processus de déterminologisation.

### 4.2.3. Remarques conclusives

Nous avons observé le comportement de six unités provenant du discours spécialisé dans la langue générale et avons pu faire quelques constats propres à chacune d'entre elles. Nous avons également pu relever que quelques éléments leur étaient communs.

Le premier élément est le pourcentage de formes verbales par rapport aux formes nominales, qui est toujours plus élevé dans les corpus de presse que dans le corpus des communiqués de presse. Voici un tableau récapitulatif des six unités.

Unité	Pourcentage des formes verbales par rapport aux formes nominales		
	Corpus aléatoire	Corpus orienté termes	Corpus des communiqués de presse de l'ESA
launch	77,2	68,8	26,4
eclipse	55,9	44,4	3,7
orbit	34	26,8	26,4
probe	43,5	46	11,9
rocket	17,6	9,9	0
star	18,4	13,8	0

Tableau 4.20 : Pourcentage des formes verbales par rapport aux formes nominales dans les trois corpus de travail

Nous pensons pouvoir expliquer ce phénomène de plusieurs façons : d'abord, dans le cas de *launch*, *eclipse*, *orbit* et *probe*, où le verbe est aussi utilisé dans le corpus des communiqués de presse, nous pensons que ces différences de pourcentage peuvent s'expliquer par le fait qu'en général, le substantif est réservé au discours spécialisé. Le verbe serait alors plutôt caractéristique de la langue générale (Condamines & Bourigault, 1999; Condamines & Picton, À paraître b; Condamines, 1998; Guilbert, 1973; L'Homme, 1995, 2003, 2012). Ainsi, même s'il est aujourd'hui largement reconnu que les verbes peuvent être des termes, qu'ils peuvent véhiculer des connaissances spécialisées, le substantif reste la catégorie grammaticale de choix dans le discours spécialisé. Les proportions formes verbales/formes nominales que nous avons observées dans nos trois corpus de travail viennent appuyer cette idée.

Dans le cas de *to rocket* et de *to star*, nous pensons que la création du verbe dans la langue générale – rappelons qu'il n'existe pas dans le corpus des communiqués de presse ni dans les ressources spécialisées – peut être une conséquence de la déterminologisation de ces unités. Puisque le processus de déterminologisation semble se produire sous la forme d'un flux, nous pensons qu'une première étape serait justement le passage de ces unités dans la langue générale, puis qu'une deuxième étape serait la création d'un verbe par conversion. On pourrait supposer qu'une fois le sens spatial largement diffusé dans la langue générale, qu'une fois que les locuteurs se le sont appropriés, la conversion devient alors quelque chose de naturel, comme pour toutes les autres unités de la langue générale. Le sens s'étire alors et recouvre une autre réalité (Meyer & Mackintosh, 2000).

Comme nous l'avons déjà mentionné à la fin de la section 4.2.1., nous avons constaté une certaine diversification des objets dans les corpus de presse dans le cas des verbes qui ont un

emploi transitif dans le discours spécialisé et dans le discours général, et ce, à des degrés différents<sup>1</sup>. Cette diversification des objets sous-entend une extension du sens des verbes considérés, le rapport avec le domaine spatial étant alors beaucoup moins évident, voire inexistant. Nous avons donc senti que le sens de *launch*, *eclipse*, et *probe* s'était « étiré » (Meyer & Mackintosh, 2000) dans les corpus de presse.

Le lien avec le domaine spatial quasi inexistant dans les corpus de presse, dans le cas de *rocket* et *star*, nous a permis de sentir le même « étirement du sens » (*Ibid.*) pour ces unités que pour celles mentionnées ci-dessus. Meyer et Mackintosh (2000 : 208) parlent également d'« extension figurée du sens terminologique d'origine », ce qui se rapproche de la notion de métaphore.

Il semble alors que *launch*, *eclipse*, *probe*, *rocket* et *star* soient employés dans le discours général par analogie avec leur emploi dans le discours spécialisé. Il semble aussi que le passage d'une unité du discours spécialisé au discours général puisse être suivi par la métaphorisation du sens de cette unité.

Le cas d'*orbit* pourrait être légèrement différent, car son emploi dans les corpus de presse est encore très fortement lié au domaine spatial (voir point 4.2.1.3.). L'emploi de cette unité dans le discours général par rapport au discours spécialisé ne diffère pas aussi nettement que dans le cas des cinq autres unités sur lesquelles nous nous sommes penchée, et ce, même si nous avons relevé quelques cas où son sens s'était « étiré » (Meyer & Mackintosh, 2000).

Nous avons pu remarquer certaines différences de pourcentage entre les formes verbales par rapport aux formes nominales dans le corpus aléatoire et dans le corpus orienté termes. Nous pensons pouvoir expliquer ces différences par la méthode de sélection des articles que nous avons choisie. En effet, la sélection des articles par mots clés sur la plateforme LexisNexis® aura concentré la recherche sur la forme nominale et non verbale ou adjectivale des unités<sup>2</sup>. Nous étions bien évidemment consciente de cette possibilité lors de la constitution des corpus, puisque notre but était justement d'observer ces termes-là, sous cette forme. Nous nous sommes néanmoins retrouvée confrontée à la réalité de la langue, ce qui s'est révélé être bénéfique, finalement, vu les résultats présentés dans cette section.

---

<sup>1</sup> Nous avons en effet remarqué une très grande diversification des objets des verbes *to launch* et *to eclipse*, une diversification des objets du verbe *to probe* moins extraordinaire mais tout de même très marquée et une diversification des objets du verbe *to orbit* nettement moins marquée.

<sup>2</sup> Sur la plateforme, nous avons effectué la recherche à partir de la forme canonique nominale des unités. De ce fait, seuls des noms au singulier ou au pluriel ont été considérés ; les autres formes des unités ont été exclues.

Ainsi, la proportion des formes des cent unités figurant sur la liste ne peut pas rendre compte fidèlement de la réalité dans le corpus orienté termes, vu que nous avons orienté la recherche sur une seule forme. C'est pourquoi nous pensons que la proportion des formes dans le corpus aléatoire représente le mieux la réalité de la langue, étant donné la sélection aléatoire des articles constitutifs de ce corpus.

Nous avons relevé un autre élément commun à ces six unités, qui vient renforcer le fait qu'elles ont pénétré la langue générale il y a un certain temps déjà, du moins bien avant 1998. Nous avons en effet observé la répartition de ces six unités dans les seize sous-corpus des corpus de presse, dans le but de voir si le nombre d'occurrences de ces unités subissait une nette augmentation de 1998 à 2013. En effet, si le nombre de leurs occurrences augmentait au fil des ans, avec peut-être très peu d'occurrences dans les premières années et beaucoup plus dans les dernières années, cela pourrait être le signe qu'une unité est entrée dans la langue générale plus ou moins à cette période-là. Néanmoins, nous n'avons pas observé un tel phénomène pour ces unités, nous pouvons donc supposer qu'elles font partie de la langue générale depuis plus longtemps que 1998 et, puisque nous ne possédons pas de données pour une période antérieure à 1998, nous ne sommes pas en mesure d'estimer quand ces unités ont fait leur entrée dans la langue générale.

## 5. CONCLUSION

Dans ce travail, nous avons observé le phénomène de la déterminologisation dans le domaine spatial en anglais. Nous avons observé dans quelle mesure une unité se comportait de manière différente dans le discours général par rapport au discours de spécialité. Nous avons ainsi cherché à repérer des emplois propres à la langue générale, qui nous semblent indiquer qu'une unité est largement diffusée dans la langue générale.

Nous nous sommes donc concentrée sur les différences de comportement du verbe dans le discours général par rapport au discours spécialisé, dans le cas où il existe une ambiguïté de forme entre le nom et le verbe. Nous avons examiné de manière systématique certains aspects de six unités en particulier (voir point 4.2.), ce qui nous a permis de montrer certaines similarités, notamment en ce qui concerne la fréquence du verbe par rapport au nom, qui est toujours plus élevée dans la langue générale que dans la langue de spécialité.

Nous avons ainsi pu constater que le lien avec le domaine spatial était plus ou moins effacé dans les exemples des six unités observées de plus près que nous avons donnés aux points 4.2.1. et 4.2.2., ce qui vient renforcer l'idée que le phénomène de la déterminologisation se produit sous la forme d'un flux. Meyer et Mackintosh (2000 : 202) font une distinction entre deux types de changement sémantique, soit « l'essence du sens terminologique demeure dans le mot déterminologisé », c'est-à-dire que le degré de compréhension du concept est plus faible dans le discours général que dans le discours spécialisé, soit les « changements conceptuels [sont] plus considérables », c'est-à-dire que le mot déterminologisé n'a plus de lien avec le concept d'origine dans le discours général. Ainsi, le flux que nous pensons avoir perçu grâce à nos observations pourrait représenter le continuum, en quelque sorte, qui existerait entre ces deux types de changement sémantique. En effet, les six unités que nous avons observées se trouvent chacune à une étape différente du processus de déterminologisation, le lien avec le spatial étant soit très marqué, dans le cas d'*orbit*, par exemple, soit très effacé, voire inexistant, dans le cas de *star*, par exemple.

Nous souhaitons maintenant présenter quelques perspectives d'amélioration de notre démarche et de nos observations. Nous présenterons également quelques propositions d'extension du champ de recherche.

Dans un premier temps, afin d'arriver à mieux cerner les changements sémantiques pressentis grâce à l'observation des verbes transitifs, nous pensons qu'il est nécessaire :

- d'étudier de manière plus systématique et plus approfondie la diversification des objets de ces verbes ;
- d'observer les sujets de ces verbes, afin de repérer les différences entre le discours spécialisé et le discours général ;
- de s'intéresser en particulier au passage inanimé – animé et/ou concret – abstrait des objets de ces verbes, une fois qu'ils ont pénétré la langue générale.

Nous pensons en effet que les différences entre les objets et les sujets des verbes transitifs dans le discours général par rapport au discours spécialisé du domaine spatial seraient de bons indicateurs des changements sémantiques que ces verbes auraient subis lors de leur passage dans la langue générale. Dans la même idée, nous pensons que le troisième point serait caractéristique de changements conceptuels plus profonds et peut-être indicateurs de métaphorisation du concept d'origine (Meyer & Mackintosh, 2000).

Dans le cas de *to star* et *to rocket*, qui n'apparaissent pas dans le corpus des communiqués de presse, nous pensons qu'il est également nécessaire d'observer de façon approfondie leurs objets et leurs sujets, ainsi que leurs contextes d'apparition dans les corpus de presse, dans le but de voir dans quelle mesure leur sens s'éloigne du domaine spatial et d'essayer de cerner la réalité qu'ils recouvrent dans la langue générale. Il serait également intéressant de construire un corpus orienté sur les unités faisant l'objet de nos observations, en sélectionnant des articles contenant au moins une de ces unités, par exemple, afin d'avoir un plus grand nombre d'occurrences à examiner.

Dans un deuxième temps, afin d'améliorer notre démarche, il conviendrait de prendre comme point de comparaison un corpus plus spécialisé que notre corpus des communiqués de presse. Il serait ainsi représentatif d'une communication d'expert à expert (Bowker & Pearson, 2002) et refléterait mieux l'usage réel des termes appartenant au domaine spatial. Même si les communiqués de presse nous ont permis de faire une première entrée en matière, nous pensons qu'il est nécessaire de confronter nos premières observations avec un corpus de ce type.

De plus, puisque nous n'avons considéré que des couples noms/verbes dont la forme canonique est la même dans les deux catégories grammaticales, ce qui est caractéristique de la langue anglaise, la conversion serait peut-être plus facile dans ce cas que dans le cas de couples noms/verbes dont la forme canonique n'est pas la même dans les deux catégories grammaticales, comme *constellation/constellate* ou *gravity/gravitate*. Il serait alors intéressant d'étendre nos observations à d'autres unités, dont la forme verbale et la forme nominale partagent ou non la

même forme canonique. Cela nous permettrait de voir si une même tendance se dessine dans d'autres cas que les six que nous avons étudiés dans ce travail.

Il serait également intéressant de prendre en considération d'autres journaux anglais, ou même écossais et irlandais, pour observer l'anglais du Royaume-Uni. Dans cette même idée, nous pourrions aussi considérer la presse anglophone d'autres pays européens.

Pour terminer, ce travail nous aura permis de nous familiariser avec un champ de recherche encore peu étudié et de nous confronter à la complexité du phénomène de la déterminologisation, en particulier, ainsi que des mouvements entre la langue générale et les langues de spécialité.

## 6. BIBLIOGRAPHIE

- Ahmad, K., & Rogers, M. (2001). « Corpus Linguistics and Terminology Extraction ». In S. H. Wright & G. Budin (Eds.), *Handbook of Terminology Management*, vol. 2, 725–760. Amsterdam, Philadelphie: John Benjamins.
- Assal, A. (1995). « La métaphorisation terminologique ». *L'Actualité terminologique* 28(2), 22-24.
- Atkins, S., Clear, J., & Ostler, N. (1992). « Corpus Design Criteria ». *Literary and Linguistic Computing*, 7(1), 1–16.
- Béjoint, H., & Thoiron, P. (2000). « Le sens des termes ». In H. Béjoint & P. Thoiron (Eds.), *Le sens en terminologie*, 3–19. Lyon: Presses Universitaires de Lyon.
- Bourigault, D., & Slodzian, M. (1999). « Pour une terminologie textuelle ». *Terminologies Nouvelles*, (19), 29–32.
- Bowker, L., & Pearson, J. (2002). *Working with Specialized Language, A Practical Guide to Using Corpora*. Londres, New York: Routledge.
- Cabré, M. T. (1998). *La terminologie : théorie, méthode et applications*. Ottawa: Presses de l'Université d'Ottawa et Armand Colin.
- Cabré, M. T., Estopà Bagot, R., & Vivaldi Palatresi, J. (2001). « Automatic Term Detection: a Review of Current Systems ». In D. Bourigault, C. Jaquemin, & M.-C. L'Homme (Eds.), *Recent Advances in Computational Terminology*, 53–78.
- Calberg-Challot, M. (2007). « Quand un vocabulaire de spécialité emprunte au langage courant : le nucléaire, étude de cas ». In J. Humbley (Ed.), *Aspects de la recherche en langue de spécialité*, 71–85. Paris: Université Paris VII.
- Celotti, N., & Musacchio, M. T. (2004). « Un regard diachronique en didactique des langues de spécialité ». *Ela. Études de Linguistique Appliquée*, 3(135), 263–270.
- Condamines, A. (1994). « Terminologie et représentation des connaissances ». *Didaskalia*, (6), 29–40.
- . (1995). « Terminology : New Needs, New Perspectives ». *Terminology*, 2(2), 219–238.

- . (1998). « Analyses des nominalisations dans un corpus spécialisé : comparaison avec le fonctionnement en corpus "général" ». In A. Clas, S. Mejri, & T. Baccouche (Eds.), *La mémoire des mots*, 351–368. Montréal: Aupelf.
- . (2005). « Linguistique de corpus et terminologie ». *Langages*, 1(157), 36–47.
- Condamines, A., & Bourigault, D. (1999). « Alternance nom/verbe : explorations en corpus spécialisés ». In B. Victorri & J. François (Eds.), *Sémantique du lexique verbal, Actes de l'atelier de Caen, 22-23 janvier 1999, Cahiers de l'Esrap*, pp. 41–48.
- Condamines, A., & Picton, A. (À paraître a). « Des communiqués de presse du Cnes à la presse généraliste : comment migrent les termes du spatial ». Colloque international *La néologie en langue de spécialité. Détection, implantation et circulation des nouveaux termes*, Université Lumière Lyon 2, 2-3 juillet 2012.
- . (À paraître b). « Étude du fonctionnement des nominalisations déverbales dans un contexte de déspecialisation ». Congrès Mondial de Linguistique française, Berlin, 19-23 juillet 2014.
- Cormier, M. (1988). « La terminologie : du terme au texte ». In *Terminologie diachronique. Actes du colloque organisé à Bruxelles les 25 et 26 mars 1998*, pp. 212–218.
- Drouin, P. (2003). « Term Extraction Using Non-Technical Corpora as a Point of Leverage ». *Terminology*, 9(1), 99–117.
- Dury, P. (2006). « La dimension diachronique en terminologie et en traduction spécialisée : le cas de l'écologie ». In D. Candel & F. Gaudin (Dirs.), *Aspects diachroniques du vocabulaire*, 109–124. Rouen: Publications des Universités de Rouen et du Havre.
- . (2008). « The Rise of Carbon Neutral and Compensation Carbone. A Diachronic Investigation into the Migration of Vocabulary from the Language of Ecology to Newspaper Language and Vice Versa ». *Terminology*, 14(2), 230–248.
- Estopà Bagot, R. (2001). « Les unités de signification spécialisées : élargissant l'objet du travail en terminologie ». *Terminology*, 7(2), 217–237.
- Galisson, R. (1978). *Recherches de lexicologie descriptive : la banalisation lexicale. Le Vocabulaire du football dans la presse sportive. Contribution aux recherches sur les langues techniques*. Paris: Nathan.

- . (1979). « Le phénomène de la banalisation lexicale ». In *Lexicologie et enseignement des langues*, 71–128. Paris: Hachette.
- Gouadec, D. (1990). *Terminologie : constitution des données*. Paris: AFNOR.
- Guilbert, L. (1971). « La néologie scientifique et technique ». *La Banque Des Mots*, 1, 45–54.
- . (1973). « La spécificité du terme scientifique et technique ». *Langue Française*, 17, 5–17.
- . (1975). *La créativité lexicale*. Paris: Larousse.
- Habert, B. (2000). Des corpus représentatifs : de quoi, pour quoi, comment ? In M. Bilger (Ed.), *Linguistique sur corpus. Études et réflexions* (pp. 11–58). Perpignan: Presses de l'Université de Perpignan.
- Humbley, J. (2004). « La réception de l'œuvre d'Eugen Wüster dans les pays de langue française ». *Les Cahiers Du CIEL*, 33–51.
- . (2006). « La traduction des métaphores dans les langues de spécialité : le cas des virus informatiques ». *Lynx*, (52), 49–62.
- Josselin-Leray, A. (2005). *Place et rôle des terminologies dans les dictionnaires généraux unilingues et bilingues. Étude d'un domaine de spécialité : volcanologie*. Thèse de doctorat en Lexicologie et Terminologie multilingues. Université Lumière - Lyon 2, Lyon.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. Londres, New York: Longman.
- Kocourek, R. (1991a). *La langue française de la technique et de la science : vers une linguistique de la langue savante*. Wiesbaden: O. Brandstetter Verlag.
- . (1991b). « Textes et termes ». *META*, 36(1), 71–76.
- L'Homme, M.-C. (1993). « Le verbe en terminologie : du concept au contexte ». *L'Actualité Terminologique*, 26(2), 17–19.
- . (1995). « Définition d'une méthode de recensement et de codage des verbes en langue technique : applications en traduction ». *TTR : Traduction, Terminologie, Rédaction*, 8(2), 67–88.

- . (2003). « Capturing the Lexical Structure in Special Subject Fields with Verbs and Verbal Derivatives. A Model for Specialized Lexicography ». *International Journal of Lexicography*, 16(4), 403–422.
- . (2004). *La terminologie : principes et technique*. Montréal: Les Presses de l'Université de Montréal.
- . (2012). « Le verbe terminologique : un portrait de travaux récents ». Congrès Mondial de Linguistique Française, Berlin.
- Lemay, C., L'Homme, M.-C., & Drouin, P. (2005). « Two Methods for Extracting “Specific” Single-word Terms from Specialized Corpora ». *International Journal of Corpus Linguistics*, 227–255.
- Lerat, P. (1995). *Les langues spécialisées*. Paris: Presses Universitaires de France.
- Meyer, I., & Mackintosh, K. (1996). « The Corpus from a Terminographer's Viewpoint ». *International Journal of Corpus Linguistics*, 1(2), 257–285.
- . (2000). « "L'étirement" du sens terminologique : aperçu du phénomène de la déterminologisation ». In H. Béjoint & Thoiron (Eds.), *Le sens en terminologie*, 198–217. Lyon: Presses Universitaires de Lyon.
- Nicolae, C., & Delavigne, V. (2009). « Naissance et circulation d'un terme : Une histoire d'exoplanètes ». Actes de sixièmes journées de Linguistique de Corpus, Lorient, *Textes et Corpus*, 4, 143–155.
- Pearson, J. (1998). *Terms in Context*. Amsterdam, Philadelphie: John Benjamins.
- Picton, A. (2009). *Diachronie en langue de spécialité. Définition d'une méthode linguistique outillée pour repérer l'évolution des connaissances en corpus. Un exemple appliqué au domaine spatial*. Thèse de doctorat en Sciences du Langage. Université Toulouse 2, Toulouse.
- Portelance, C. (1991). « Fondements linguistiques de la terminologie ». *META*, 36(1), 64–70.
- Rastier, F., & Valette, M. (2009). « De la polysémie à la néonymie ». *La Problématique Du Mot*, 77, 97–116.
- Rey, A. (1979). *La terminologie : noms et notions*. Paris: Presses Universitaires de France.
- Rondeau, G. (1981). *Introduction à la terminologie*. Montréal: Gaëtan Morin.

- Slodzian, M. (1995). « Comment revisiter la doctrine terminologique aujourd'hui ? » *La Banque Des Mots : Terminologie et Intelligence Artificielle*, 7, 11–18.
- . (2000). « L'émergence d'une terminologie textuelle et le retour du sens ». In H. Béjoint & P. Thoiron (Eds.), *Le sens en terminologie*, 61–83. Lyon: Presses Universitaires de Lyon.
- Toma, A. (2003). « Terminologie interdisciplinaire ». Colloque international *Traduction et francophonie - Traduire les francophonies*, Rennes, 12-13 septembre, 257-268.
- . (2011). « La terminologie multilingue - une étude autoréférentielle et multilingue ». Actes de VIIe Journée scientifique de REALTIER, Québec, *Multilinguisme et pratiques terminologiques*.
- Ungureanu, L. (2003). *L'interpénétration langue générale-langue spécialisée dans le discours d'internet*. Université Paris 13/Université Technique de Moldova.
- Valette, M., Estacio-Moreno, A., Petitjean, E., & Jacquey, E. (2006). « Éléments pour la génération de classes sémantiques à partir de définitions lexicographiques. Pour une approche sémique du sens ». In P. Mertens, C. Fairon, A. Dister, & P. Watrin (Eds.), *Verbum ex machina, Actes de la 13ème conférence sur le traitement automatique des langues naturelles (TALN 06)*, 357–366. Presses Universitaires de Louvain.
- Van Campenhoudt, M. (2006). « Que nous reste-t-il d'Eugen Wüster ? ». Colloque international *Eugen Wüster et la terminologie de l'École de Vienne*, Université de Paris 7, 3-4 février. (<http://www.termisti.refer.org/wuster.pdf>, consulté le 22 mai 2014.)

## 7. ANNEXES

### 7.1. Annexe 3.1 : extrait du corpus des communiqués de presse de l'ESA

<615>

Media event at ESOC: closest encounter between ESA's comet chaser Rosetta and Mars  
16 February 2007

On Sunday 25 February, ESA's probe Rosetta, currently on a ten-year journey to comet Churyumov-Gerasimenko, will make its closest approach to the planet Mars, coming within 250 kilometres of its surface.

The critical close swingby of Mars is needed to use the gravity of Mars to modify the spacecraft's speed and direction. Rosetta will emerge from its martian encounter pointed towards its next target, Earth ! It arrives for a second swingby of our home planet on 13 November (the first having already taken place on 4 March 2005).

To take advantage of this upcoming closest of encounters with the Red Planet, Rosetta's instruments - as well as those on its lander - will be switched on over predefined time slots to perform a series of scientific observations, including planetary imaging. Flight controllers at ESA's European Space Operations Centre (ESOC) have already set everything ready for this crucial manoeuvre.

Launched on 2 March 2004 on an Ariane 5 rocket, Rosetta is the first probe ever designed to enter orbit around a comet's nucleus and release a lander onto its surface. Arriving at comet Churyumov-Gerasimenko in 2014, the probe will take over a year to conduct a thorough scientific study of this remnant of the primitive nebula which gave birth to our solar system some 4.6 billion years ago. By the end of its epic journey, Rosetta will have performed three Earth and one Mars swingbys in all. It will also have studied asteroids Steins and Lutetia, in September 2008 and July 2010 respectively.

Media representatives wishing to follow this Rosetta Mars swingby from the ESOC control centre in Darmstadt/Germany are requested to complete and return the attached reply form.

For further information, please contact :

ESA Communication Department

Media Relations Office

Tel: +33(0)1.53.69.7155

Fax: +33(0)1.53.69.7690

Programme

Rosetta Mars swingby

25 February 2007, 2 a.m. start

02:00 - Doors open & Filming opportunity in Mission Control Room

02:40 - Welcome by David Southwood, ESA Director of Science Programme

02:50 - Rosetta Mars swingby the manoeuvres and flight dynamics, Uwe Feucht, Head of Flight Dynamics Division/Team

03:00 - En route science, first images, Gerhard Schwehm, Head of Solar System Science Operations Division & Rosetta Mission Manager

03:10 - Introductory comments on approach ; Paolo Ferri, Head of Solar and Planetary Missions Division and Rosetta Flight Operations Director Comments on eclipse, Andrea Accomazzo, Rosetta Spacecraft Operations Manager Comments on closest approach/eclipse, Andrea Accomazzo & Paolo Ferri

03:15 - Closest approach to Mars, eclipse starts

03:17 - Ground stations, Manfred Lugert, Head of Ground Facilities Operations Division

03:28 - Occultation ends - signal back

03:30 - Imagery from Rosetta and Mars Express , Uwe Keller, Mas-Planck Institute

03:40 - Comments on eclipse end and telemetry acquisition, Andrea Accomazzo,

03:52 - Conclusions, Manfred Warhaut, Head of Mission Operations Department

04 :00 - End of event

ROSETTA MARS SWINGBY

ESA/ESOC – Robert-Bosch Strasse, 5 – Darmstadt (Germany)

Sunday 25 February 2007 – 2.a.m. start

Name : \_\_\_\_\_ : Surname: \_\_\_\_\_

Media : \_\_\_\_\_

Address : \_\_\_\_\_

Tel.: \_\_\_\_\_ Fax: \_\_\_\_\_

Mobile : \_\_\_\_\_ E-mail: \_\_\_\_\_

I will participate

I will not participate

Please return this form back to :

Jocelyne Landeau-Constantin

ESA/ESOC Communications Department

Tel: +49(0)6151.90.2696

Fax: +49(0)6151.90.2961

For further information:

ESA Media Relations Office

Tel: +33(0)1.53.69.7155

Fax: +33(0)1.53.69.7690

<616>

Rosetta comet-chaser takes a close look at planet Mars

25 February 2007

There was considerable relief today at ESA's space operations centre (ESOC) in Darmstadt, Germany. In the early hours, spacecraft controllers, orbital mechanics experts, engineers and scientists were able to witness a spacecraft playing 'cosmic billiards'. Between 03:13 and 03:40 CET, ESA's comet chaser, Rosetta, swung by Mars at a distance of only 250 kilometres and then whooshed away from the Red Planet, on a brand new path, continuing on a journey that will ultimately take it beyond Jupiter's orbit.

Its final destination is comet Churyumov-Gerasimenko, which it will reach only in 2014, after travelling some 6000 million kilometres in 10 years (its epic voyage began on 2 March 2004 with a launch by an Ariane 5 rocket). Rosetta will next be heading for the Sun, and its journey will require two more swing-bys around the Earth, in November this year and November 2009.

Once at its destination, Rosetta will first deposit, from a height of about one kilometre, a small but very complex lander on the comet's nucleus. This lander, a sort of miniature chemical laboratory packed with sophisticated instruments, will analyse the surface and provide information on the nucleus. The Rosetta probe will then chase the comet for one year and observe its nucleus as it continues on its trip towards the inner solar system at a speed of 135,000 km per hour.

There is still a long way to go, but so far everything seems to be going exactly according to plan. ESA's Director of Science, David Southwood, witnessing the Mars swing-by at ESOC with scientists involved in the mission and the operations teams, said: "Interplanetary expeditions rely on very complex communication links. ESA's mission operations centre here in Darmstadt is

doing a great job. I and all the scientists involved in the mission are grateful to the experts who are taking such good care of 'our baby'. And this is only the beginning. The true excitement of targeting and releasing the lander on the comet's nucleus is yet to come. Today we have reached another milestone on the way to finding an answer to questions such as whether life on Earth began with the help of comets."

"The successful Mars swingby of the ESA Rosetta spacecraft has been the most critical event in the mission since launch. Now we are heading back to Earth in order to gain, in November this year, further momentum for the subsequent visits of the asteroids and the comet. I would like to thank all those who have contributed to this achievement", said Manfred Warhaut, Head of the Mission Operations Department.

During the approach to Mars, instruments onboard Rosetta - as well as on its lander - were switched on at predefined times to observe the environment and take imagery of the Red Planet. In September 2008 and July 2010, when it is deep inside the asteroid belt between Mars and Jupiter, Rosetta will also observe the asteroids Stein and Lutetia close up.

For further information, please contact:

ESA Media Relations Office  
Communication Department

Phone: + 33 1 5369 7155

Queries: [media@esa.int](mailto:media@esa.int)

ESA/ESOC

Communication Office

Phone: + 49 6151 90 26 96

For further information:

ESA Media Relations Office

Tel: +33(0)1.53.69.7155

Fax: +33(0)1.53.69.7690

<617>

Official opening of the Soyuz launch base construction site in French Guiana

26 February 2007

The construction site of the Soyuz launch base in French Guiana was officially opened today by Jean-Jacques Dordain, ESA Director General, Yannick d'Escatha, President of CNES, Jean-Yves Le Gall, Director General of Arianespace, and Anatoly Perminov, Head of Roscosmos. The ceremony took place in the presence of many French authorities and representatives of all the European and Russian entities contributing to the startup of the project.

On this occasion, a commemorative plaque was unveiled and a stone from the Baikonur launch pad from which the cosmonaut Yuri Gagarin took off in 1961 was deposited on the site. This stone is a powerful symbol of the continuity between the Russian launch site, which saw the start of the space age with Sputnik followed by the first human space flight, and the setting-up of a Soyuz launch operation in French Guiana. The arrival of Soyuz marks the culmination of 40 years of space cooperation between France and Russia, a decade of industrial cooperation in the Starsem framework, and the longstanding cooperation between ESA and Russia's Federal Space Agency (Roscosmos).

Soyuz, renowned for its legendary exploits in space exploration, has been launched 1713 times so far, thus demonstrating both its robustness and reliability. It has to date placed 1661 satellites in orbit around the Earth and sent 91 Russian and 40 non-Russian cosmonauts into space.

Launching Soyuz from the Guiana Space Centre (CSG) will make it possible to use French Guiana's equatorial position to significantly increase the launcher's lift capability. After an inaugural flight at the end of 2008, Soyuz will launch communication satellites into geostationary orbit, navigation satellites forming part of the European Galileo constellation, Earth observation

satellites into polar orbit, and interplanetary probes. Soyuz at the CSG, alongside the Ariane 5 heavy-lift launcher and Europe's small Vega launcher, will harmoniously complete the range of ESA launchers. Their exploitation is handled by Arianespace, which will be able to offer launch services for all types of mission.

This project is being co-funded by ESA, the European Union and Arianespace and carried out with CNES as system prime contractor, in the framework of an ESA programme. Although the site is being officially opened today, excavation work was begun several months ago by the French firm VINCI with the help of numerous European and local firms; the first Russian teams are due to arrive at Sinnamary between now and the end of the year to build the launch and functional support infrastructure for the launch pad.

As Jean-Jacques Dordain, Director General of the European Space Agency, said: "We are entering a new era for launchers for Europe, which is the positive outcome of good cooperation between ESA and Russia, initiated by France, and which will enhance the launch flexibility offered by Arianespace."

Yannik d'Escatha, President of CNES, observed that: "The continuous relationship developed in the space domain between France and Russia over the last 40 years and its extension to cooperation on launchers as from 1996 has led today to this tremendous joint project between Europe and Russia: Soyuz at the CSG. With the construction work on this launch site at the Guiana Space Centre, Europe's Spaceport, the link that has been developed will now be very much a tangible one on French territory, for ESA, for decades to come."

Jean-Yves Le Gall, Chief Executive Officer of Arianespace, said that: "Soyuz at the CSG is a new phase in the amazing undertaking begun 10 years ago by Europe and Russia within Starsem. Soyuz at the CSG will be a great asset for Arianespace's commercial strategy and will greatly benefit our customers, shareholders and partners."

As Anatoly Perminov, Head of Roscosmos, put it: "We can clearly see today that the most ambitious projects, such as this construction of the Soyuz launch pad at the CSG, cannot be carried out by a single entity. It is only through the joint efforts of Europe, Russia and France that such a bold ambition is becoming a tangible reality before our very eyes. This mutually beneficial cooperation provides a guarantee that the future exploration of space can only be for peaceful purposes and that it will always be in line with the aspirations of humanity at large."

For further information, please contact:

Media Relations Office

ESA Communication Department

Tel: +33(0)1.53.69.7155

Fax: +33(0)1.53.69.7690

For further information:

ESA Media Relations Office

Tel: +33(0)1.53.69.7155

Fax: +33(0)1.53.69.7690

## **7.2. Annexe 3.2 : extrait du corpus orienté termes**

<TEXT=16814 - The Guardian>

<DATE=02.01.2012>

International: Iran announces successful test of smart missile

Iran has claimed to have successfully tested a new medium-range ground-to-air missile during naval exercises in the Gulf, amid rising tensions over the country's nuclear programme.

State news agency Irna yesterday quoted Iran's naval commander, Mahmoud Mousavi, as saying the missile was equipped with the latest technology. The missile test was made during 10 days of

naval exercises to the east of the Straits of Hormuz, the narrowest section of the Gulf, which Iran has threatened to close if the west interferes with its oil exports.

The exercises come a few weeks before EU foreign ministers meet to consider further sanctions, possibly including an oil embargo, after an International Atomic Energy Agency (IAEA) report in November confirmed western allegations that Iran had worked on nuclear weapons designs, at least until 2003, and may have carried out experiments more recently.

Amid constant speculation that Israel or the US could use air strikes against Iran's nuclear programme, Iran has tried to buy long-range surface-to-air missiles from Russia. After Moscow cancelled a deal to sell missiles last year, Tehran said it would develop its own as an independent deterrent. Yesterday's statement appeared designed to show it was making progress.

Iran insists its nuclear programme is entirely for peaceful purposes and on Saturday announced "a breakthrough" in its independent nuclear programme - the production of its first nuclear fuel rod.

<TEXT=16815 - The Guardian>

<DATE=02.01.2012>

International: China: This is no video game: China's historic treasures plundered by tomb raiders: Thieves use dynamite and bulldozers to steal relics to meet global demand

China's extraordinary historical treasures are under threat from increasingly aggressive and sophisticated tomb raiders, who destroy archaeological evidence as they swipe irreplaceable relics. The thieves use dynamite and even bulldozers to break into the deepest chambers - and night vision goggles and oxygen canisters to search them. The artefacts they take are often sold on within days to international dealers.

Police have already stepped up their campaign against the criminals and the government is devoting extra resources to protecting sites and tracing offenders. This year it set up a national information centre to tackle such crimes.

Tomb theft is a global problem that has gone on for centuries. But the sheer scope of China's heritage - with thousands of sites, many of them in remote locations - poses a particular challenge.

"Before, China had a large number of valuable ancient tombs and although it was really depressing to see a tomb raided, it was still possible to run into a similar one in the future," said Professor Wei Zheng, an archaeologist at Peking University. "Nowadays too many have been destroyed. Once one is raided, it is really difficult to find a similar one."

His colleague, Professor Lei Xingshan, said: "We used to say nine out of 10 tombs were empty because of tomb-raiding, but now it has become 9.5 out of 10."

Their team found more than 900 tombs in one part of Shanxi and almost every one had been raided.

They spent two years excavating two high grade tombs from the Western Zhou and Eastern Zhou periods (jointly spanning 1100BC to 221BC) and found both had been completely emptied by thieves. "It really is devastating to see it happening," Zheng said. "Archaeologists are now simply chasing after tomb raiders."

Experts say the problem became worse as China's economy opened up, with domestic and international collectors creating a huge market. Zheng said a phrase emerged in the 1980s: "If you want to be rich, dig up old tombs and become a millionaire overnight."

But he added that a crackdown by authorities was helping to contain the problem to an extent. According to the ministry of public security, police investigated 451 tomb-raiding cases in 2010 and another 387 involving the theft of relics. In the first six months of that year, they smashed 71 gangs, detained 787 suspects and recovered 2,366 artefacts.

Those caught face fines and jail terms of three to 10 years, or life in the most serious cases.

Officials say tomb thefts have become increasingly professionalised. Gangs from the provinces worst hit - Shanxi, Shaanxi and Henan, which all have a particularly rich archaeological heritage -

have begun exporting their expertise to other regions. One researcher estimated that 100,000 people were involved nationally.

Wei Yongshun, a senior investigator, told China Daily in 2011 that crime bosses often hired experienced teams of tomb thieves and sold the plunder on to middlemen as quickly as they could.

Other officers told how thieves paid farmers to show them the tombs and help them hide from police.

Local officials have insufficient resources to prevent the crimes and often do not see the thefts as a priority. Others turn a blind eye after being bribed by gangs.

Often, raiders return to a site repeatedly over months. In some cases, thieves have reportedly built small "factories" next to tombs - allowing them to break in without being noticed.

But international collectors bear as much responsibility for the crimes as the actual thieves: the high prices they offer create the incentive for criminals.

Wei said: "Stolen cultural artefacts are usually first smuggled out through Hong Kong and Macao and then taken to Taiwan, Canada, America or European countries to be traded."

The sheer size as well as value of the relics demonstrates the audacity of the raiders - last year, the Chinese authorities recovered a 27-tonne sarcophagus that had been stolen from Xi'an and shipped to the US.

It took four years of searching before China identified the collector who had bought the piece - from the tomb of Tang dynasty concubine Wu Huifei - for an estimated \$1m (£650,000), and secured its return.

Luo Xizhe of the Shaanxi provincial cultural relics bureau told China Daily: "If we don't take immediate and effective steps to protect these artefacts, there will be none of these things left to protect in 10 years."

He said provincial and national authorities planned to spend more than 100m yuan (£10m) on surveillance equipment for tombs in Shaanxi over the next five years. But video surveillance and infrared imaging devices for night-time monitoring cost 5m yuan for even a small grave, he added.

Spending on protecting cultural relics as a whole soared from 765m yuan in 2006 to 9.7bn in 2011.

Wei, the archaeologist, said precious evidence such as how and when the tomb was built was often destroyed in raids, even if relics could be recovered. "Quite apart from the valuable objects lost, the site is also damaged and its academic value is diminished," he said.

In a particularly alarming case last year, raiders simply bulldozed their way through 10 newly discovered tombs in eastern Jiangxi province.

The Global Times newspaper reported that pieces of coffins and pottery and iron items were scattered across the ravaged site, which was thought to date back 2,000 years. Archaeologists said further excavation was impossible because the destruction was so bad.

Additional research by Han Cheng

Captions:

Terracotta warriors were first found in a Shaanxi tomb in 1974, but new finds are threatened by thieves, with one expert suggesting that up to 95% of tombs have been raided

Photograph: Jason Lee/Reuters

<TEXT=168 - The Guardian>

<DATE=02.01.2012>

National: Calls to pardon Turing grow as he is given stamp honour

The mathematician and second world war codebreaker Alan Turing is to be celebrated on a special stamp as an online petition calls for a posthumous pardon to quash his conviction for gross indecency.

The computer pioneer is one of 10 prominent people chosen for the Royal Mail's Britons of Distinction stamps, to be launched in February, which includes the allied war heroine Odette Hallowes of the Special Operations Executive, composer Frederick Delius, and architect Sir Basil Spence, to mark the golden jubilee of Coventry Cathedral.

Turing worked as part of the team that cracked the Enigma code at Bletchley Park, and went on to help create the world's first modern computer. This year marks the centenary of his birth.

He was convicted of gross indecency in 1952, when homosexual acts were illegal in the UK, and sentenced to chemical castration. He killed himself two years later by taking cyanide. The e-petition says his treatment and death "remains a shame on the UK government and UK history".

In 2009, the then prime minister Gordon Brown issued an unequivocal apology on behalf of the government to Turing, describing his treatment as "horrifying" and "utterly unfair". Brown said the country owed him a huge debt.

Hallowes, who was born in Amiens, France, in April 1912 and married an Englishman in 1931, was awarded the George Cross (the only woman to receive the honour while alive) and the Legion d'honneur for her work in Nazi occupied France. She was imprisoned, tortured and condemned to death in 1943 after being betrayed, but survived Ravensbruck concentration camp and the war. She died in 1995 in Walton-on-Thames, Surrey.

The Olympic Games and the Queen's diamond jubilee also feature in the 2012 special stamp programme.

House of Windsor stamps will feature the five monarchs from the start of the 20th century, with the Queen taking pride of place on the final stamp to mark the 60th year of her reign. A special miniature sheet issued on 6 February will bring together six portraits of her taken from stamps, coinage and banknotes.

Giant peaches and a famous chocolate factory will feature in a set of six stamps on 10 January to celebrate the work of Roald Dahl.

The first of several stamp issues to mark the London Games go on sale on 5 January. Other stamp subjects during the year include Charles Dickens, born 200 years ago this year, great British fashion, comics, dinosaurs, space science and classic locomotives of Scotland.

Captions:

Alan Turing helped crack German codes during the second world war

### **7.3. Annexe 3.2 : extrait du corpus aléatoire**

<TEXT=3768 - The Guardian>

<DATE=30.04.2001>

Comment & Analysis: Gerry's smash and grab: The replacement of regional arts boards with a national organisation is compellingly awful viewing

Here's a strange little drama of how New Labour modernises government - and it's showing near you. Audiences so far have been modest, because the script is somewhat confusing, but it's worth persevering with: it's a classic farce rich in irony and paradox and by the end of the it you'll be rolling in the aisles.

The central character is complex: a charming, very successful businessman, with a ruthless record as corporate raider and a rags-to-riches background. The plot may seem a little obscure since it involves a small corner of government. But add in millions of unexpected extra dosh released by the Treasury, along with the corporate decision to take the public service ethos by the scruff of its neck, and Gerry's Smash and Grab is a tale for our times.

It opened in mid-March when Gerry Robinson (chairman of the Arts Council) launched a dawn raid on the England's 10 regional arts boards (RABs). He was freshly glowing from having wowed Gordon Brown (who plays a cameo role) with his Granada-Compass history of efficiency and cost-cutting contract catering to wrestle the biggest budget out of the Treasury in memory. A fat pounds 100m was swilling about for the arts, and with the plaudits still ringing in his ears, Robinson pounced.

The action is pacy and full of surprises. Plans were drawn up in secret and sprung on astonished RABs, which were given six weeks to hand over all their staff contracts and assets to Gerry. The six weeks are up today and no one has handed over anything. Now there's a twist in the plot Gerry wasn't expecting: a debate in the Commons on Wednesday amid general uproar.

The script has been a virtually incomprehensible document, the Prospectus for Change, which is liberally sprinkled with keywords such as "modern" and "new". This is supposed to explain to the RABs' several hundred staff and thousands of artists whom they help what the purpose of the upheaval is all about. The idea is one national arts organisation with regional offices. A sweetener was added for the RABs : we may be abolishing you, but you'll get bigger budgets.

Confused by now as to what Gerry's up to? Well, you're in good company. Eight out of the 10 RABs are in a state of outraged rebellion. Instead of arts administrators getting on with their jobs, they're all in meetings trying to puzzle out what the prospectus means and what they should do about it. Arts Council staff are zipping about the country trying to explain the inexplicable prospectus to angry meetings of the arts community and the public. Gerry's apparatchiks are trying to pour oil on very troubled waters and, six weeks too late, they are lavishing soothing words about consultation and devolution on the ruffled regional sensitivities.

In the midst of this muddle there are three things on which the prospectus is clear. First, while it primly insists that this is not primarily a cost-cutting exercise, it admits that saving money will be one of the results. Second, it is honest about jobs: yes, there will be some job losses; yes, some could be compulsory redundancies; and yes, all redundancy agreements will have to be renegotiated.

Third and most importantly, whatever else this prospectus hopes to achieve, one thing is clear: it talks of a national "single voice" for the arts. It's the idea of Robinson with a loud-hailer that has really freaked out the RABs, who pride themselves on having built up distinctive and accountable regionalism. Local authorities are incensed at the high-handedness with which Robinson's plans sweep them aside.

This is the most contentious issue. Even Chris Smith at the Department of Culture, Media and Sport has made his nervousness clear in a conveniently leaked letter. The suspicion is that the Labour luvvies have nobbled Robinson, complaining of bureaucracy, and, as one northerner admitted, the Antony Gormleys and Anish Kapoor of this world would much prefer to deal straight with the bigwigs in London. Where the RABs play a crucial role is in building up and developing the work of local artists and communities who are unlikely to get to bend Robinson's ear.

It's the same hook New Labour gets caught on every time. All its rhetoric is to devolve, but its instincts are to centralise. After the Livingston/Morgan debacles in London and Wales, the talk is of having learnt lessons, but the Arts Council is flying in the face of this, defending its plans for this "single voice" on the grounds that artists should be able to expect consistency in arts policy across the country. But is consistency such a prize that it's worth sacrificing local accountability and regional devolvement for it? The bitter irony is that New Labour is finishing off the work begun by the Tories in their mid-80s centralising spree when the power to appoint the RABs' chairs was shifted from the regions to the secretary of state.

Dawn raids may be good drama but they don't work in the public sector. The first sentence of the prospectus says that all the restructuring would be carried out swiftly while "consultation and implementation" would come in the second stage. Nothing now is going to happen swiftly, the

sides are digging in for a long trench war. Meanwhile the trust between London and the regions (never high) has been spectacularly spent. This could end up a tragedy.  
madeleine.bunting@guardian.co.uk

<TEXT=3769 - The Guardian>

<DATE=30.04.2001>

Leading article: Mr Bush's bad dream: Conscience-less conservatism runs amok

That Field of Dreams is George Bush's favourite film provides a clue to his presidency, now 100 days old. Set in mid-west Iowa farm country, the movie is about baseball, Mr Bush's abiding passion. And it is about fantasy as, one by one, the all-time greats of the game's bygone eras emerge from the towering maize, bat and gloves in hand, to live and play again. Picking up the theme, Mr Bush's main concern since taking office has been to put the clock back. His has been shown to be a peculiarly narrow, inflexible interpretation of America's conservative tradition. He appears overly influenced, if not led by the nose, by hawkish old-timers such as vice-president Dick Cheney and defence secretary Donald Rumsfeld. But more than this, Mr Bush seems set in a time-warp, to be harking back to an America that he thinks perhaps existed in his 1950s childhood but has long since disappeared back into the corn.

Mr Bush's inherited political fantasy is of an America that is homogeneous at home and dominant and unchallengeable abroad. It is of an American society that is God-fearing, uniform and structured, that upholds the old values because they are simply the best, that leads the world because the world simply knows it is right. That Mr Bush is not a deep thinker has become apparent to all. Nor is that so terrible. But what is terribly troubling is that this president does not seem to question or delve beneath the superficial images that flicker across his screen, be they of China (atheistic communists), Iraq (evil dictatorship), or environmentalists (anti-business liberals). He is beguiled by fatuous ideas of the make-believe past, by its false certainties and old totems. Mr Bush pays lip service to multiculturalism and other aspects of modern, changing America. But that is all. He has a sure grasp on what never was but little purchase on what may be in future. His father called it the "vision thing". He lacked it - and so, too, does his cliché-confounded son. Mr Bush portrays himself as a moral man. He is a regular churchgoer, a fitness fanatic who sees in a healthy body the foundation of a healthy mind. Having vowed to give the Oval Office a "good scrubbing" after Bill Clinton's depredations, he has banned informal dress in the White House and ordered that "naughty bits" (sex and violence) be excised from films shown aboard Air Force One. Mr Bush has invited faith-based groups to play a role in government and, for example, banned overseas aid for organisations offering abortion services. His increased education funding is linked to his concern for traditional family values. He repeatedly justifies his decisions as "the right thing to do".

Were he a more insightful man, Mr Bush might glimpse the contradictions between reality and this risible self-righteousness - and the outrage that results. What is remotely moral about the systematic rollback of industrial, business, environment and workplace regulations and safeguards described last week by Julian Borger in G2? How does Mr Bush square his sworn duty to lead the nation as a whole with his regressive tax cuts and his abject surrender to special interests such as Big Oil, Big Tobacco and what used to be called the military-industrial complex? There can be very little moral justification for building new generations of treaty-busting missiles, risking a revived cold war with Russia, unnecessarily antagonising China over Taiwan, pouring scorn on the Korean peace process, and giving Israel's Ariel Sharon a virtual free run. When he blithely tears up the Kyoto global warming protocol and instructs his treasury secretary to block IMF-World Bank bailouts for developing world economies, little wonder many ask what ethical universe Mr Bush inhabits. Born again Christian or not, these and other Bush actions at home and abroad in the first 100 days have been deeply immoral. They proceed not from faith and principle but from arrogance and ignorance.

This is fantasy politics at its most destructive. And yet the biggest contradiction of all is that this so-say moral man did not rightly or fairly win last November's election - and but for the Florida fix, would now be back in Texas watching ballgames. Living with Mr Bush in future will require a clear understanding of his limitations but also of these basic self-deceptions. Let us meanwhile bury that "compassionate" tag once and for all. This is conscience-less conservatism. And it is dangerous to your health.

<TEXT=3770 - The Guardian>  
<DATE=30.04.2001>

Leading article: The wrong man resigns: Ukraine's struggle is just beginning

Viktor Yushchenko's forced resignation as Ukraine's prime minister has intensified the sense of political crisis that has gripped Kiev since last autumn when President Leonid Kuchma was accused of ordering the murder of an investigative journalist. Mr Yushchenko's tearful departure after losing a parliamentary confidence vote sparked renewed street protests of a kind that in recent months have frequently threatened to spiral out of control. But his overthrow has had a broader impact, affecting confidence in Ukraine's still immature democracy and in the country's ability to tackle its economic problems. The vanquished prime minister's market reforms and pro-western stance had begun to show results, with GDP growing by 6% last year, industrial production up by 12% and steel exports to the US and the EU on the rise. The figures reflected the first yearly growth recorded since Ukraine became independent 10 years ago, a period that saw real GDP plummet by about two-thirds. Now big lenders such as the US and the World Bank-IMF are likely to hold back until the dust settles.

That may take a while. Mr Yushchenko alienated powerful business interests - the Ukrainian equivalent of Russia's oligarchs - who feared his reforms would loosen their often corrupt hold on Soviet-era state monopolies. He was badly served by his divided allies and fiercely opposed by old-style communists. In courting the west, he increasingly represented an obstacle to President Vladimir Putin's evident ambition to re-establish Moscow's sway in the most important part of what Russians proprietorially call their "near abroad". And Ukraine is in any case dependent for its energy supplies on Russia, still its biggest trading partner. Mr Yushchenko was further undermined by Mr Kuchma, a weakened and oppressive figure who has good cause to fear his prime minister's grassroots popularity. In truth, it is the president who should have resigned.

These underlying tensions and divisions will not be dispelled by this latest upheaval. On the contrary, they may only grow worse, ultimately undermining Ukraine's briefly-held independence (as in the case of Belarus). "I am not abandoning politics. I am going away in order to return," Mr Yushchenko said. He is now expected to try to rally the opposition ahead of elections next year. That offers a glimmer of hope. But it also suggests the crisis will deepen.

<TEXT=3771 - The Guardian>  
<DATE=30.04.2001>

Leading article: Visa for the final frontier: The sky's not the limit - if you're rich

That Dennis Tito is the first space tourist is an earthly assumption of galactic proportions. As X-Files fans, crop circle cognoscenti, and Arizona flying saucer specialists all know, interplanetary visitations are by no means unusual. All the same, all credit to Mr Tito, who was orbiting to his heart's desire yesterday, camera and spare rolls of film to hand, while awaiting docking with the international space station. And credit, too, in the bank account of the cash-strapped Russians who blasted the US-born millionaire off the face of the earth for a cool Dollars 20m and a promise to make good any breakages. Spoilsports at Nasa tried to ground Mr Tito because they said he might sue if zapped by Martians, was otherwise obliterated, or lost his return ticket. Fortunately, this grubby attempt to apply US legal practice to outer space failed to take off.

But Mr Tito's exploit raises a serious issue: the fundamental question of extraterrestrial egalitarianism. To be blunt, why is it only rich people who get to fly about the universe? In pursuit of equal opportunities on earth and in heaven, we volunteer to be the first newspaper in space. Can't afford Dollars 20m, though.

<TEXT=3772 - The Guardian>  
<DATE=30.04.2001>

Letter: Against global capitalism

We, activists who have worked within the framework of the "official" trade union and labour movement for many years, declare solidarity with the "globalise resistance" demonstration outside the World Bank offices and other "May Day Monopoly" events. We support the demonstrators against the fantastic frenzy of aggression being whipped up by the police, much of the media and the mayor of London, Ken Livingstone. The announcement of a "zero tolerance" policing policy and a campaign of preventive arrests is a brazen abuse of civil liberties.

Livingstone's call for police to arrest anyone "whose intention it is to engage in criminal activities" (how will they be selected?) is shameful. It is an incitement to officers who have no sympathy for progressive causes to abuse the powers given them by Labour's dictatorial Criminal Justice Act - as Livingstone knows perfectly well. The publication of photo-graphs of people "suspected" by police of "intending" to cause violence is witch-hunting, not journalism.

The establishment's paranoid reaction to the movement against global capitalism, and the indifference or hostility of most "official" labour organisations, is symptomatic. Instead of embracing the passionate enthusiasm with which a new generation are confronting the evils of an immoral society, they shudder in fear. We welcome that enthusiasm with open arms.

Richard Abernethy, Clare Cowen, Bridget Leach, Simon Pirani and 21 others  
smpirani@compuserve.com

## 8. INDEX DES TABLEAUX

Tableau 3.1 : Nombre d'occurrences dans les sous-corpus du corpus des communiqués de presse de l'ESA.....	24
Tableau 3.2 : Nombre d'occurrences dans les sous-corpus du corpus de presse généraliste orienté termes.....	34
Tableau 3.3 : Nombre d'occurrences dans les sous-corpus du corpus de presse généraliste aléatoire.....	34
Tableau 4.1 : Les dix unités les plus fréquentes dans les deux corpus de presse .....	36
Tableau 4.2 : Unités dont la fréquence est plus élevée dans le corpus aléatoire que dans le corpus orienté termes .....	38
Tableau 4.3 : Unités dont la différence de fréquence est de plus de 0,4 % .....	39
Tableau 4.4 : Exemples d'unités dont la fréquence par rapport aux cent termes est comparable dans les deux corpus de presse .....	39
Tableau 4.5 : Proportion des formes verbales de LAUNCH par rapport aux formes nominales .	43
Tableau 4.6 : Répartition de la forme passive <i>launched</i> dans le corpus des communiqués de presse de l'ESA.....	43
Tableau 4.7 : Répartition des différentes formes de LAUNCH dans les trois corpus de travail....	46
Tableau 4.8 : Proportion des formes verbales de ECLIPSE par rapport aux formes nominales ..	46
Tableau 4.9 : Répartition des différentes formes de ECLIPSE dans les trois corpus de travail.....	48
Tableau 4.10 : Répartition en pourcents des occurrences de l'unité <i>eclipse</i> dans les seize sous-corpus de chaque corpus de travail .....	49
Tableau 4.11 : Proportion des formes verbales de ORBIT par rapport aux formes nominales ....	50
Tableau 4.12 : Répartition des différentes formes de ORBIT dans les trois corpus de travail.....	51

Tableau 4.13 : Répartition en pourcents des occurrences de ORBIT dans les seize sous-corpus de chaque corpus de travail.....	52
Tableau 4.14 : Proportion des formes verbales de PROBE par rapport aux formes nominales ...	54
Tableau 4.15 : Répartition des différentes formes de PROBE dans les trois corpus de travail .....	56
Tableau 4.16 : Proportion des formes verbales de PROBE par rapport aux formes nominales ...	57
Tableau 4.17 : Répartition des différentes formes de ROCKET dans les trois corpus de travail..	59
Tableau 4.18 : Proportion des formes verbales de STAR par rapport aux formes nominales .....	59
Tableau 4.19 : Répartition des différentes formes de STAR dans les trois corpus de travail .....	61
Tableau 4.20 : Pourcentage des formes verbales par rapport aux formes nominales dans les trois corpus de travail .....	62