

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

**Vulgarisation et traduction : représentation
discursive des notions scientifiques
biomédicales en français et en espagnol**

par

Mariana Raffo

Département de linguistique et de traduction
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de Maîtrise ès arts
en Traduction
option Recherche

Décembre, 2007



© Mariana Raffo, 2007

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

**Vulgarisation et traduction : représentation
discursive des notions scientifiques
biomédicales en français et en espagnol**

présenté par :
Mariana Raffo

a été évalué par un jury composé des personnes suivantes :

Jeanne Dancette, présidente-rapporteuse
Sylvie Vandaele, directrice de recherche
Nathan Ménard, membre du jury

Résumé

La caractérisation des discours spécialisés en vue de la traduction constitue un défi en raison du nombre de variables à considérer. Le présent travail de mémoire constitue l'étape exploratoire indispensable afin de dégager les pistes pour une recherche approfondie. Dans ce dessein, une méthodologie adaptée à l'étude de la mise en discours des notions spécialisées dans une optique bilingue a été mise au point. Elle vise la mise en correspondance de l'expression des notions spécialisées dans deux langues (le français et l'espagnol) tout en tenant compte des paramètres liés à une situation de communication spécifique : la vulgarisation. À la fin de cette étape préliminaire, nous pouvons dresser un bilan positif. La démarche méthodologique a permis de recueillir des données non seulement nombreuses, mais pertinentes au regard des objectifs de cette étude. Ainsi, il a été possible de mettre en évidence des différences et des similitudes au sein des corpus ainsi qu'entre les textes de vulgarisation et les textes scientifiques, qui constituent autant de pistes pour des futures recherches.

Mots-clés : Traduction spécialisée; médiation de savoirs spécialisés; approche cognitive; approche discursive; modes de conceptualisation.

Abstract

The characterization of specialized discourse from a translator's point of view poses a challenge since a great number of variables have to be taken into account. This master's thesis constitutes an exploratory study, which is crucial in order to set relevant research topics. Thus, a methodology adapted to the study of the specialized concepts' presentation in discourse from a bilingual standpoint was created. This methodology allows us to compare the expression of specialized concepts in both languages (French and Spanish) according to the parameters of a specific communicative situation: popularization. This preliminary study was successful: a great amount of relevant data was collected. Several differences and similarities between corpora, and between the characteristics of popularization and scientific text were shown. Besides, a certain number of research avenues were uncovered.

Keywords : Specialized translation; mediation of specialized knowledge; discursive approach; cognitive approach; conceptualization modes.

Table des matières

RÉSUMÉ.....	I
ABSTRACT.....	II
TABLE DES MATIÈRES.....	III
LISTE DES TABLEAUX.....	VI
LISTE DES FIGURES.....	VII
REMERCIEMENTS.....	XI
CONVENTIONS D'ÉCRITURE	XII
1. INTRODUCTION	1
1.1. Problématique	2
1.2. Questionnement	4
2. CADRE THÉORIQUE	8
2.1. Le texte, plaque tournante du sens.....	8
2.2. Les discours spécialisés : recontextualisation des savoirs.....	9
2.3. Les notions, catégorisation du réel (et de l'imaginaire)	11
2.4. Le problème du terme	13
2.5. Les notions spécialisées en discours	15
2.6. Modes de conceptualisation	19
3. ÉTAT DE LA QUESTION	21
3.1. Les langues	21
3.2. Degré de spécialité.....	22
3.3. La vulgarisation	26
3.4. Les modes de conceptualisation métaphorique.....	28
3.5. Récapitulation.....	31
4. MÉTHODOLOGIE	32
4.1. Corpus.....	32

4.2. Collecte de données.....	35
4.2.1. Éléments relevés.....	35
4.2.2. Système d'annotation.....	38
4.2.3. Récupération et traitement des données.....	43
5. RÉSULTATS.....	44
5.1. Notions	45
5.1.1. Nombre de notions et densité notionnelle.....	45
5.1.2. Actualisation des notions	46
5.1.3. Groupes de notions communes et noyau thématique.....	46
5.1.4. Hyper-catégories	51
5.2. Expressions linguistiques.....	54
5.2.1. Nombre d'expressions et expressions communes.....	54
5.2.1.1. <i>Expressions différentes</i>	58
5.2.2. Variation linguistique	60
5.2.2.1. <i>Nombre d'expressions renvoyant à une même notion</i>	60
5.2.2.1.1. Rapport avec le nombre d'actualisations.....	65
5.2.2.1.2. Rapport avec l'hyper-catégorie	67
5.2.2.2. <i>Ensembles paraphrastiques</i>	72
5.2.3. Structure morphosyntaxique	74
5.2.3.1. <i>Noms et syntagmes nominaux</i>	77
5.2.3.2. <i>Verbes et syntagmes verbaux</i>	78
5.2.3.3. <i>Adjectifs et syntagmes adjectivaux</i>	79
5.2.3.4. <i>Constructions elliptiques et formes réduites</i>	80
5.2.3.5. <i>Structure morphosyntaxique et variation linguistique</i>	80
5.2.3.6. <i>Partie du discours et hyper-catégorie</i>	81
5.3. Stratégies discursives.....	84
5.3.1. Diaphore	84
5.3.2. Reformulation.....	86
5.3.3. Définition.....	86
5.3.4. Explication.....	87
5.4. Modes de conceptualisation	89
6. DISCUSSION.....	102
6.1. Démarche méthodologique.....	102
6.1.1. Le corpus	103
6.1.2. La collecte des données	103
6.1.3. Validité des données obtenues	105

6.2. Expression linguistique des notions	105
6.2.1. Variation linguistique	106
6.2.1.1. <i>Nombre d'expressions renvoyant à une même notion</i>	106
6.2.1.2. <i>Ensembles paraphrastiques</i>	107
6.2.1.2.1. Expressions communes.....	107
6.2.1.2.2. Expressions différentes.....	110
6.2.2. Structure morphosyntaxique des expressions	111
6.2.3. Stratégies discursives.....	112
6.3. Conceptualisation métaphorique des notions.....	114
6.3.1. Indices de conceptualisation métaphorique	115
6.3.2. Modes de conceptualisation	117
6.4. Structures notionnelles.....	120
7. CONCLUSION.....	122
BIBLIOGRAPHIE.....	124
ANNEXE I.....	XIII

Liste des tableaux

Tableau 1. Composition des corpus.....	34
Tableau 2. Balises et attributs	42
Tableau 3. Notions actualisées dans les quatre sous-corpus	49
Tableau 4. Proportion moyenne des hyper-catégories dans les quatre sous-corpus	52
Tableau 5. Expressions communes aux sous-corpus Espagne et Argentine	54
Tableau 6. Expressions communes aux sous-corpus France et Québec	56
Tableau 7. Expressions employées dans le corpus en espagnol pour renvoyer aux notions /cellule souche/ et /blastocyste/	58
Tableau 8. Expressions employées dans le corpus en français pour renvoyer aux notions /blastocyste/ et /cellule souche pluripotente/	58
Tableau 9. Notions faisant l'objet d'une explication dans les quatre sous-corpus.....	87
Tableau 10. Modes de conceptualisation métaphorique et indices de conceptualisation repérés dans les corpus	92
Tableau 11. Expressions communes aux sous-corpus Espagne et Argentine retrouvées dans le moteur de recherche <i>Google Académico</i>	108
Tableau 12. Expressions communes aux sous-corpus France et Québec retrouvées dans le moteur de recherche <i>Google Scholar</i> en français	109
Tableau 13. Nombre de notions classées sous chaque hyper-catégorie dans chaque texte	xiii

Liste des figures

Figure 1. Taille moyenne des sous-corpus	34
Figure 2. Balises <fichier>, <nom>, <nature>, <langue> et <texte>	39
Figure 3. Balises <surtitre>, <titre>, <chapeau>, <sous-titre>, <par> et <phr>	39
Figure 4. Balise <lingNot> et ses attributs	40
Figure 5. Balise <expl> et ses attributs	41
Figure 6. Balise <concInd> et ses attributs	41
Figure 7. Balise <sic> et ses attributs.....	42
Figure 8. Densité notionnelle moyenne par 1 000 mots dans chaque sous-corpus.....	45
Figure 9. Densité moyenne d'actualisation de notions par 1 000 mots dans chaque sous-corpus	46
Figure 10. Proportion moyenne dans chaque sous-corpus des notions communes à l'ensemble du corpus en espagnol.....	47
Figure 11. Proportion moyenne dans chaque sous-corpus des actualisations des notions communes à l'ensemble du corpus en espagnol.....	47
Figure 12. Proportion moyenne dans chaque sous-corpus des notions communes à l'ensemble du corpus en français.....	48
Figure 13. Proportion moyenne dans chaque sous-corpus des actualisations des notions communes à l'ensemble du corpus en français	48
Figure 14. Proportion moyenne dans chaque sous-corpus des notions du noyau thématique	50
Figure 15. Proportion moyenne dans chaque sous-corpus des actualisations des notions constituant le noyau thématique.....	51
Figure 16. Distribution moyenne des notions des quatre sous-corpus en fonction de l'hyper-catégorie.....	52
Figure 17. Distribution des notions communes du corpus en espagnol en fonction de l'hyper-catégorie.....	53
Figure 18. Distribution des notions communes du corpus en français en fonction de l'hyper-catégorie.....	53
Figure 19. Distribution des notions en fonction de l'hyper-catégorie dans le noyau thématique	54

Figure 20. Proportion des expressions communes dans l'ensemble d'expressions renvoyant à des notions actualisées dans les deux sous-corpus (Espagne et Argentine).....	56
Figure 21. Proportion des expressions communes dans l'ensemble d'expressions renvoyant à des notions actualisées dans les deux sous-corpus (France et Québec).....	57
Figure 22. Distribution des notions actualisées au moins deux fois dans Espagne en fonction du nombre d'expressions employées pour les exprimer.....	61
Figure 23. Distribution des notions actualisées au moins deux fois dans Argentine en fonction du nombre d'expressions employées pour les exprimer.....	61
Figure 24. Distribution des notions actualisées au moins deux fois dans France en fonction du nombre d'expressions employées pour les exprimer.....	62
Figure 25. Distribution des notions actualisées au moins deux fois dans Québec en fonction du nombre d'expressions employées pour les exprimer.....	62
Figure 26. Distribution des notions communes au corpus en espagnol (actualisées au moins deux fois) en fonction du nombre d'expressions employées pour les exprimer.....	63
Figure 27. Distribution des notions communes au corpus en français (actualisées au moins deux fois) en fonction du nombre d'expressions employées pour les exprimer.....	64
Figure 28. Distribution des notions du noyau thématique (actualisées au moins deux fois dans chaque sous-corpus) en fonction du nombre d'expressions employées pour les exprimer.....	64
Figure 29. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Espagne.....	66
Figure 30. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Argentine.....	66
Figure 31. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus France.....	67
Figure 32. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Québec.....	67

Figures 33 et 34. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Espagne (regroupement en fonction de l'hyper-catégorie).....	68
Figures 35 et 36. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Argentine (regroupement en fonction de l'hyper-catégorie).....	69
Figures 37 et 38. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus France (regroupement en fonction de l'hyper-catégorie).....	70
Figures 39 et 40. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Québec (regroupement en fonction de l'hyper-catégorie)	71
Figure 41. Distribution des expressions des quatre sous-corpus en fonction de la structure morphosyntaxique.....	75
Figure 42. Distribution des expressions renvoyant aux notions communes du corpus en espagnol en fonction de la structure morphosyntaxique	76
Figure 43. Distribution des expressions renvoyant aux notions communes du corpus en français en fonction de la structure morphosyntaxique	76
Figure 44. Distribution des expressions renvoyant aux notions communes aux quatre sous-corpus en fonction de la structure morphosyntaxique	77
Figure 45. Distribution des notions en fonction de la partie du discours des expressions employées (notions exprimées par au moins deux expressions)	81
Figure 46. Distribution dans Espagne des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours.....	82
Figure 47. Distribution dans Argentine des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours.....	82
Figure 48. Distribution dans France des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours.....	83
Figure 49. Distribution dans Québec des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours.....	83
Figure 50. Proportion des indices de conceptualisation métaphorique au sein des expressions de chaque sous-corpus	89

A Tita

Remerciements

À Sylvie Vandaele, ma *maestra*, qui a su me guider dans les méandres de l'apprentissage de la recherche et m'a transmis sa passion du métier et du travail bien fait. Et de la linguistique, de la biologie, des livres, du cinéma, de la bonne conversation... Je souhaite que nous puissions nous amuser encore longtemps ensemble.

À Sylvie Boudreau, faiseuse de miracles, qui au cours des innombrables heures au labo, a été généreuse en conseils, sourires, soutien technique et moral, graines de citrouille, écoute, thé, livres, commentaires pertinents et chocolat.

Aux professeurs du Département de linguistique et de traduction de l'Université de Montréal, qui m'ont inspirée et encouragée.

À mes parents, Daniel et Susana, qui ont le talent d'être près à distance.

À Juanma, mon *compañero*, qui m'aime comme j'aime, à sa façon.

Et à Milton, bien sûr.

Conventions d'écriture

Notion théorique (première occurrence)	<i>moment discursif</i>
Notion spécialisée	<i>/cellule/</i>
Unité linguistique ou syntagme tiré du corpus en français	cellule
Unité linguistique ou syntagme tiré du corpus en espagnol	<u><i>célula</i></u>

Système d'annotation du corpus

Balises	<lingNot>
Attributs	@notion
Valeurs	" <i>cellule souche</i> "

1. Introduction

« Ce dont il s'agit ici, ce n'est pas de neutraliser le discours, d'en faire le signe d'autre chose et d'en traverser l'épaisseur pour rejoindre ce qui demeure silencieusement en deçà de lui, c'est au contraire de le maintenir dans sa consistance, de le faire surgir dans la complexité qui lui est propre »

M. Foucault, *L'archéologie du savoir*, Paris, Gallimard, 1969, p. 65.

Nous les traducteurs, nous partons des mots et, une fois le travail fini, il ne nous reste à nouveau que des mots. Ils sont le support matériel de la communication et, en même temps, ils en constituent l'élément le plus volatil, car la forme linguistique d'origine disparaît au profit d'une autre. Ce qui doit rester intact, c'est le *sens*. Mais, où réside-t-il? Peut-on le trouver au-delà de cette matière instable constituée par les mots d'une langue, ou y est-il solidement imbriqué? Cette question vieille comme l'histoire est à l'origine de la dichotomie canonique de la réflexion traductologique. Depuis l'Antiquité, au cœur des polémiques sur la meilleure façon de traduire se trouve l'opposition entre le contenu et la forme. Dès 395, des détracteurs forcent Saint Jérôme à défendre sa méthode de traduction : « *ce n'est pas un mot par un mot, mais une idée par une idée que j'exprime* » (Saint Jérôme 1953, p. 59). Mais des siècles plus tard, en 1955, il y a toujours un Nabokov pour affirmer de façon lapidaire que « *the clumsiest literal translation is a thousand times more useful than the prettiest paraphrase* »¹ (Nabokov 1955, p. 496).

Quoi qu'il en soit, le cas de la traduction spécialisée semble pouvoir être réglé assez vite, car les textes dits pragmatiques ont d'emblée été exclus du débat au profit des œuvres littéraires, philosophiques et religieuses². En effet, contrairement aux écrits littéraires, dont la matérialité graphique ou acoustique fait aussi partie du sens (songeons au célèbre exemple de Racine : « *Pour qui sont ces serpents qui sifflent sur vos têtes...* »³), les textes scientifiques et techniques semblent constituer l'exemple par excellence d'indépendance vis-à-vis de la forme. Étant donné que ces textes visent à transmettre un savoir encyclopédique, réputé objectif et universel, leur sens se situerait exclusivement au niveau

¹ [Notre traduction :] « *La plus maladroite des traductions littérales est mille fois plus utile que la plus belle des paraphrases* ».

² Schleiermacher ne fait que continuer et sanctionner une tradition lorsqu'il limite le champ d'action du véritable traducteur aux textes où « *l'objet ne domine plus, mais est dominé par la pensée et la sensibilité, voire n'existe fréquemment que dans et par le discours* ». Dans Schleiermacher, F. (1985). « Des différentes méthodes du traduire ». Dans A. Berman (dir.), *Les Tours de Babel. Essais sur la traduction*, Mauvezin: Trans-Europ-Repress. S'il est vrai que sa définition vise les domaines de l'art et de la science, elle exclut catégoriquement les textes pragmatiques et va même à l'encontre des principes qui définissent la démarche scientifique.

³ Racine, J. (1994). *Andromaque* (préface de R. Picard, édition de J.-P. Collinet), Paris : Gallimard. Acte V, Scène V.

du contenu. C'est ce qui rendrait la tâche à la fois difficile et facile. Difficile parce que, le traducteur n'étant pas nécessairement un expert du domaine, la compréhension des notions représente un défi de taille. Facile parce qu'il n'aurait qu'à s'assurer de l'équivalence notionnelle sans se heurter aux écueils de l'expression linguistique.

1.1. Problématique

Les choses ne sont pourtant pas si simples, car le savoir est étroitement lié au langage. L'homme découvre le monde à travers ses sens, mais le rend intelligible avec des mots. Connaître vraiment quelque chose exige de pouvoir le nommer; et pas de n'importe quelle façon. Malgré l'absence, en règle générale, d'un souci esthétique, la transmission de connaissances impose un emploi particulier de la langue. Ainsi, ces formulations ne sont pas interchangeables :

(1) « *Placez les boutons de commande sur les tiges de valve situées sur le panneau de commande.* »¹

(2) « *Mettez les deux trucs ronds sur les deux petits bâtons qui sont sur la planche en métal.* »

S'il est possible d'admettre que les deux phrases ont le même contenu notionnel, un locuteur de langue française n'hésitera pas à signaler la première, et non la deuxième, comme relevant du discours d'un expert. Se pose donc la question de comment caractériser linguistiquement les textes spécialisés.

Il y a d'abord les termes. Ils permettent de (dé)nommer les notions et « *tissent la trame du domaine* » (Balliu 2005). Selon les principes de la Théorie générale de la terminologie fondée par Wüster, dite « théorie classique », ils désigneraient les notions de manière précise et biunivoque, c'est-à-dire que chaque terme ne servirait qu'à nommer une seule notion et que chaque notion ne serait désignée que par un seul terme (Cabré 2003). Ainsi, en traduction spécialisée, l'accent a surtout été mis sur la recherche et l'apprentissage de la terminologie du domaine et des équivalents dans la langue d'arrivée. Encore une fois, les choses ne sont pas si simples. Dans les textes, les termes n'apparaissent pas de la même façon que dans les dictionnaires; on y trouve fréquemment des variantes morphosyntaxiques, des sigles, des formes réduites, etc. Or, comment être sûr que l'on fait référence à la même notion lorsqu'on dit « *imagerie par résonance magnétique nucléaire* », « *imagerie par résonance magnétique* », « *imagerie à résonance magnétique* » et « *IRM* »?

¹ MasterChef (2007). *Guide d'utilisation, d'entretien et d'assemblage. H405 barbecue au propane*, Toronto.

Qui plus est, la précision et la biunivocité de la langue scientifique prônées par la théorie classique disparaissent souvent en discours. Une même notion peut recevoir plusieurs dénominations pour des motifs très divers. L'usage spécifique à la région géographique, par exemple, justifie qu'une même maladie soit appelée « mucoviscidose » (en France) ou « fibrose kystique » (au Québec). Et la prise en compte des savoirs de l'interlocuteur déterminera l'emploi, en espagnol, de « tripanosomiasis cruzi » ou de « mal de Chagas ». De plus, les termes ne sont pas toujours dépouillés de toute indétermination et ambiguïté. Que dire ainsi du terme « inhibition enzymatique »? Au sein d'un même domaine, il désigne l'arrêt ou la diminution soit de l'activité d'une enzyme, soit d'une réaction où l'une d'entre elles intervient. Il s'ensuit que, de la même façon qu'un dictionnaire bilingue ne suffit pas pour traduire un roman, connaître par cœur une liste de termes ne garantit pas une bonne traduction d'un rapport technique.

Par ailleurs, la communication spécialisée comporte d'autres contraintes sur le plan linguistique. Autant les traducteurs et les rédacteurs que les terminologues reconnaissent « *l'importance que revêt l'usage d'une phraséologie adéquate, ce qui est généralement évoqué sous le vocable un peu flou de "génie de la langue"* » (Vandaele 2005, p. 276). Pavel définit la phraséologie comme « *la combinatoire syntagmatique des unités terminologiques relevant d'une structure conceptuelle cohérente* » (Pavel 1993, p. 70). Cela veut dire que l'environnement linguistique des termes, d'une part, contribue à l'idiomaticité du texte et, d'autre part, met en évidence la structuration des savoirs du domaine. En effet, ces unités souvent témoignent des divers modes de conceptualisation sous-jacents au domaine, ce qui pose parfois des problèmes pour la traduction. C'est le cas de l'adjectif (participe passé) « ancré » dans la phrase :

(3) « *La protéine E24 est **ancrée** dans la membrane par sa région C-terminale.* »⁵

Le contraste avec

(4) « *La protéine E24 est **fixée** dans la membrane par sa région C-terminale.* »

fait ressortir son caractère idiomatique et met en évidence un mode de conceptualisation propre aux sciences biomédicales : concevoir les protéines comme des bateaux, ou à tout le moins comme des corps flottant sur la membrane, lipidique et fluide (Vandaele *et al.* 2006).

Enfin, au niveau textuel, plusieurs auteurs (notamment, Hoffmann 1985, Sager *et al.* 1980, Kocourek 1991) ont mis en évidence les spécificités propres à ce qu'on appelle la « langue - ou les langues - de spécialité » ou « LSP » : la prédominance des formes

⁵ Teyssier E. (1997) *Caractérisation de protéines de l'enveloppe des chloroplastes d'épinard*, Thèse de Doctorat, Université de Grenoble 1, Saint-Martin-d'Hères, France.

impersonnelles et du présent de l'indicatif, la préférence pour des phrases et des syntagmes longs, la condensation syntaxique du contenu, la tendance à la nominalisation et à l'adjectivation, l'emploi d'unités brachygraphiques, etc. Il ne s'agit pas cependant de traits nécessaires et suffisants : ils ne se trouvent pas tous systématiquement dans chaque texte technique et scientifique, et aucun d'entre eux n'est exclusif à la communication spécialisée. Kocourek nous met en garde contre l'identification de la langue de spécialité à un style ou registre « *du fait qu'elle se différencie elle-même en styles et registres* » (Kocourek 1991, p. 17). En effet, les connaissances nées dans les laboratoires débordent des cercles des experts et circulent au sein de différents « circuits du savoir » dans un large éventail de discours produits par divers énonciateurs à l'intention d'un vaste public de lecteurs initiés et non initiés (Calsamiglia et Cassany 1999).

Il devient ainsi évident que la problématique de la traduction scientifique ne se borne pas à la compréhension d'un contenu ou à l'apprentissage des termes d'un domaine. Elle est constatée en discours, lorsque le traducteur doit prendre des décisions liées à la variation terminologique, à la phraséologie et aux divers registres de la communication spécialisée. En effet, il faut garder à l'esprit que le point de départ d'une traduction se trouve toujours dans une énonciation : « *un acte au cours duquel ces phrases s'actualisent, assumées par un locuteur particulier, dans des circonstances spatiales et temporelles précises* » (Ducrot et Todorov 1972, p. 405). Les textes sont ainsi contraints par de multiples variables qui interagissent à différents niveaux : non seulement le contenu que l'on veut transmettre et la façon dont il est conceptualisé, mais aussi la langue choisie pour l'exprimer, les savoirs, intentions et attitudes de l'auteur et des lecteurs visés ainsi que le lieu et le moment de l'énonciation, entre autres. Les textes spécialisés ne constituent pas une exception; ils s'actualisent dans des situations réelles de communication dont les paramètres ont un impact autant sur la forme que sur le contenu (Calsamiglia et al. 1998, Cassany et al. 2000).

1.2. Questionnement

Reconnaître que le discours constitue la matière première du travail du traducteur nous pousse à aborder les difficultés sous un nouvel angle et à étudier l'expression linguistique en rapport avec les variables susceptibles de l'influencer. Nous nous interrogeons sur la façon dont s'exerce cette influence, et surtout, sur le rôle joué par la langue employée (dans notre cas, le français ou l'espagnol) au sein de cette interaction. Bien sûr, le grand nombre de facteurs linguistiques et extralinguistiques à prendre en compte décourage toute tentative de donner une réponse globale dans le cadre d'un mémoire de maîtrise.

Cependant, à partir de cette première question très générale, nous pouvons déceler d'autres questions de singulière importance pour la communication et la traduction spécialisées en ciblant particulièrement la présentation en discours – l'expression linguistique et les modes de conceptualisation – des notions spécialisées.

Puisque tout locuteur adapte son propos au public auquel il s'adresse, la façon dont les notions sont conceptualisées et exprimées linguistiquement dépend largement du niveau de connaissance des destinataires visés (en terminologie, Cabré 2000, Sager 2000, en analyse du discours, Calsamiglia et Van Dijk 2004, Cassany et al. 2000). Il est pourtant nécessaire de préciser qu'il ne s'agit pas du savoir que les lecteurs pourrait effectivement posséder, mais de celui que l'auteur suppose chez son lectorat⁶. Ainsi, en tenant compte des savoirs supposés, il est possible de déterminer au niveau pragmatique le degré de spécialisation des textes et de postuler une échelle allant des articles scientifiques produits par et pour des experts (les plus spécialisés) jusqu'aux textes de vulgarisation qui s'adressent au grand public (les moins spécialisés). Plusieurs typologies s'appuient sur ce critère, notamment celle de Loffler-Laurian (1983), mais aussi celles de Gläser (1993), Weise (1993) et Ciapuscio (2003a).

Situer les textes dans cette échelle de spécialisation permet d'étudier l'expression linguistique et les modes de conceptualisation des notions spécialisées en fonction du niveau de connaissances supposé des destinataires. Dans une perspective de traduction, il est particulièrement intéressant de se pencher sur la variation qui pourrait être entraînée par la langue employée, d'une part, et par le lieu où le texte est publié (son environnement socioculturel), d'autre part. Notre travail porte spécifiquement sur les textes qui se trouvent à l'extrémité de l'échelle : articles journalistiques de vulgarisation. Ayant pour tâche de présenter des notions spécialisées au grand public, ils opèrent une recontextualisation des savoirs spécialisés (Hall *et al.* 1999, cité dans Ciapuscio 2003b) où la dimension sociale devient fondamentale. Au cours de cette recherche, nous cherchons à répondre aux questions suivantes :

1. Comment sont exprimées linguistiquement les notions spécialisées dans les articles journalistiques de vulgarisation?

Nous nous intéressons principalement au nombre d'unités et syntagmes employés pour renvoyer à une même notion (variation linguistique), aux caractéristiques morphosyntaxiques de ces expressions et aux liens de coréférence qui les unissent. Les travaux de certains chercheurs (Ciapuscio 2003a, Freixa 2002) évoquent un taux de

⁶ La distinction entre le lecteur véritable et l'image que le locuteur s'en fait a été théorisée par des chercheurs en analyse du discours comme Charaudeau, qui distingue le Tu interprétant du Tu destinataire (construit par l'auteur).

variation plus élevée dans les textes moins spécialisés. Ainsi, il devient logique d'envisager que, pour chaque notion véhiculée, un grand nombre d'expressions soient employées et qu'elles se présentent sous des formes linguistiques diverses, constituant un riche ensemble paraphrastique. En outre, plusieurs auteurs (Calsamiglia et Van Dijk 2004, Cassany *et al.* 2000, Cassany et Martí 1998) mettent en évidence la présence dans le discours de vulgarisation des stratégies discursives qui aident le lecteur à bâtir une représentation de la notion et à l'intégrer à une structure des connaissances. Nous faisons donc l'hypothèse que les textes journalistiques privilégieront les stratégies qui renforcent la cohérence textuelle en instaurant des liens entre les expressions : la diaphore (reprise anaphorique et cataphorique), la reformulation, la définition et l'explication.

2. Quels sont les modes de conceptualisation des notions spécialisées véhiculés dans les articles journalistiques de vulgarisation et comment se manifestent-ils dans le discours?

Dans cette étude, nous ciblons spécifiquement la conceptualisation métaphorique. En effet, de nombreux chercheurs ont souligné le rôle de la métaphore en tant que stratégie de vulgarisation de connaissances (Knudsen 2003, Liakopoulos 2002, Loffler-Laurian 1994, pour ne nommer que quelques-uns). Par ailleurs, il est reconnu que la conceptualisation métaphorique sous-tend le discours hautement spécialisé ainsi que l'élaboration des théories scientifiques (Temmerman 2002b, Vandaele *et al.* 2006, Fox Keller 1995). Nous nous attendons donc à retrouver dans les textes analysés une densité importante d'expressions métaphoriques, mais pensons aussi à la possibilité de l'usage d'analogies, faisant appel à une comparaison explicite. Nous tenterons également de savoir si les métaphores des textes de vulgarisation sont compatibles avec celles des textes scientifiques.

- 3.a Quel est l'impact de la langue dans laquelle le texte est écrit (le français ou l'espagnol) sur la présentation des notions spécialisées dans les articles journalistiques de vulgarisation?
- 3.b Quel est l'impact du contexte socioculturel dans lequel le texte apparaît (la France, le Québec, l'Espagne et l'Argentine) sur la présentation des notions spécialisées dans les articles journalistiques de vulgarisation?

Il est évident que l'emploi d'une langue spécifique détermine la forme linguistique employée pour exprimer les notions, mais l'intérêt du point de vue traductologique réside surtout dans le rapport entre le contexte socioculturel et le choix des expressions. De plus, nous nous interrogeons sur l'influence de la langue employée par l'auteur du texte et du lieu de publication sur le nombre d'expressions renvoyant à une même notion et sur leurs caractéristiques morphosyntaxiques. Il reste également à déterminer si les stratégies

discursives déployées dans le texte dépendent de la langue dans laquelle il est rédigé ou du contexte socioculturel dans lequel il apparaît. Ces variables pourraient aussi jouer un rôle dans la façon dont les notions spécialisées sont conceptualisées. En effet, en plus des différences dans le niveau des connaissances du public, l'intérêt accordé par la société à un sujet donné (la recherche sur les cellules souches, par exemple) et les enjeux qui y sont liés peuvent varier d'une société à l'autre, ce qui influence la manière dont on en parle.

Afin de tester ces hypothèses, nous avons réalisé une étude sur corpus bilingue français-espagnol. Les textes compilés se situent au pôle le moins spécialisé de l'échelle de spécialisation : des articles journalistiques de vulgarisation. Il s'agit de textes originaux (non traduits) produits dans quatre contextes socioculturels différents : la France, le Québec, l'Espagne et l'Argentine. La méthode d'analyse consiste à repérer les unités et syntagmes utilisés pour référer à des notions spécialisées (liées à la recherche sur les cellules souches et le clonage) au moyen d'un système d'annotation des textes créé *ad hoc*. L'avantage de notre démarche réside dans la possibilité d'étudier les séquences linguistiques sans devoir juger a priori de leur statut terminologique et de leur structure morphosyntaxique. En revanche, l'étude se voit limitée aux notions et aux expressions présentes dans le discours vulgarisé, car une comparaison avec des textes scientifiques détaillée et menée avec rigueur exige un travail d'une plus grande envergure.

Cette recherche poursuit des objectifs pratiques, mais surtout méthodologiques et théoriques. En ce qui concerne la pratique de la traduction, nous sommes convaincue que la connaissance des mécanismes de mise en discours des notions spécialisées dans différentes langues est fondamentale pour prendre des décisions de traduction éclairées. Les questions posées se trouvent au cœur du travail de tout traducteur; les rendre conscientes permet de systématiser autant l'analyse du texte de départ que la production du texte d'arrivé. Sur le plan méthodologique, nous offrons notre apport à l'étude du discours spécialisé, non pas comme une sous-langue s'opposant à une langue générale, mais considéré dans sa dimension pragmatique. Nous proposons ainsi une démarche méthodologique originale pour l'étude de la réalisation linguistique des notions spécialisées en discours. Sur le plan théorique, il s'agit d'insister sur le fait que la traduction ne se réduit pas à un transfert linguistique et sur le besoin de prendre en compte les divers paramètres de la situation de communication. Le nombre d'éléments qui contribuent au sens du texte et la complexité des liens existants entre eux nous forcent à diviser la recherche en plusieurs étapes. Ce travail en constitue la première et se veut également une phase exploratoire qui nous aidera à déterminer la pertinence et la faisabilité d'une étude plus approfondie de certains phénomènes.

2. Cadre théorique

Afin de mener à bien cette étude, il était essentiel de se doter d'un appareil théorique cohérent qui guiderait nos choix méthodologiques. Plutôt que de dresser un tableau d'un ou de plusieurs courants de pensée, nous exposerons dans ce chapitre le point de vue adopté pour aborder la problématique ainsi que les éléments théoriques qui sous-tendent notre analyse. Il s'agit donc de préciser notre conception du texte, du discours spécialisé et de la réalisation linguistique des notions spécialisées, mais surtout de définir les concepts clés pour notre travail.

2.1. Le texte, plaque tournante du sens

Du fait qu'il renvoie à une instance d'énonciation, le **texte** se transforme en une structure multidimensionnelle. Il s'agit certes d'un objet linguistique, mais qui réalise une intention communicative en fonction de paramètres situationnels spécifiques. Il instaure un espace d'interaction - les lecteurs ne sont pas des récepteurs passifs! - où convergent des connaissances linguistiques et extralinguistiques ainsi que des expériences et des visions du monde. Ainsi, son sens ne saurait être le résultat d'une somme de phrases. Il faut ici indiquer ce que nous entendons par **sens** et **signification**, qui peuvent donner lieu à confusion en raison des emplois divers qu'ils ont dans la littérature. Dans ce travail, nous adoptons l'usage proposé par Ducrot (1980) : nous utilisons le premier lorsqu'il s'agit d'interpréter un énoncé en tenant compte de la situation extralinguistique, et réservons le deuxième pour renvoyer aux propriétés sémantiques de la langue. Ce choix nous permet d'articuler les divers horizons qui constituent notre cadre théorique et de situer notre réflexion par rapport aux discussions traductologiques.

Notre vision du texte est principalement redevable des courants de l'**analyse du discours**, qui envisagent l'étude du langage dans sa relation avec la **situation de communication**. En effet, autant les circonstances temporelles et spatiales que les participants contraignent la façon dont la langue est employée. Tout sujet parlant le sait : c'est ce qui lui permet de comprendre une pièce de Molière, d'arriver à midi à un dîner à Montréal et de se méfier des promesses des politiciens. La production et l'interprétation des textes se fondent précisément sur les savoirs des interlocuteurs sur le monde, d'une part, et sur la situation elle-même, d'autre part. Charaudeau (1983) le précise : « *parler ou écrire c'est, pour le sujet énonçant, faire des hypothèses sur le savoir du sujet interprétant et sur les points de*

vue de celui-ci vis-à-vis du propos langagier et de lui-même, sujet énonçant » (p. 24). Les circonstances du discours sont donc constituées par « *l'ensemble des savoirs supposés qui circulent entre les protagonistes du langage »* (Charaudeau 1983, p. 25).

Au sein de la traductologie, ce sont surtout les *théories fonctionnalistes* qui se sont penchées sur les choix et contraintes entraînés par la production d'un texte en fonction d'une nouvelle situation de communication. Ces approches s'articulent autour de la notion de *skopos* (« but ») pouvant servir pour rendre compte autant de la visée communicative du traducteur que des fonctions du texte original et de sa traduction (Reiss et Vermeer 1984/1991). Lors de la traduction, les stratégies et les procédés utilisés dépendent du *skopos* de l'acte traductionnel et de celui du texte d'arrivée; ainsi, le traducteur d'un roman pourrait tenter de calquer la syntaxe afin de mettre en évidence l'usage particulier qu'en fait l'auteur ou, au contraire, traduire les noms propres pour que « ça ne sente pas la traduction ». Dans un cas comme dans l'autre, « *the receiver, or rather the addressee, is the main factor determining the target-text Skopos »*⁷ (Nord 1997, p. 29). Les théories fonctionnalistes tiennent compte également des hypothèses faites par l'énonciateur : « *the translator offers this new audience a target text whose composition is, of course, guided by the translator's assumptions about their needs, expectations, previous knowledge, and so on. »*⁸ (Nord 1997, p. 35).

2.2. Les discours spécialisés : recontextualisation des savoirs

Les savoirs supposés déterminent les « *styles et registres* » auxquels fait référence Kocourek (1991, p. 17). Plusieurs typologies textuelles tiennent compte des paramètres situationnels pour la classification, et se servent particulièrement du niveau de connaissances que l'on attribue aux lecteurs pour établir le *degré de spécialisation* des textes et les hiérarchiser (Ciapuscio 2003a, Loffler-Laurian 1983). En effet, prendre en compte les circonstances du discours mène à envisager la spécialisation d'un texte en termes de gradation, et à réviser l'opposition entre la « langue générale » et les « langues de spécialité ».

Dans les faits, il y aurait plutôt un *continuum* entre les discours qui ont pour but de véhiculer des savoirs spécialisés. Jacobi et Schiele voient une « *continuité des pratiques de sociodiffusion des connaissances au sein du champ scientifique »* (Jacobi et Schiele 1988,

⁷ [Notre traduction :] « *le récepteur, ou plutôt le destinataire, est le facteur qui détermine le skopos du texte cible »*.

⁸ [Notre traduction :] « *le traducteur offre à ce nouveau public un texte cible dont l'écriture est guidée, bien évidemment, par ses suppositions quant à leurs besoins, leurs attentes, leurs savoirs préalables, etc. »*.

p. 36). Ces pratiques configurent un **circuit discursif**, c'est-à-dire, un circuit social où toutes les formes discursives, des présentations des résultats de recherche dans des congrès et des publications savantes jusqu'aux articles de vulgarisation des quotidiens, se trouvent reliées.

« [Los] discursos especializados (los que formulan con más precisión el conocimiento y que han sido reconocidos y legitimados por la comunidad científica) son el resultado de un conjunto previo de actividades investigadoras -y discursivas - y el punto de partida de otras prácticas y tareas de divulgación y adaptación del conocimiento »⁹
(Cassany et al. 2000, p. 75).

C'est à cet ensemble de productions discursives visant à diffuser des savoirs spécialisées que nous faisons référence lorsque nous parlons de **communication spécialisée**. Elle comprend des textes dont le degré de spécialisation est variable : des **discours scientifiques**, produits par et pour des experts (comme les articles scientifiques publiés dans les revues savantes), mais aussi des **discours de vulgarisation**, des textes rédigés par un expert ou par un médiateur (journaliste ou vulgarisateur) pour une audience plus large. Nous tenons à attirer l'attention sur le fait que nous considérons les discours de vulgarisation comme des discours spécialisés au même titre que les discours scientifiques (spécialisé fonctionnant alors comme terme générique).

Le fait d'intégrer un circuit discursif donné n'implique pas nécessairement une uniformité des discours. Au contraire, Cassany et ses collaborateurs (2000) attirent l'attention sur l'hétérogénéité de la vulgarisation, non seulement en raison des contenus variés, mais aussi quant aux buts (annoncer une découverte, sensibiliser le public à une cause, etc.), aux auteurs (journaliste, spécialiste) et aux contextes dans lesquels ils se produisent (pédagogique, journalistique, etc.). Ainsi, il est possible de voir le texte comme une **offre cognitive** (Ciapuscio 2003a). Il propose à son destinataire, non seulement un ensemble de notions, mais aussi une certaine représentation de celles-ci ainsi qu'une configuration particulière du réseau notionnel.

Dans ce cadre, à l'instar de Ciapuscio, nous pouvons considérer la communication scientifique, et spécifiquement la vulgarisation, comme une **recontextualisation** des connaissances :

« *the activities of communicating and transmitting science may be regarded as a continuous process of recontextualization and reformulation - one that starts with*

⁹ [Notre traduction :] « les discours spécialisés (ceux qui expriment le plus précisément les connaissances et qui ont été reconnus et légitimés par la communauté scientifique) découlent d'un ensemble d'activités de recherche - et discursives - préalables, et sont à l'origine d'autres pratiques et d'autres tâches de vulgarisation et d'adaptation des connaissances ».

press releases or scientific articles and probably ends with scientific news for the lay reader »¹⁰ (Ciapuscio 2003b, p. 210)

Au cours de cette opération de recontextualisation, selon Cassany et ses collaborateurs (2000), l'auteur d'un texte de vulgarisation exécute trois tâches cognitives : la réélaboration du réseau conceptuel, le choix de dénominations¹¹ et la mise en place de stratégies discursives de façon à ce que le lecteur puisse intégrer les notions à son système de connaissances. En effet, un chercheur n'abordera pas son sujet de la même façon dans un article pour une revue savante qu'au cours d'une présentation dans un congrès. Il modulera son discours en fonction de ce qu'il sait et de ce qu'il pense que l'autre sait, d'une part, sur le monde (son article sera plus ou moins détaillé selon l'état d'avancement de ses travaux et selon la spécificité de la revue par rapport à sa discipline) et, d'autre part, sur la situation de communication (comme la possibilité de poser des questions après une présentation).

2.3. Les notions, catégorisation du réel (et de l'imaginaire)

Avant d'aller plus loin, il est essentiel de donner quelques précisions sur ce que nous entendons par **notion**. Pour cette recherche, nous ne considérons pas nécessaire de distinguer notion et **concept**. Nous utilisons la première dénomination au détriment de la seconde pour des raisons purement opératoires : afin d'éviter les confusions lorsque nous traitons de la conceptualisation métaphorique. À partir de la psychologie cognitive (notamment Le Ny 2005), nous considérons la notion comme une représentation mentale permettant de catégoriser les **référents**, qui peuvent être des objets réels (les oiseaux) ou imaginaires (les licornes). Il peut s'agir également d'instances référentielles (un embryon) ou de situations référentielles en entier (le clonage d'un embryon par un chercheur) apparentées aux « *scènes* » de Fillmore (1977).

Une telle conception peut sembler trop étrangère à la pensée terminologique. En effet, nous nous éloignons de la définition classique de la notion en tant que constituant du terme, associé à une dénomination (Rondeau 1984). Rappelons pourtant que, sous cet angle, Sager (1990) distingue « *the individual objects of our sensation, perception and*

¹⁰ [Notre traduction :] « *les activités de communication et de transmission du savoir scientifique peuvent être considérées comme un processus continu de recontextualisation et de reformulation - il commence dans les communiqués de presse ou les articles scientifiques et finit dans les rubriques scientifiques lues par le grand public* ». Ciapuscio emprunte la notion de recontextualisation à Hall, C. *et al.* (1999). « *Speech Representation in Social Work Discourse* », *Text*, 19(4), p. 539-570.

¹¹ Ces auteurs donnent à « dénomination » un sens très large, semblable à celui que nous donnons à « expression » (voir ci-dessus).

imagination and the abstract categories, i. e. the concepts which represent them »¹² (p. 22). Il définit les notions comme « *constructs of human cognition processes which assist in the classification of objects by way of systematic or arbitrary abstraction* »¹³ (Sager 1990, p. 22).

Les notions peuvent à leur tour être regroupées dans des catégories représentant un niveau d'abstraction plus élevé : les **hyper-catégories** (Le Ny 2005). Ainsi, la notion de /cellule/ peut être rangée sous les entités concrètes, tandis que la notion /pluripotence/ sera classée dans les propriétés. Le Ny rappelle que les hyper-catégories ne sont pas « *clairement conceptualisées, et désignées de façon stable, par tous les locuteurs. Ce qui caractérise les hyper-catégories, c'est au contraire d'être le plus souvent floues* » (Le Ny 2005, p. 248). C'est ce qui pourrait expliquer l'absence d'un consensus dans la littérature quant à la liste des hyper-catégories.

Sager reconnaît quatre classes de notions plus vastes : les entités (« *entities* »), les activités (« *activities* ») – sous laquelle sont rangés les processus, les opérations et les actions dans lesquels interviennent les entités –, les qualités (« *qualities* ») – englobant les propriétés et les dispositions permettant de différencier les entités –, et les relations (« *relations* »). Il conclut cependant que chaque domaine exige une structure propre permettant de rendre compte des types de notion qui le constituent.

Nous avons établi une liste de sept hyper-catégories spécifiques à notre étude dont nous nous servirons pour classer les notions présentées dans les corpus analysés : **entité concrète, entité abstraite, action, processus, propriété, état et phénomène**¹⁴. N'ayant pas retrouvé dans la littérature des balises claires pour les reconnaître, nous fournissons pour chacune d'entre elles une brève définition tirée du *Trésor de la langue française informatisé* (ATILF 2004).

- Entité : Chose réelle, existante mais représentable uniquement par un concept.
- Action : Opération d'un agent (animé ou inanimé, matériel ou immatériel) envisagée dans son déroulement.
- Processus : Suite continue de faits, de phénomènes présentant une certaine unité ou une certaine régularité dans leur déroulement.
- Propriété : Caractère distinctif qui appartient à un être, une espèce, mais qui ne lui appartient pas toujours exclusivement.

¹² [Notre traduction :] « *les objets individuels appréhendés au moyen des sens, de la perception et de l'imagination et les catégories abstraites, c'est-à-dire, les concepts qui les représentent* ».

¹³ [Notre traduction :] « *des constructions résultant de processus cognitifs humains qui collaborent à la classification des objets par l'entremise de l'abstraction systématique ou arbitraire* ».

¹⁴ Certaines hyper-catégories ont été mises en relation avec les classes sémantiques et grammaticales (Cordier, F. et J. François (dirs.) (2002). *Catégorisation et langage*, Paris: Hermès Science Publications.), mais nous ne les reprenons que pour un classement conceptuel.

- État : Manière d'être (soit stable, soit sujette à des variations) d'une personne ou d'une chose.
- Phénomène : Ce qui apparaît, ce qui se manifeste aux sens ou à la conscience, tant dans l'ordre physique que dans l'ordre psychique [...].

Pour nous, la notion constitue le pivot entre, d'une part, les référents qu'elle permet de classer et, d'autre part, les formes linguistiques qui y renvoient en discours, qu'il s'agisse d'unités terminologiques ou non. Dans un texte, nous pouvons identifier les **actualisations** d'une notion, c'est-à-dire, les fois où l'on y fait référence, et examiner les **expressions** qui sont employées pour l'actualiser.

Situer la notion au plan strictement cognitif impose de postuler une différence avec le plan sémantique. Nous distinguons donc le **signifié** d'une forme linguistique de la notion à laquelle elle renvoie. Le premier correspond à la signification linguistique d'une expression, tandis que la notion permet de mettre cette expression en rapport avec un certain nombre de référents pouvant être désignés par elle. Certains théoriciens de la terminologie défendent également cette position. Depecker, par exemple, considère que cette distinction permet de « *regrouper de façon indubitable autour de chaque concept les synonymes qui y renvoient* » en précisant que « *la synonymie à l'œuvre ici est d'ordre référentiel* » (Depecker 2005, p. 9). Il reconnaît une difficulté supplémentaire dans le cas de la traduction.

2.4. Le problème du terme

En principe, à partir de la Théorie générale wüstérienne, le lien entre la notion et sa réalisation linguistique est concrétisé dans un **terme** entendu comme une étiquette linguistique renvoyant de façon monosémique et univoque à la notion. Pourtant, les récents débats terminologiques remettent cette définition en question, car elle ne permet pas de caractériser les unités terminologiques de façon réaliste. Ne voulant pas nous engager dans une description détaillée des divers courants modernes de la terminologie¹⁵,

¹⁵ Pour une discussion approfondie, nous renvoyons à quelques auteurs et travaux clés. La **socioterminologie** s'organise principalement autour des chercheurs de l'école de Rouen, notamment Gambier et Gaudin (Gaudin, F. (1993). *Pour une socioterminologie : Des problèmes pratiques aux pratiques institutionnelles*, Rouen: Publications de l'Université de Rouen.). Pour sa part, Temmerman propose une **approche sociocognitive** de la terminologie (Temmerman, R. (2000). *Towards New Ways of Terminology Description, the Sociocognitive Approach*, Amsterdam/Philadelphia: John Benjamins.). Enfin, Cabré tient compte des diverses positions critiques dans sa **théorie communicative** ou théorie des « portes » (Cabré, M. T. (1999). *La terminología, representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos*, Barcelona: IULA.). Un ouvrage collectif récent réunit les apports de plusieurs chercheurs, dont Sager, Depecker et Condamines, qui examinent la nature linguistique du terme et les liens entre terminologie et linguistique (Béjoint, H. et P. Thoiron (dirs.) (2000). *Le sens en terminologie*, Lyon: Presses Universitaires de Lyon.).

nous mentionnerons trois critiques qui rendent, à notre avis, la notion de terme problématique en regard des objectifs de notre recherche.

La première concerne le rapport d'univocité entre le terme et la notion. Comme le dit Depecker,

« loin que la terminologie se résume à de simples bijections - un concept pour une désignation et une désignation pour un concept -, les processus de la synonymie¹⁶ sont constamment à l'œuvre : image de satellite, image du satellite, image transmise par le satellite, image satellitaire, etc » (Depecker 2005, p. 8).

Le choix parmi ces « synonymes » fait intervenir des critères autres que référentiels. En effet, en plus d'une dimension cognitive, qui les met en lien avec la structure notionnelle du domaine, les termes comportent une dimension communicative et une dimension linguistique - ce qui est notamment mis en évidence dans la « théorie des portes » avancée par Cabré (2003 et travaux précédents).

On reproche également à la théorie classique son approche essentiellement prescriptive.

« Les praticiens de la terminologie [...], sauf ceux qui ont participé à une situation dont le but était la standardisation des notions et des termes, ont signalé certaines incohérences et des décalages entre ce qui, en théorie, constitue quelques principes et les données que l'on trouve en réalité » (Cabré 2000, p. 26).

Un regard strictement prescriptif empêche de rendre compte de l'usage réel des termes, autrement dit, d'examiner le comportement de ces unités en discours et leur rapport, d'une part, avec l'environnement linguistique et, d'autre part, avec la situation et les participants de la communication. Ainsi, certains chercheurs soulignent l'importance de considérer les unités terminologiques comme des signes linguistiques et de les étudier en relation avec les mots (notamment L'Homme 2005, mais aussi Sager 2000, Cabré 2000). La nature sociale et historique des emplois terminologiques est mise de l'avant par les tenants de la socioterminologie (par exemple, Gambier 1991).

Enfin, la terminologie classique n'accorde le statut de terme qu'à des unités nominales. Cabré (2003) et L'Homme (2005) voudraient étendre ce privilège à des unités appartenant à d'autres parties du discours, tel que des verbes et des adjectifs et des adverbes. Ce besoin d'élargir la notion de terme laisse entrevoir l'existence de plusieurs phénomènes au statut terminologique flou, à mi-chemin entre la terminologie et la phraséologie. Surtout, il révèle les profonds désaccords qui subsistent entre les différents courants sur ce qu'est un terme.

¹⁶ Nous rappelons que, pour Depecker, il s'agit d'une « synonymie référentielle » et qu'il prône la distinction entre signifié et notion. Cette appellation est toutefois problématique, car elle perpétue plutôt la confusion.

À l'origine des critiques se trouve l'évolution dans les méthodes du travail terminographique, opérée surtout grâce à l'application des techniques de la linguistique de corpus ainsi qu'au développement des technologies d'extraction semi-automatique de termes (Slodzian 2000). En effet, c'est lors de l'observation des termes dans leur contexte linguistique que l'on se heurte aux limites de l'approche strictement onomasiologique. Comme le signale Slodzian (2000), l'étude des textes met au jour l'importance de la polysémie et de la variation terminologiques. Il devient alors impossible de nier le rôle de l'interprétation contextuelle en terminologie. De plus, l'opposition entre le mot et le terme s'estompe à partir de la présence d'éléments phraséologiques et de ce qu'elle appelle « *item tiers* », apparenté au « *vocabulaire général d'orientation scientifique* » de Phal (1971).

La réflexion suscitée par les nouvelles pratiques terminologiques bat encore son plein. Elle a entraîné un éclatement de la notion de terme, qui a assurément gagné à être nuancée, mais dont on n'a pas fourni, jusqu'à maintenant, une définition précise. Très souvent se pose encore la difficulté pour distinguer les termes d'autres éléments linguistiques plus ou moins figés : termes « médians », collocations, éléments phraséologiques, syntagmes libres. Pourtant, comme le rappelle Slodzian (2000), les expressions telles qu'elles se trouvent dans les textes sont la seule réalité concrète analysable, la seule clé pour accéder aux notions spécialisées. Voici le point de départ de notre recherche : nous nous intéressons à toutes les formes servant à la mise en discours des notions. Ainsi, nous ne chercherons pas dans notre travail à identifier les unités terminologiques. Nous ne ferons pas appel à la notion de terme en raison de son indétermination et nous nous servirons d'autres principes afin de décrire les expressions relevées.

2.5. Les notions spécialisées en discours

En discours, la marge de manœuvre du locuteur s'élargit considérablement. Il compte sur le contexte, sur la situation de communication et sur l'interlocuteur pour *faire sens*. Les notions spécialisées peuvent donc être exprimées de façon très diverse. Par exemple, dans un texte la même notion peut être rendue par **clonage humain, obtenir un embryon humain par clonage, créer un embryon** ou **fabrication d'un embryon de clone humain**, et, dans un autre texte, on peut y faire référence par **la question**. Nous considérons le groupe d'expressions employées pour rendre une même notion comme un ***ensemble paraphrastique***. Cette notion se fonde sur la « *conception énonciative* » de la ***paraphrase*** développée par Fuchs (1982). Elle définit cette relation

« *entre deux séquences linguistiques X et Y comme un jugement de nature métalinguistique porté par un sujet parlant sur ces deux séquences - jugement par*

lequel il identifie le sémantisme qu'il attribue à X et le sémantisme qu'il attribue à Y »¹⁷ (Fuchs 1982, p. 116).

Ce jugement n'est possible qu'en « gommant » les différences que comportent nécessairement ces séquences par l'entremise d'une « réduction par rapport à la complexité des relations entre les séquences en langue » (Fuchs 1982, p. 177). Fuchs souligne également que l'identification est « effectuée par un sujet en situation »; il ne s'agit pas « d'une relation d'identité donnée a priori et une fois pour toutes » (Fuchs 1982, p. 116)¹⁸. Le recours à la paraphrase témoigne ainsi d'un vouloir dire et se révèle comme une stratégie du locuteur au service d'une intention communicative. Dans notre travail, nous nous penchons spécifiquement sur la **relation de coréférence** qui fonde l'identification des expressions renvoyant à une même notion spécialisée.

Dans le cadre particulier de la communication spécialisée, les liens paraphrastiques ont été étudiés par Mortureux (1993)¹⁹, qui propose de les décrire sur deux plans. Le premier consiste à examiner au sein des textes la présence ou l'absence de marques linguistiques explicitant la relation, ce qui détermine des **paradigmes définitionnels** - lorsqu'elle est indiquée - et des **paradigmes désignationnels** - lorsqu'elle ne l'est pas. Les marqueurs peuvent être des verbes métalinguistiques comme **appeler** (dans **un ovule fécondé est appelé fœtus à 14 jours; quand apparaît l'ébauche du système nerveux**), des marques de coordination métalinguistique comme **ou** (dans **couronne ou cumulus cellulaire**), des moyens typographiques (dans **l'objet de leur quête : des cellules souches pluripotentes**) ou des marques de reprise anaphorique comme les adjectifs démonstratifs (**ces cellules**). Il faut souligner que les paradigmes définitionnels ne comportent pas seulement des définitions, tel que l'on peut le remarquer dans les exemples; il s'agit de l'ensemble d'expressions dont les liens sont marqués linguistiquement.

Le deuxième niveau d'analyse concerne les divers types de relations existantes entre les expressions, qui peuvent s'établir par l'intermédiaire de leur signification linguistique, comme les synonymes et les hyperonymes, ou en discours, comme les expressions métaphoriques. Mortureux (1993) indique que, dans ce dernier cas, la paraphrase peut servir à introduire une **appréciation** (« qualification ») du locuteur sur la notion en question. Les liens peuvent être identifiés grâce à des marques linguistiques (les marqueurs déterminant les paradigmes définitionnels) ou à un effort d'interprétation du lecteur (constituant ainsi des paradigmes désignationnels).

¹⁷ Mise en relief de l'auteur.

¹⁸ Ainsi, la paraphrase s'inscrit dans la tension entre le *même* et l'*autre* que les traducteurs connaissons si bien.

¹⁹ Mortureux se penche sur les « reformulants pour un vocable donné ». Nous préférons parler d'ensembles paraphrastiques, car nous ne voulons pas mettre l'accent sur l'orientation de la relation entre les paraphrases, mais plutôt sur le fait qu'elles font référence à une même notion.

Comme nous l'avons mentionné plus haut, pour caractériser le **lien de référence** existant entre les expressions et la notion spécialisée à laquelle elles renvoient, nous avons décidé d'écartier la division entre termes et mots, car elle est problématique et ne semble pas opératoire dans le cadre de notre travail. Les rapports de **dénomination** et de **désignation**²⁰ décrits par Kleiber (1984) pourraient s'avérer plus utiles dans l'analyse de l'expression des notions en discours. Tous les deux servent à « *met[tre] en rapport les signes avec les choses* », mais le premier « *consiste en l'institution entre un objet et un signe X d'une association référentielle durable* » en tant que « *résultat d'un acte de dénomination effectif ou celui d'une habitude associative* » (Kleiber 1984). Pour Kleiber (1984), la désignation est « *uniquement momentanée, transitoire et contingente* », tandis que la dénomination établit « *une règle de fixation référentielle qui permet l'utilisation ultérieure du nom pour l'objet dénommé* ». En effet, si pour désigner on n'a pas besoin d'un codage préalable, les dénominations sont des « *unités codées* » pouvant faire l'objet d'un apprentissage (Kleiber 1984).

Les relations de dénomination et de désignation sont donc étroitement reliées à l'usage, puisqu'elles se fondent sur le caractère figé des expressions. Kleiber (1984) laisse entrevoir qu'il s'agit d'un processus graduel, surtout dans le cas des syntagmes qui seraient « *en voie de lexicalisation* ». D'autres chercheurs préfèrent parler de **degré de figement** afin de mettre en évidence « *le continuum entre constructions syntaxiques "libres" et expressions idiomatiques* » (François et Mejri 2006, p. 7). À notre avis, cette notion serait mieux adaptée que celle de terme à la caractérisation des discours spécialisés, et tout particulièrement de la vulgarisation. Cependant, la mesure du degré de figement au sein de la discipline scientifique des unités et syntagmes qui y sont employés exige la réalisation d'une étude comparative.

Le choix des expressions et l'emploi des divers marqueurs est au service des **stratégies discursives** mises en place par l'auteur du texte pour la diffusion des savoirs. Dans les discours de vulgarisation, on a relevé de nombreux procédés visant à faire le lien entre les nouvelles connaissances et celles que le lecteur possède déjà, dont la dénomination, la définition, la description, la reformulation ou paraphrase, la narration, la modalisation et les constructions analogiques et métaphoriques (plusieurs travaux sont cités dans Calsamiglia et Van Dijk 2004). Qui plus est, plusieurs chercheurs considèrent les stratégies discursives comme un trait caractéristique de la vulgarisation (Cassany *et al.* 2000, Calsamiglia et Van Dijk 2004, Ciapuscio 2003a). Notre recherche cible cinq mécanismes qui servent à cerner les notions et à structurer les connaissances : la **diaphore** (anaphore et cataphore), la **reformulation**, la **définition**, l'**explication** et la **métaphore**.

²⁰ À ne pas confondre avec la notion de *paradigme désignationnel* que nous venons de présenter, qui rend compte du rapport paraphrastique non marqué linguistiquement.

Nous reprenons la dénomination générique de « diaphore » introduite par Maillard (1974) pour désigner la reprise anaphorique et cataphorique²¹. Nous la concevons à partir de la définition d'anaphore proposée par Kleiber, « *comme une expression dont l'interprétation référentielle dépend d'une autre expression (ou d'autres expressions) mentionnée dans le texte et généralement appelée son antécédent* » (Kleiber 1993, p. 22, cité dans Charaudeau et Maingueneau 2002). La diaphore permet d'exprimer une notion différemment sans pour autant postuler une identité sémantique entre les unités ou syntagmes employés, mais plutôt un lien coréférentiel; de ce fait, elle assure principalement la progression thématique. Elle peut être grammaticale, lorsque l'antécédent est repris par un pronom, par un adjectif possessif ou par un syntagme comportant une marque linguistique du lien diaphorique (article défini, adjectif démonstratif ou possessif, etc.), ou lexicale, dans les cas de répétition de l'antécédent et de reprise par un synonyme ou par un quasi-synonyme (hyperonyme, hyponyme, etc.).

La reformulation est le plus souvent entendue au sens large de paraphrase ou même comme une notion générique²², mais nous préférons la circonscrire aux cas où le locuteur postule une relation d'identité pleine entre deux formes linguistiques (ou plus). Dans ce sens, elle remplit surtout une fonction métalinguistique. Cette fonction se retrouve également dans la définition, que Sager envisage comme « *a linguistic description of a concept, based on the listing of a number of characteristics, which conveys the meaning of the concept* »²³ (Sager 2000, p. 39). Ayant distingué la notion du signifié, nous dirons qu'il s'agit plutôt d'une description linguistique de traits conceptuels. La définition est liée à une autre expression faisant référence à une notion.

À la différence des stratégies décrites précédemment, l'explication ne met pas en relation des expressions renvoyant à une notion spécialisée, mais permet de relier l'une d'entre elles à une autre expression mettant en relief un aspect de la notion. Par conséquent, si la définition peut, dans certaines circonstances, se substituer à l'unité ou syntagme définis, cela est impossible pour l'explication. En effet, il ne s'agit pas d'établir un lien coréférentiel entre deux expressions, mais d'introduire une nouvelle information à propos de la notion.

Enfin, la métaphore est souvent conçue comme une stratégie employée dans la vulgarisation pour « *"éclaircir" des concepts, pour les rendre plus "accessibles"* » (Loffler-

²¹ « Il serait utile de disposer d'un vocable qui transcende l'opposition anaphore/cataphore et désigne la référence contextuelle en général : le terme de DIAPHORE, qui n'est pas vectoriellement orientée, pourrait convenir assez bien » (Maillard, M. [1974], « Essai de typologie des substituts diaphoriques », *Langue française*, 21, p. 55-71).

²² Ainsi, certains auteurs distinguent la reformulation paraphrastique de la reformulation non paraphrastique (voir principalement Gülich, E. et T. Kotschi (1983). « Les marqueurs de la reformulation paraphrastique », *Cahiers de linguistique française*, 5, p. 305-351.).

²³ [Notre traduction :] « *une description linguistique d'un concept, comportant l'énumération d'un certain nombre de caractéristiques, qui véhicule la signification de celui-ci* ».

Laurian 1994). Il s'agit pourtant d'un mécanisme conceptuel complexe liée au processus même de la cognition. Nous ne le classons donc pas parmi les procédés discursifs, même si parfois il peut servir comme tel. Nous préférons parler de **conceptualisation métaphorique**, ce qui mérite une description plus détaillée.

2.6. Modes de conceptualisation

La **conceptualisation métaphorique** a été décrite du point de vue de la linguistique cognitive par Lakoff comme un mécanisme qui consiste dans « *understanding and experiencing one kind of thing in terms of another* »²⁴ (Lakoff et Johnson 1980/2003, p. 5). Loin d'être exceptionnel, il s'agit d'un processus de raisonnement fréquemment à l'œuvre, autant dans la vie de tous les jours que dans la pensée scientifique. Il est donc nécessaire de faire la distinction entre les **expressions métaphoriques**, effectivement employées dans le discours, et la **métaphore conceptuelle**, qui opère au niveau cognitif par la projection d'un cadre conceptuel source sur un autre cadre cible. L'exemple classique est celui de la conceptualisation de la discussion comme une guerre (LA DISCUSSION EST LA GUERRE²⁵) qui se manifeste dans des expressions comme celles-ci :

(5) « *Il a démoli tous mes arguments.* »

(6) « *Ta position est indéfendable.* »

Pour sa part, Talmy (2001) parle de **factivité** et **fictivité** afin de rendre compte de l'existence de deux représentations simultanées de la même situation référentielle : on sait que la discussion n'est pas vraiment la guerre (représentation factive), mais la représentation de la guerre sert à *penser* la situation de discussion et à en *parler* (représentation fictive). La notion d'**intégration conceptuelle** développée par Fauconnier et Turner (1998/2002) met en relief le fait que, pour que l'expression métaphorique prenne un sens, les deux représentations doivent s'intégrer de façon cohérente. Ce qui revient à dire qu'il existe des restrictions quant aux cadres conceptuels source pouvant se projeter sur un cadre cible donné : ce dernier sélectionne certains **modes de conceptualisation métaphorique** (Vandaele *et al.* 2006, Vandaele et Lubin 2005).

Les différents modes de conceptualisation interviennent dans l'élaboration et la structuration de connaissances scientifiques (par exemple, Fox Keller 1999, Kordon 1991) et laissent leur trace dans les discours, ce qui est particulièrement intéressant pour la traduction (Temmerman 2002a, Vandaele *et al.* 2006, Vandaele et Raffo à paraître, dans le

²⁴ [Notre traduction :] « *de comprendre et d'expérimenter une sorte de choses comme si c'était une autre* ».

²⁵ Nous reprenons la convention typographique proposée par Lakoff et indiquons les métaphores conceptuelles en petites majuscules.

domaine biomédical). Afin de les caractériser, Vandaele analyse les **indices de conceptualisation**, « *tout élément d'un texte témoignant de la réalisation linguistique d'une métaphore conceptuelle* » (Vandaele 2005, p. 279). Les analyses montrent qu'il s'agit souvent d'unités prédicatives évoquant un cadre conceptuel différent de celui du texte; la projection du premier sur le second a lieu par l'intermédiaire des actants du prédicat (la classe actantielle typique du cadre source est projetée sur les actants réalisés dans le texte) (Vandaele et Lubin 2005). Ainsi, par exemple, dans la phrase

(7) « *Les cellules des crêtes neurales [...] migrent vers la région ventrale de l'embryon où elles colonisent différents territoires [...].* »²⁶

on peut voir que le verbe « migrer » renvoie à un cadre conceptuel distinct et que la classe des êtres vivants (humains, animaux) se projette sur les cellules. La méthode automatique étant exclue en raison du degré de connaissance extralinguistique nécessaire à l'analyse, le repérage de ces indices s'effectue en ayant recours à la **dissonance cognitive**, « *qui résulte de la comparaison mentale entre l'expression relevée et une expression proche exprimant le cadre conceptuel source* » (Vandaele et Lubin 2005).

Les indices de conceptualisation forment des réseaux lexicaux qui renforcent un mode de conceptualisation en particulier (Vandaele et al. 2006). En effet, l'occurrence dans les textes d'indices appartenant habituellement à la classe d'êtres vivants (dans l'exemple, « région », « colonisent » et « territoires ») rend compte de l'importance de ce mode de conceptualisation des cellules dans le domaine. Ces réseaux de conceptualisation métaphorique sont donc étroitement liés à l'idiomaticité de l'expression. Vandaele souligne que « *la phraséologie des langues de spécialité est largement gouvernée par les modes de conceptualisation sous-jacents, et plus précisément par la conceptualisation métaphorique* » (Vandaele et Lubin 2005, p. 419).

Une meilleure connaissance des cadres cognitifs constitutifs d'un domaine et les contraintes qu'ils exercent au plan linguistique - autant sur les termes que sur la phraséologie - constitue un puissant outil pour la traduction. Elle permet d'intégrer plus facilement de nouveaux savoirs et de mieux fonder les décisions traductionnelles, d'autant plus que les expressions linguistiques équivalentes de deux ou plusieurs langues ne relèvent pas nécessairement d'une même conceptualisation. De plus, les phénomènes d'abus de la traduction littérale et de l'interférence linguistique apparaissent sous une lumière nouvelle. Cette approche fera donc partie de notre boîte à outils pour la réalisation de cette recherche.

²⁶ Exemple emprunté à Vandaele, S. et al. (2006). « La conceptualisation métaphorique en biomédecine : indices de conceptualisation et réseaux lexicaux », *Glottopol*, 8, p. 73-94.

3. État de la question

La façon dont les notions spécialisées sont exprimées dans les textes a été l'objet d'un bon nombre de recherches sous divers angles. Dans ce chapitre, nous allons passer en revue quelques-unes qui, à notre connaissance, comportent des aspects pertinents en rapport avec notre propre recherche, même si elles ont été menées avec des objectifs différents. En effet, réalisés dans le cadre de la terminologie, de la linguistique textuelle, de l'analyse du discours et de la traduction, ces travaux s'avèrent intéressants dans la mesure où ils prennent en compte des variables que nous étudions. Ils permettent également de situer l'originalité de notre contribution. Bien que cela puisse paraître étonnant, les études traitant des différences sémantiques et morphosyntaxiques entre les langues ne semblent pas compatibles avec notre travail (Carreño 2004, Suárez 2004). Nous trouvons plus de points de contact avec les recherches qui explorent l'influence du degré de spécialisation du texte dans l'expression linguistique (Freixa 2002, Ciapuscio 2003a) et celles qui mettent en relief la dimension sociale de la vulgarisation (Cassany et Martí 1998, Cataldi 2004). De plus, nous considérons plusieurs travaux de Vandaele (travaux divers), qui examine les modes de conceptualisation sous-jacents au discours scientifique, et un article de Knudsen (2003) où elle se penche sur les métaphores de la vulgarisation.

3.1. Les langues

Iveth Carreño et María Mercedes Suárez se sont penchées sur la variation terminologique dans un contexte de traduction de l'anglais vers l'espagnol, la première dans une perspective de repérage automatique et la deuxième dans un cadre comparatif. À partir d'un corpus parallèle de textes spécialisés du domaine de l'environnement (23 textes originaux en anglais avec leur traduction en espagnol), Carreño (2004) a analysé les variantes proposées dans les textes espagnols comme équivalents de 50 termes anglais (25 termes simples et 25 termes complexes). Elle relève divers types de variation, dont certains, comme l'occurrence du terme anglais dans le texte traduit et l'omission, sont exclusifs ou caractéristiques de la situation de traduction - ce qui semble évident. Son travail ne se penche ni sur les motivations ni sur les conséquences de la variation; il se borne à une description formelle en vue du traitement automatique.

Pour sa part, Suárez (2004) a choisi d'étudier un corpus d'articles sur le génome humain parus dans la revue *Scientific American* et traduits pour être publiés dans sa version

espagnole, *Investigación y Ciencia* (34 textes originaux avec leur traduction). À la différence de Carreño, elle analyse les variantes dans les deux langues. Cependant, seule est prise en compte la variation dénomminative « explicite », exprimée au moyen de marqueurs discursifs tels que *called, known as, named, or, referred to (as), termed (as), viewed as* ainsi que les parenthèses. Suárez étudie les rapports qui s'établissent entre les variantes faisant intervenir une analyse sémantique componentielle. Ce type d'analyse, d'une part, ne convient pas toujours à une analyse de syntagmes terminologiques, dont le « sens » n'est pas nécessairement componentiel et, d'autre part, a le désavantage de mettre l'accent sur la signification linguistique au détriment des phénomènes discursifs. L'aspect de cette recherche qui est le plus intéressant pour nous est la mise en évidence des changements entraînés dans la relation sémantique entre les variantes, mais aussi des différences au niveau textuel (cohésion, progression thématique), cognitif (point de vue) et communicatif (niveau de spécialisation) vis-à-vis du texte original.

Ces deux travaux attirent l'attention sur les diverses solutions de traduction des termes. En revanche, observer des textes traduits ne permet pas de tirer de conclusions assurées quant au comportement des locuteurs natifs. Il est nécessaire de compter sur des études sur corpus unilingues ou comparables afin de cerner les caractéristiques propres à une langue.

3.2. Degré de spécialité

Freixa et Ciapuscio ont mené des travaux comparant des textes spécialisés à des textes s'adressant au grand public. Se situant dans la lignée de la théorie communicative prônée par Cabré (1999, entre autres), Freixa (2002) offre dans sa thèse de doctorat, *La variació terminològica*, une réflexion théorique et un travail empirique sur la variation dénomminative au sein de textes de différent degré de spécialisation. Elle part « *de la base que un mateix emissor, en aquest cas un especialista, pot expressar un coneixement especialitzat de maneres diferents segons els coneixements que pressuposa en el receptor* »²⁷ (p. 135). Les corpus, constitués de textes du domaine du traitement de déchets et de la pollution, représentent les pôles de l'échelle de spécialisation établie en fonction des émetteurs et des récepteurs des textes (à partir de la typologie de Loffler-Laurian 1983). Il s'agit d'actes de communications dont l'auteur et les destinataires sont des experts et de brochures et publications destinées au grand public. La sélection des

²⁷ [Notre traduction :] « *du principe qu'un même émetteur, dans ce cas, un spécialiste, peut exprimer des connaissances spécialisées de façon différente en fonction des savoirs qu'il suppose chez le récepteur* ».

séquences linguistiques à analyser repose sur deux critères : d'une part, la pertinence thématique des unités et des syntagmes faisant référence à des notions du domaine, d'autre part, des contraintes morphosyntaxiques ont été introduites dans le but de limiter l'analyse aux structures considérées « typiques » des unités terminologiques. Ces critères n'établissent pas cependant une distinction nette entre ce qu'on entend habituellement par dénomination (« *unitats més fixades, lexicalitzades i normalment més sintètiques* »²⁸) et d'autres « unités de discours » (désignation, paraphrase). En effet, Freixa reconnaît avoir adopté une conception plus vaste de la dénomination en raison des difficultés pour les discriminer en discours.

Les résultats obtenus montrent une grande dispersion, malgré une majorité de notions exprimées par 1 ou 2 dénominations : beaucoup de notions sont « *unidénominatives* » tandis que beaucoup d'autres sont « *polydénominatives* » (jusqu'à 14 dénominations dans le corpus de vulgarisation et jusqu'à 12 dans le corpus scientifique). Freixa ne trouve pas de corrélation entre le nombre d'occurrences (nos « actualisations ») de la notion et la variation. Pour mesurer la variation dans ses corpus, Freixa fait appel à ce qu'elle appelle « *indice de dénomination* », mais qui n'est en réalité que la moyenne (nombre de dénominations par notion), et conclut à un taux de variation plus bas des communications scientifiques par rapport aux textes de vulgarisation (1,58 contre 2,02). Pourtant, ce calcul ne permet pas de rendre compte des résultats, qui ne présentent pas une distribution normale. Ce biais a un grand impact sur les conclusions que l'on peut tirer de cette étude. De plus, les notions figurant dans les deux corpus étudiés ne coïncidaient que partiellement. Il n'a donc été possible de comparer la variation entre les textes scientifiques et les textes de vulgarisation que sur 120 notions représentant 26,54 % du total. Freixa admet que « *qualsevol conclusió seria més determinant si els dos corpus terminològics hagessin estat nocionalment més coincidents* »²⁹ (p. 227). Dans ce cas, il aurait été souhaitable de connaître le nombre d'actualisations de ces notions afin de mesurer leur représentativité dans les corpus. Elle constate que seules 42 (35 %) notions étaient exprimées de la même façon dans les deux corpus tandis que 78 (65 %) présentaient une coïncidence partielle ou des différences dans la dénomination.

En ce qui concerne la partie du discours à laquelle appartiennent les dénominations, il n'est pas surprenant de découvrir une nette prédominance des noms (unités « monolexicales », noms suivis d'un syntagme prépositionnel et noms suivis d'adjectif notamment) par rapport aux verbes et aux adjectifs. Cette régularité se retrouve dans les

²⁸ [Notre traduction :] « *unités plutôt fixées, lexicalisées et généralement plus synthétiques* ».

²⁹ [Notre traduction :] « *il aurait été possible d'arriver à des conclusions plus probantes s'il y avait eu une coïncidence notionnelle plus grande entre les deux corpus* ».

groupes de dénominations exprimant une même notion. Il y a lieu de s'interroger sur le rôle joué par le deuxième critère de repérage des dénominations (la structure morphosyntaxique) dans l'obtention de ces résultats. La variation se vérifie surtout parmi les expressions syntagmatiques, qui subissent des réductions, la siglaison ou des modifications de la base ou des expansions. Certaines variantes ne sont équivalentes qu'en contexte : les hyperonymes, hyponymes, méronymes et cohyponymes. Si dans les textes plus spécialisés prédomine la réduction, les modifications sur le plan lexical sont plus fréquentes dans la vulgarisation. Par ailleurs, Freixa élabore une typologie des variations qui complète les classifications existantes et montre que les types de variations ne sont pas présents dans la même proportion dans les deux corpus.

Dans le cadre de la linguistique textuelle, Guiomar Ciapuscio propose d'aborder l'étude du discours spécialisé dans plusieurs de ses travaux (Ciapuscio 1993, 1999, 2003a). Plus particulièrement, son livre *Textos especializados y terminología* (Ciapuscio 2003a) effectue un rapprochement entre la linguistique textuelle d'origine allemande (par ex., de Beaugrande et Dressler 1981, Heinemann et Viehweger 1991), l'approche communicative de la terminologie (notamment, Cabré 1999) et la lexicologie verticale (Wichter 1994). L'auteure articule les postulats de ces cadres théoriques complémentaires et emprunte leurs outils méthodologiques pour analyser des textes dont la situation communicative varie en fonction des participants. Elle tient compte de la communication entre spécialistes (articles scientifiques et encyclopédies spécialisés), entre spécialistes ou journalistes spécialisés et lecteurs initiés (articles de semi-vulgarisation) et entre journalistes et grand public (articles de presse). Pourtant, ses corpus ne sont constitués que d'un texte de chaque classe considérée; son étude est menée à titre « exemplaire » : elle cherche surtout à étayer une position théorique et à valider une démarche d'analyse. Par l'observation du fonctionnement des termes en contexte, elle tente d'établir des fondements linguistiques pour déterminer le degré de spécialisation des textes, son but ultime étant de constituer une typologie.

Dans un premier temps, Ciapuscio analyse le contenu thématique de trois textes : une entrée de dictionnaire spécialisé (*Diccionario de Ciencias Médicas* de Stedman), un texte de semi-vulgarisation (revue *Scientific American*) et un article de vulgarisation (paru dans plusieurs quotidiens argentins). Nous devons ici soulever le fait étonnant qu'elle utilise pour son analyse un texte traduit par elle-même (celui de *Scientific American*). Ciapuscio constate une nette différence entre le contenu de la description de l'encyclopédie spécialisée et celui des articles de semi-vulgarisation et de presse, ce qui peut être corrélé à une divergence des fonctions (informer, convaincre) et des situations de communication. De plus, l'analyse montre dans l'article de vulgarisation une diminution de la précision conceptuelle qui se manifeste dans la terminologie et le traitement paraphrastique de

l'information (páncreas au lieu de células isletas del páncreas). Cette section met certainement en évidence une relation entre la classe du texte et son contenu. Par contre, la variation référentielle des dénominations et la variation terminologique sont peu détaillées. Si le modèle de la lexicologie verticale permet de représenter comparativement le contenu notionnel des textes, il semble offrir peu d'outils de description linguistique.

Dans le chapitre suivant, l'auteure cherche les traces d'un ensemble de notions dans le résumé et dans la conclusion d'un article scientifique, dans une entrevue réalisée à un des chercheurs et publiée dans une revue destinée à un public universitaire, et dans deux articles parus dans deux journaux argentins. Elle se sert du modèle cognitif-procédural de De Beaugrande et Dressler (1981) pour identifier les noyaux conceptuels figurant dans le résumé (« *abstract* ») et les rapports existants entre eux. Puis, elle décrit les changements d'ordre thématique qui se produisent au fil des autres textes et compare les réalisations linguistiques des noyaux conceptuels. L'analyse du choix, de la variation et du traitement paraphrastique ou non des termes selon le type textuel fournit la preuve d'une relation entre la « *surface* » linguistique, la fonction, la situation communicative et le degré de spécialisation d'un texte. Ainsi, dans les articles de vulgarisation, en plus de la simplification, voire du détournement conceptuel, elle voit la terminologie spécifique du résumé scientifique disparaître au profit de termes « banalisés », de paraphrases et de séquences descriptives. Les données encyclopédiques, l'explicitation des rapports entre certains concepts et la simplification de la phraséologie témoignent d'une volonté d'offrir un ancrage notionnel au lecteur. Au cours de ses observations, l'auteure effleure les sujets du changement de point de vue sous lequel le concept est abordé et des frontières floues entre la terminologie spécialisée et le lexique dit général.

Enfin, Ciapuscio propose une typologie de textes spécialisés prenant en compte plusieurs niveaux : fonctionnel, situationnel, sémantique et formel-grammatical. Nous ne nous attarderons pas ici à la décrire, car les tentatives formelles de classement s'éloignent des objectifs du présent travail. Soulignons seulement qu'elle ne conçoit pas de limites étanches entre les diverses classes de textes, mais les situe sur un axe vertical en fonction de leur degré de spécialisation. Malgré la taille réduite des corpus et une analyse linguistique peu exhaustive, ce travail démontre la pertinence de l'approche textuelle : autant l'étude de la terminologie et de la variation conceptuelle permet de mieux cerner la classe textuelle, autant il est utile de prendre en compte l'impact de la fonction et de la situation communicative, niveaux supérieurs de caractérisation du texte, sur le contenu thématique et les structures linguistiques de surface.

3.3. *La vulgarisation*

L'analyse du discours s'intéresse aux rapports existants entre la mise en discours des notions spécialisées (depuis l'organisation textuelle jusqu'aux unités lexicales employées) et les circonstances de production et de réception des textes (notamment les intentions communicatives des locuteurs). Les analyses menées sous cet angle sont non seulement nombreuses, mais extrêmement diverses quant aux aspects étudiés et aux méthodologies employées. Ainsi, par exemple, Méndez Cendón et López Arroyo (2003) identifient des marqueurs rhétoriques et phraséologiques permettant de rendre compte de la structuration de l'information dans des articles scientifiques et dans leurs résumés. Moirand (2003) étudie le phénomène de la polyphonie au sein de textes journalistiques à partir de sa conception de « *nouvelle vulgarisation* »³⁰. Dans le cadre d'un projet plus vaste visant à bâtir une théorie du contexte, Van Dijk (2003) explore les liens entre savoir spécialisé et discours mettant en évidence la structuration de la présentation des notions dans un manuel de génétique, sans pourtant offrir une description linguistique approfondie.

Nous centrerons notre intérêt sur quatre travaux pertinents au regard de notre étude. Parmi les plus intéressants se trouve celui de Cassany et ses collaborateurs (2000), qui présentent un modèle théorique de la vulgarisation permettant de rendre compte des transformations des réseaux conceptuels spécialisés aux plans conceptuel, textuel et discursif. Ce groupe de travail entend la vulgarisation comme une réélaboration du matériel à vulgariser avec d'autres ressources linguistiques. Leur but est donc d'identifier et d'analyser les tâches cognitives et discursives réalisées par l'auteur-vulgarisateur ainsi que d'étudier les moyens linguistiques employés pour faire comprendre les notions spécialisées aux non initiés. Ils illustrent le modèle avec des exemples tirés d'un corpus de 27 textes parus dans la presse espagnole traitant du concept « transgénique ». Nous croyons que le modèle théorique élaboré comporte un pouvoir heuristique certain, malgré le flou méthodologique, concernant notamment le repérage et l'analyse des séquences linguistiques considérées, et l'attribution aux différents discours (spécialisés et non spécialisés) de traits caractéristiques, qui devrait s'appuyer sur une justification plus étoffée.

Calsamiglia et Van Dijk (2004) prônent une étude de la vulgarisation dans un cadre multidisciplinaire réunissant la linguistique, l'analyse critique du discours et les sciences

³⁰ Elle distingue le modèle traditionnel de la vulgarisation, où prime la transmission de connaissances, d'un nouveau modèle axé sur la rencontre dialogique de la communauté scientifique et des autres acteurs sociaux concernés.

cognitives. Dans cet article, ils présentent les éléments théoriques qui fondent leur vision de la vulgarisation en tant que pratique discursive sociale visant une recontextualisation d'un ensemble de savoirs. Puis, dans un corpus de 42 articles sur le génome humain publiés dans le journal espagnol *El País*, ils étudient la structuration des connaissances proposée et quatre « *structures explicatives* » : dénomination, métaphore, définition et description. Les mêmes réserves émises pour le travail de Cassany et ses collaborateurs quant à la méthodologie adoptée s'appliquent ici : la nature des données analysées et les critères pour leur repérage ne sont pas explicités, et il manque de définitions de notions clés, comme celle de dénomination.

Enfin, deux études ont ciblé plus particulièrement l'expression linguistique des notions. Cataldi (2004) relève les diverses expressions dénominatives employées pour nommer les organismes génétiquement modifiés et les techniques biotechnologiques dans les journaux espagnols entre 1999 et 2000. Ayant constaté que les textes sur le sujet reflètent un débat, ce qui donne lieu à un « *discurso mediático plurilogo* »³¹ (Moirand 2000, Reboul-Touré 2000), elle décrit les dénominations en regard des domaines dont elles sont issues (social et économique) et des locuteurs (professeur ou chercheur, politicien, écologiste, entreprise de biotechnologie) qui les utilisent. De cette façon, elle peut mettre en relation les formes choisies pour nommer les concepts avec des facteurs d'ordre scientifique, social, politique et économique constituant le contexte de la communication. Les résultats présentés mettent en évidence une « *intensa divulgación debate* »³² et permettent de conclure que « *los textos periodísticos, además de informar, transmiten un punto de vista que puede orientar la interpretación del público a partir de una perspectiva positiva o negativa respecto a las plantas, cultivos y alimentos transgénicos* »³³ (Cataldi, 2004, p. 67). Le cadre théorique dans lequel se situe cette étude permet de jeter une lumière originale sur certains aspects de la vulgarisation. En revanche, Cataldi ne fournit pas une description linguistique rigoureuse des expressions dénominatives relevées ni ne cerne clairement les concepts auxquels elles font référence (encore une fois la définition de dénomination fait défaut).

Cassany et Marti (1998) examinent les stratégies utilisées dans la presse espagnole pour vulgariser le concept de prion lors de l'écllosion de l'affaire de la « *vache folle* » en 1996. Son intérêt particulier réside dans le fait qu'il s'agit d'une notion complexe mais essentielle à l'explication de la maladie et, en même temps, à ce moment-là, d'un concept

³¹ [Notre traduction :] « *discours médiatique plurilogo* ».

³² [Notre traduction :] « *vulgarisation-débat intense* ». Mise en relief par l'auteur.

³³ [Notre traduction :] « *les textes journalistiques, en plus d'informer, transmettent un point de vue pouvant orienter l'interprétation du public à partir d'une perspective positive ou négative vis-à-vis des plantes, des cultures et des aliments génétiquement modifiés* ».

nouveau et controversé au sein de la communauté scientifique. D'ailleurs, la plupart des textes traitant du sujet ne présentent pas que des informations strictement scientifiques, mais s'occupent surtout des enjeux politiques et socioéconomiques liés à la maladie. Le corpus comprend des articles publiés en espagnol et en catalan, cependant, les éventuelles différences entre ces deux langues ne sont pas signalées. Les procédés analysés sont d'ordre lexical, discursif et rhétorique, à l'exclusion de la métaphore. Étonnamment, la première stratégie observée consiste à éviter la notion de prion et à offrir une explication vague des causes des maladies concernées. Par contre, lorsque la notion est exprimée, le terme *prión* ou des termes hyperonymes sont utilisés sans variantes. La variation dénomminative concerne plutôt des notions périphériques à celle de prion (formation, types et rapports entre les prions de diverses espèces et avec les maladies). Ici nous ne trouvons pas non plus une définition précise du concept de dénomination.

En outre, Cassany et Marti mettent en évidence plusieurs stratégies discursives liées à la sélection de l'information et sa présentation : la contextualisation, qui fournit au lecteur un cadre de référence avant d'introduire le concept, la présence de séquences narratives et de questions rhétoriques et la modalisation du discours visant à mettre en relief l'attitude de l'auteur et rapprocher ses propos du public. Selon les chercheurs, ces moyens discursifs sont caractéristiques de la vulgarisation, puisqu'ils n'appartiennent pas au registre technique-scientifique ni ne sont imposés par les genres journalistiques utilisés. Pourtant, cette conclusion ne découle pas des résultats présentés dans cette étude ou dans un autre travail qui y serait mentionné.

3.4. Les modes de conceptualisation métaphorique

Le recours à la métaphore dans les discours spécialisés a suscité dans les derniers temps un nombre considérable de travaux (Meyer *et al.* 1998, Temmerman 2002a, Koller 2004, Liakopoulos 2002, Loffler-Laurian 1994, Van Rijn-Van Tongeren 1997, Knudsen 2003, Vandaele *et al.* 2006). Nous allons cependant centrer la revue aux deux dernières auteures dont les travaux nous semblent être les plus pertinents pour notre analyse.

En mettant à profit les acquis des sciences cognitives³⁴ – notamment la théorie de la conceptualisation métaphorique développée par Lakoff (Lakoff et Johnson 1980/2003) –, Vandaele et ses collaborateurs³⁵ se penchent sur les modes de conceptualisation et leur représentation linguistique dans une perspective de traduction. À cette fin, plusieurs

³⁴ Les bases théoriques de cette approche sont traitées en détail à la section 2.4 du « Cadre théorique ».

³⁵ Leur réflexion s'étend sur plusieurs travaux (voir citations ponctuelles ci-dessous).

corpus comparables en anglais et en français ont été constitués en compilant des textes appartenant aux domaines de la biologie cellulaire et moléculaire, d'une part, et de l'anatomie, d'autre part. Il s'agit essentiellement d'articles scientifiques et d'ouvrages de référence, ce qui exclut totalement le discours de vulgarisation et presque complètement le registre didactique. La description du mécanisme métaphorique est réalisée sous l'angle cognitif (Lakoff et Johnson 1980/2003, Talmy 2001, Fauconnier et Turner 1998/2002) et linguistique faisant appel à l'analyse actancielle (Tesnière 1959/1965, Mel'čuk *et al.* 1995).

À partir de l'analyse de certaines métaphores conceptuelles concernant les cellules, les molécules et les protéines qui parcourent le discours biomédical (Vandaele 2000, 2002, 2003), Vandaele conclut que la conceptualisation métaphorique sous-tend largement la phraséologie, et même un grand nombre de dénominations dans les discours spécialisés (Vandaele 2004, Vandaele *et al.* 2006). Elle constate également que la relation entre le cadre cognitif source et le cadre cible s'établit par l'intermédiaire d'unités lexicales prédicatives, le lien métaphorique se nouant grâce à la projection de classes d'actants typiques de l'unité en question sur ceux apparaissant dans le corpus (Vandaele 2004, 2005). Par exemple, si les actants typiques du verbe *résider* sont une personne et un espace géographique mais que, dans les textes spécialisés, ce rôle est rempli par une molécule et une partie de la cellule, il est possible de postuler la métaphore UNE MOLÉCULE EST UN INDIVIDU HABITANT UN LIEU (Vandaele 2004, p. 625).

Ainsi, l'analyse met surtout l'accent sur les unités qui ne sont pas prises habituellement en compte par les dictionnaires et encyclopédies spécialisées (noms n'ayant pas le statut de terme, verbes, adjectifs, adverbes). Dans tous les cas, elles reflètent un mode de conceptualisation propre au domaine et constituent de ce fait des « *indices de conceptualisation métaphorique* » (Vandaele 2005). La mise en relief de ces unités dont le statut terminologique n'est pas clairement défini attire l'attention sur la difficulté de tracer les limites entre la langue générale et la langue de spécialité. Par ailleurs, cette étude éclaire plusieurs aspects de la traduction de textes spécialisés, tels que le phénomène de l'interférence linguistique et la notion d'équivalence.

Dans une étude ultérieure, Vandaele et Raffo (2007) présentent une comparaison de réseaux d'indices de conceptualisation métaphorique relevés dans des textes scientifiques et des textes de vulgarisation³⁶, et mettent en évidence une continuité entre les deux. En effet, certains indices employés dans le discours scientifique, se retrouvent également dans des textes moins spécialisés, comme le verbe reparer, qui véhicule le mode de conceptualisation des organes et des tissus comme des machines. D'autres indices repérés

³⁶ Cette comparaison est en partie fondée sur des données tirées des corpus compilés pour ce mémoire.

dans des textes de vulgarisation prolongent un réseau lexical présent dans le corpus scientifique. C'est le cas de fabricar, hyponyme d'un indice courant dans les textes scientifiques (producir) et de collar, indice compatible avec la conceptualisation des aminoacides comme une chaîne (cadena polipeptídica). On trouve également dans la vulgarisation des indices servant à remplacer des dénominations du domaine compatibles avec ceux des textes scientifiques (les noms palabra et letra pour faire référence aux codons et aux nucléotides sont cohérents avec la métaphore bien connue de la traduction du code génétique). Il peut s'agir même de dérivés (decodificador pour ribosoma). Parfois, pourtant, on introduit un indice témoignant d'un nouveau mode de conceptualisation, comme receta, qui projette le cadre conceptuel de la cuisine sur l'activité scientifique. En plus des expressions métaphoriques, Vandaele et Raffo signalent la présence dans les textes de vulgarisation de constructions analogiques dont les indices peuvent être compatibles ou non avec les réseaux identifiés dans le discours scientifique.

Pour sa part, Knudsen (2003) examine la distinction réalisée par Boyd (1993) entre les métaphores ³⁷ « *constitutives* » des théories scientifiques et les métaphores « *pédagogiques* » ou « *exégétiques* ». Dans ce dessein, elle étudie la manifestation linguistique de la métaphore du code génétique dans trois articles scientifiques parus dans *Science* et trois textes de semi-vulgarisation publiés dans *Scientific American*. Ayant constaté la présence de ce mode de conceptualisation dans les deux corpus, elle montre que cette classification n'est pas toujours pertinente, les limites entre les deux étant plutôt floues :

« *The point is that any categorization of a metaphor has to rely on a pragmatic, diachronic analysis as well, because, as we have seen, theory-constructive metaphors can be used for pedagogical purposes, and—as will be suggested below—perhaps even the other way round* »³⁸ (Knudsen 2003, p. 1259)

Ainsi, elle postule que le but visé par la métaphore n'est pas propre à chaque mode de conceptualisation spécifique mais dépend du contexte dans lequel il se manifeste. Cette vision est compatible avec l'idée d'une continuité plutôt qu'une opposition entre les discours scientifique et de vulgarisation.

³⁷ Lorsque Knudsen parle de « métaphore », elle fait référence autant à ce que Vandaele, et nous même, appellons « conceptualisation métaphorique » qu'aux expressions métaphoriques, ce qui est confus.

³⁸ [Notre traduction :] « *Le point essentiel est que la catégorisation d'une métaphore doit être fondée sur une analyse pragmatique et diachronique, car, comme nous l'avons constaté, les métaphores constitutives peuvent servir des buts pédagogiques et - comme nous le suggérerons plus tard - inversement* ».

3.5. Récapitulation

Ces recherches nous offrent un horizon suffisamment vaste sur lequel situer notre propre travail. Les travaux menés dans le cadre de la terminologie fournissent des descriptions assez fines des éléments linguistiques repérés. Pourtant, ils mettent l'accent sur les dénominations et ce, d'un point de vue formel au détriment de la dimension discursive. La voie de l'étude des textes authentiques dans une perspective de traduction reste donc encore à explorer. Les approches textuelles et discursives ne décrivent pas en détail les phénomènes linguistiques. Ces dernières omettent parfois des précisions fondamentales quant à la démarche méthodologique et la terminologie adoptées. Les analyses sont donc difficilement reproductibles et comparables entre elles; l'ampleur des phénomènes observés reste difficile à cerner. Par exemple, il semble symptomatique qu'aucun des auteurs ne définisse la notion de dénomination, pourtant centrale dans leurs travaux. À partir des exemples présentés, il y a lieu même de se demander si l'on parle de la même notion. Ces courants ont toutefois le mérite de situer les phénomènes liés à l'expression des notions dans un cadre global permettant de faire le lien avec les circonstances de production et de réception des textes, notamment avec l'intention communicative du locuteur. Malheureusement, le fait que ces recherches soient réalisées sur une seule langue et la diversité de méthodes d'analyse ne permet pas de tirer des conclusions quant aux différences entre les langues. Enfin, l'étude des mécanismes cognitifs sous-jacents au discours spécialisé a révélé une architecture conceptuelle qu'il serait intéressant de mettre en perspective avec des aspects pragmatiques, comme le degré de spécialisation des textes.

4. Méthodologie

Pour décrire la représentation des notions spécialisées dans la vulgarisation des connaissances au grand public en fonction de la langue et du contexte socioculturel, il était nécessaire de se pencher sur des textes authentiques dans une perspective synchronique. Ainsi, nous avons décidé de réaliser une étude sur un corpus nous permettant d'observer l'interaction de ces variables : des articles journalistiques dans deux langues (français-espagnol) ayant été publiés dans des sociétés différentes (l'Espagne, l'Argentine, la France et le Québec). Pour ce travail, qui constitue une première étape d'un projet de recherche visant la caractérisation des discours spécialisés, nous avons mis au point une démarche méthodologique adaptée à notre objet. Elle devrait servir à éclairer des aspects nouveaux concernant les variables ciblées et à recueillir des données sur lesquelles fonder une étude plus exhaustive.

4.1. Corpus

Les corpus faisant l'objet de l'analyse sont constitués par des textes écrits se trouvant à l'une des extrémités de l'échelle de spécialisation : la vulgarisation pour des lecteurs non experts. Ce choix a été effectué à partir de la typologie présentée par Loffler-Laurian (1983), fondée uniquement sur des critères pragmatiques (le « niveau de compétence » de l'émetteur et du récepteur et le support du message). Il s'agit donc de deux corpus de textes de vulgarisation (le premier en langue espagnole et le deuxième en langue française) rédigés par des journalistes et ayant paru dans la presse destinée au grand public.

Notre choix des langues s'est arrêté sur le français et l'espagnol en raison surtout de l'absence de travaux comparatifs. En effet, la plupart des recherches se penchent sur une de ces langues ou incluent l'anglais. Si ce dernier est la *lingua franca* au sein de la communauté scientifique, la part du français et de l'espagnol dans la communication spécialisée reste toutefois significative. La francophonie et l'hispanophonie regroupent un grand nombre de locuteurs dans le monde, ce qui est à l'origine de nombreux circuits de diffusion de connaissances produisant des textes de différents niveaux de spécialisation. Certes, il s'agit de deux langues latines qui, de ce fait, ont un grand nombre de traits communs, mais cette parenté ne peut nous assurer que le comportement discursif et la conceptualisation des notions seront toujours les mêmes. Puis, dans le but de déceler une éventuelle influence du contexte socioculturel dans la présentation des connaissances

scientifiques, deux lieux de publication différents ont été choisis pour chaque langue : la France et le Québec dans le cas du français et l'Espagne et l'Argentine pour l'espagnol. Afin d'attester l'usage propre à chacune, le corpus a été constitué de textes comparables (non traduits).

Pour la sélection des textes de vulgarisation, nous avons fait appel à la notion de **moment discursif**, défini par Moirand comme « *the surge of intense and diversified media activity in connection with a single event* »³⁹ (Moirand 2003, p. 178). Il s'agit de la période, souvent réduite à quelques semaines, pendant laquelle un thème s'impose dans les médias à partir d'un événement ponctuel et devient un sujet d'actualité. Cet engouement de la presse est alors à l'origine de textes non seulement nombreux mais aussi de genres divers. Nous pouvons citer comme exemple l'ensemble d'articles, d'éditoriaux et de lettres des lecteurs autour du mariage et de l'homosexualité lors de la reconnaissance légale de l'union de conjoints de même sexe en juillet 2005. Adopter le moment discursif comme critère de constitution du corpus permet de déterminer des balises claires pour une analyse menée dans une perspective synchronique.

Ainsi, nous avons restreint la recherche à l'annonce, en février 2004, de l'obtention par clonage de cellules souches embryonnaires humaines par une équipe de recherche coréenne. Les textes en français ont été rassemblés à partir de la base de données *Bibliobranché*, qui réunit les principales ressources périodiques du Canada et de l'Europe francophone, tandis que les textes en espagnol ont été tirés des sites Internet de journaux argentins (*La Nación*, *Clarín* et *Página12*) et espagnols (*El País*, *ABC* et *El Mundo*). Nous n'avons tenu compte que des quotidiens distribués à l'échelle nationale, et provinciale dans le cas de Québec. La recherche des articles a été effectuée en fonction de la date (février 2004) et de mots clés susceptibles d'apparaître dans le titre ou dans le corps du texte, comme **cellules souches** / *células madre* et le nom du chercheur coréen responsable de l'équipe (D' Hwang).

Une fois une banque de textes constituée, nous avons décidé de ne conserver que les articles dont le but dominant est de transmettre un savoir scientifique, ce qui exclut les éditoriaux, les lettres des lecteurs et autres textes d'opinion, ainsi que les articles rendant compte des réactions des divers acteurs sociaux (politiques, religieux, etc.) et les brèves. De cette manière, 11 articles de vulgarisation ont été retenus; ils constituent deux corpus (espagnol et français) formés par quatre sous-corpus (espagnol, argentin, français et québécois). Chaque texte téléchargé a été sauvegardé sous un même nom de fichier (encodant le nom du journal, la date et un mot-clé du titre) en format .pdf, afin de conserver la mise en page originale, et en format .txt, ce qui permettait la conversion en langage XML pour

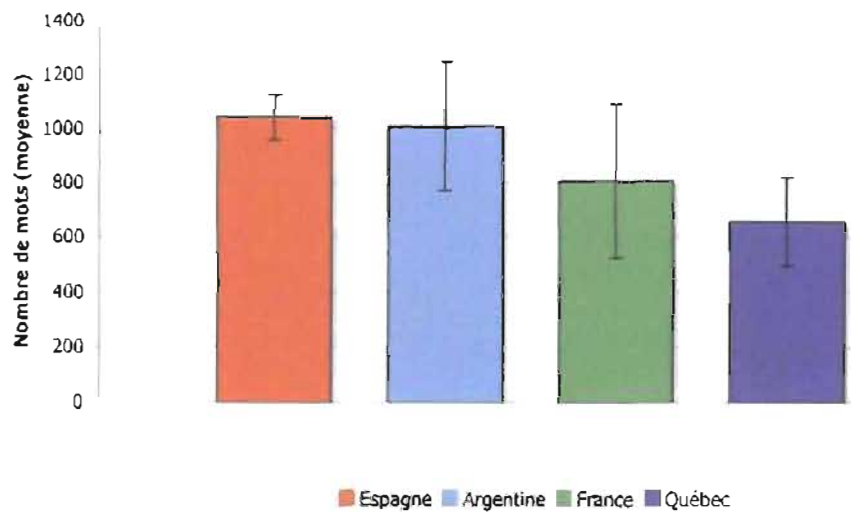
³⁹ [Notre traduction :] « *la vague d'activité médiatique intense et diversifiée autour d'un événement ponctuel* ».

l'annotation. Pour l'étiquetage des textes, des fichiers .xml ont ainsi été créés. Les informations bibliographiques ont été inscrites dans un fichier *EndNote* qui a servi à alimenter la base de données en ligne regroupant les résultats. Le tableau 1 et la figure 1 montrent la composition des corpus.

Tableau 1. Composition des corpus

Corpus en espagnol : 6125 mots		
Espagne : 3112 mots		
ABC (A040213Primera) : 1135 mots		Moyenne : 1037,33 mots (± 84,91)
<i>El Mundo</i> (EM040212Coreanos) : 996 mots		
<i>El País</i> (LP040213Coreanos) : 981 mots		
Argentine : 3013 mots		
<i>Clarín</i> (C040213Primera) : 857 mots		Moyenne : 1004,33 mots (± 237,23)
<i>La Nación</i> (LN040213NuevaEra) : 878 mots		
<i>Página 12</i> (Pag040213Paso) : 1278 mots		
Corpus en français : 3720 mots		
France : 2412 mots		
<i>Libération</i> (L040213Seoul) : 489 mots		Moyenne : 804 mots (± 284,15)
<i>Le Figaro</i> (LF040213Embryons) : 1041 mots		
<i>Le Monde</i> (LM040213Coreens) : 882 mots		
Québec : 1308 mots		
<i>Le Devoir</i> (LD040213Embryon) : 538 mots		Moyenne : 654 mots (± 164,05)
<i>La Presse</i> (LP040213Premiere) : 770 mots		
Total : 9845 mots		

Figure 1. Taille moyenne des sous-corpus



La différence de taille entre les corpus découle des critères de compilation. Nous avons déterminé que les textes retenus devaient avoir été publiés dans des journaux ayant une portée nationale (provinciale pour le Québec), et qu'il ne s'agirait que des textes essentiellement informatifs. Dans le cas du Québec, seuls deux journaux répondent à ce critère, *La Presse* et *Le Devoir*, tandis que l'on pouvait compter sur trois quotidiens pour chacun des autres pays. Seulement 2 textes, contre 3 textes dans les autres sous-corpus, ont pu être conservés. De plus, les articles en espagnol sont plus longs en moyenne que les articles en français. L'écart entre les corpus a été analysé au moyen d'un test t de Student, qui a révélé une différence significative ($p = 0,044^{40}$). Pourtant, en raison des caractéristiques de nos échantillons, ce résultat est à considérer avec précaution.

4.2. Collecte de données

4.2.1. Éléments relevés

La première étape de notre étude consiste dans le repérage au sein des textes des expressions employées pour présenter des notions spécialisées⁴¹. Seules ont été considérées celles qui appartenaient au domaine principal des articles : la recherche sur les cellules souches et sur le clonage. Nous avons exclu les notions reliées à sa réalisation concrète (acteurs, exécution technique, équipement et diffusion des résultats) afin de nous concentrer exclusivement sur l'objet d'étude scientifique. Nous n'avons pas seulement ciblé les entités concrètes ou abstraites (/cellule/, /maladie/) mais aussi les actions (/traiter un patient/), les processus (/développement d'un embryon/), les phénomènes (/présence d'une aberration chromosomique/), les états (/indifférenciation/) et les propriétés (/pluripotence/)⁴².

Le repérage des éléments des corpus se fonde sur un principe d'ordre cognitif : ils n'ont pas été relevés en fonction de leur structure syntaxique ni de la signification des unités lexicales qui les composent, mais parce qu'ils renvoient en discours aux notions du domaine délimité. Cela veut dire que le sens référentiel a toujours primé sur la signification linguistique : ainsi, l'unité célula n'a pas toujours été considérée comme la réalisation linguistique de la notion /cellule/, car parfois, elle renvoyait à une notion plus spécifique en tant que reprise anaphorique (/cellule souche/). Afin de pouvoir regrouper les expressions relevées en fonction de la notion à laquelle elles font référence, nous avons donné à chaque notion un nom-étiquette qui prend la forme d'une unité ou d'un

⁴⁰ Pour tous les tests statistiques, le seuil de rejet a été établi à 0,05.

⁴¹ Ce que nous entendons par *notion* a été discuté à la section 2.3. du « Cadre théorique ».

⁴² Il s'agit des hyper-catégories dans lesquelles nous avons rangé les notions (section 2.3. du « Cadre théorique »). Ce classement nous a permis de déceler quelques régularités lors de l'analyse des résultats (section 5.1.4. des « Résultats »).

syntagme nominal pour les entités, les processus, les phénomènes, les états et les propriétés et d'un syntagme verbal pour les actions (voir les exemples ci-dessus). Cette étiquette se veut une brève description de la notion et ne correspond pas nécessairement au terme couramment employé dans le domaine. Par exemple, nous avons décidé de nommer /remplacer le noyau d'un ovocyte par le noyau d'une cellule/ ce que les spécialistes appellent « transfert nucléaire ».

Nous avons tenu également compte des expressions renvoyant à des notions qui étaient contenues dans une expression plus importante déjà repérée, dans la mesure où elles y occupaient une place d'actant ou de circonstant. Par exemple, la séquence clonación de embriones humanos, qui fait référence à la notion /créer un embryon humain par clonage/, contient le syntagme embrión humano, qui renvoie à /embryon humain/. En revanche, nous n'avons pas étiqueté les expressions qui étaient imbriquées dans d'autres expressions et qui renvoyaient à une notion plus générique, comme **embryon** dans **embryon humain**. Relever **embryon** aurait été redondant et faussé les résultats.

Comme nous l'avons mentionné, les expressions ont été considérées indépendamment de leur forme linguistique. En effet, l'un des objectifs de notre étude étant de décrire la structure morphosyntaxique des expressions employées, il ne nous semblait pas méthodologiquement pertinent d'introduire des critères formels dans le repérage des éléments linguistiques à analyser. Cependant, nous n'avons retenu ni les pronoms (**on les [cellule souche pluripotente] y [produire tous les types de cellules différenciées de l'organisme humain] pousse**, algunas [célula madre obtenida de un embrión clónico] se transplantaron a ratones), ni les adjectifs possessifs reprenant un antécédent par eux-mêmes (**les chercheurs ont coupé court à son [embryon] développement**).

En revanche, une fois une séquence linguistique isolée, il était important de connaître sa structure morphosyntaxique. Nous avons repéré des noms (**embryon**), des verbes (clonar) et des adjectifs (adulto) ainsi que des syntagmes nominaux (**cellules souches**), verbaux (**produire des embryons humains par clonage**) et participiaux (islote pancreático transplantado a una persona diabética), voire des propositions complètes (las células madre obtenidas de un embrión clónico eran pluripotentes). Lorsqu'il s'agissait d'une structure prédicative, le prédicat et les actants réalisés dans la phrase ont été identifiés (dans la proposition **un embryon humain est produit par clonage** : **produire** est le prédicat, **embryon humain** réalise le deuxième actant tandis que **clonage** réalise le troisième; le premier actant n'est pas réalisé). En plus de la notion exprimée par l'unité ou le syntagme, nous avons étiqueté l'hyper-catégorie à laquelle celle-ci appartient. Pour la distinction et le décompte des expressions, nous avons tenu compte des formes lemmatisées; dans le cas de prédicats, nous nous sommes fondée sur la structure actantielle et la classe d'actants réalisés.

En outre, dans l'annotation nous avons mis en évidence des stratégies discursives qui contribuent à la cohérence et à la cohésion textuelles en reliant les expressions repérées : diaphore, reformulation, définition et explication.

- Diaphore

(8) « A partir de ces 242 ovocytes, les chercheurs coréens ont obtenu 30 *embryons clonés*, qui ont pu être cultivés jusqu'au stade blastocyste. Des *cellules-souches* n'ont pu être identifiées que chez 20 de *ces embryons* et des lignées de *ces cellules* n'ont pu être isolées et cultivées qu'à partir d'un seul embryon. »

Dans ce passage, nous trouvons deux reprises anaphoriques : **embryons** reprend **embryons clonés**, de la phrase précédente, tandis que **cellules** reprend **cellules-souches**, dans la même phrase. Nous avons tenu compte des séquences cataphoriques comme **objet de leur quête** dans

(9) « [...] en extrayant de sa paroi interne l'*objet de leur quête* : des cellules souches "pluripotentes" »,

Comme nous l'avons mentionné, les éléments anaphoriques pronominaux n'ont pas été relevés.

- Reformulation

(10) « [...] et provenant dans le cas présent de la *couronne* ou *cumulus* qui entoure naturellement l'ovocyte. »

Cumulus est proposé en tant que reformulation de **couronne**⁴³.

- Définition

(11) « [...] permettant ainsi d'obtenir 30 blastocystes [sic], un *embryon aux premiers stades de son développement qui ne contient que 100 cellules* [...] ».

Le syntagme **embryon aux premiers stades de son développement qui ne contient que 100 cellules** est fourni comme définition de **blastocyste**.

- Explication

(12) « L'annonce risque de relancer la polémique sur le "*clonage thérapeutique*", qui *vis*e non pas à créer des bébés mais des traitements médicaux. »

La proposition relative **qui vise non pas à créer des bébés mais des traitements médicaux** constitue une explication de **clonage thérapeutique**.

⁴³ Dans cet exemple, l'absence de la virgule crée une ambiguïté.

Enfin, à l'intérieur des expressions repérées, nous avons isolé les indices de conceptualisation métaphorique. Afin de caractériser les modes de conceptualisation, nous avons identifié la structure actancielle réalisée dans le texte et indiqué celle correspondant au cadre conceptuel source évoqué par l'indice en question. Ainsi, dans le syntagme **production de cellules souches embryonnaires**, nous reconnaissons **production** comme un indice de conceptualisation métaphorique. L'actant de cette unité prédicative est réalisé dans le texte par **cellules souches embryonnaires**, mais il évoque le cadre conceptuel de la production industrielle, où l'actant aurait été réalisé par le nom d'un produit manufacturé.

4.2.2. Système d'annotation

Pour la collecte des données, nous avons mis sur pied en collaboration avec Sylvie Boudreau un système d'annotation des textes permettant d'isoler les éléments linguistiques visés et de rendre compte de leurs caractéristiques. D'emblée, nous cherchions un système simple, ne nous forçant ni à faire des hypothèses au préalable sur les formes sous lesquelles les notions sont représentées linguistiquement (unités, syntagmes ou propositions) ni à nous prononcer *a priori* sur le statut terminologique ou le degré de figement des éléments relevés. Il devait aussi nous permettre de faire une distinction claire entre les niveaux d'analyse (cognitif et formel) et de modifier ou corriger l'annotation facilement. Enfin, nous souhaitions disposer d'une méthode permettant d'extraire les données en fonction des différents paramètres utilisés pour l'annotation. Nous avons donc adopté - et adapté en fonction de nos besoins - le système de balisage de textes faisant appel au langage XML conçu par Vandaele et Boudreau (2006). Cette méthode « *présente l'avantage d'une grande flexibilité, tant pour les corpus annotés (ajout de textes dans différentes langues sans perturbation du travail déjà effectué, processus de révision de l'annotation), que pour les jeux de balises et d'attributs (déterminés en fonction des caractéristiques à étudier)* » (Vandaele et Boudreau 2006, p. 947)⁴⁴. Ainsi, nous avons déterminé un jeu de balises et d'attributs permettant d'identifier et de décrire les éléments relevés dans le corpus.

L'étiquetage a été réalisé à l'aide du logiciel *Oxygen v.8.1*, lequel convertit chaque fichier en format .txt dans un fichier en format .xml. Dans chacun d'entre eux, la balise <fichier> englobe tout le document et établit le lien entre le fichier et le schéma

⁴⁴ Pour une description technique du système, on peut se référer à Vandaele, S. et S. Boudreau (2006). *Annotation XML et interrogation de corpus pour l'étude de la conceptualisation métaphorique*, JADT 2006 : 8^{es} Journées internationales d'analyse statistique des données textuelles, Besançon.

permettant la validation de l'annotation. Au niveau immédiatement inférieur se trouvent les balises <nom>, <nature>, <langue> et <texte>. Les trois premières contiennent des informations sur le document tandis que la dernière entoure le texte étiqueté. La figure 2 montre la disposition de ces balises dans le document.

Figure 2. Balises <fichier>, <nom>, <nature>, <langue> et <texte>

```
<fichier xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.biomettico.com"
xsi:schemaLocation="http://www.biomettico.com/schemasXBC
http://www.biomettico.com/schemasXBC/kallaway.xsd">
  <nom>L040213Seoul</nom>
  <nature>vulgarisation</nature>
  <langue>fr-FR</langue>
  <texte>
  [...]
  </texte>
</fichier>
```

Notre système d'annotation comporte également des balises indiquant les parties et l'organisation du texte : le surtitre, s'il y a lieu, (<surtitre>), le titré (<titre>), le chapeau (<chapeau>), les sous-titres lorsqu'il y en a (<sous-titre>), les paragraphes (<par>) et les phrases (<phr>). Cette dernière balise permet de numérotter toutes les phrases du texte à l'aide de l'attribut @position. Dans la figure 3, afin d'illustrer l'utilisation de ces balises, la balise <texte> est donnée comme point de repère et les balises internes ont été omises.

Figure 3. Balises <surtitre>, <titre>, <chapeau>, <sous-titre>, <par> et <phr>

```
<texte>
  <surtitre>
    <phr position="1">SCIENCES COREE DU SUD</phr>
  </surtitre>
  <titre>
    <phr position="2">Des biologistes coréens ont créé des embryons humains par
clonage et obtenu des cellules-souches</phr>
  </titre>
  <chapeau>
    <phr position="3">Cette première marque une nouvelle étape dans la maîtrise d'une
technique controversée visant à mettre au point des traitements cellulaires contre des
maladies dégénératives incurables.</phr>
  </chapeau>
  <par>
    <phr position="4">QUATORZE biologistes travaillant dans différentes structures
scientifiques de Corée du Sud associés à un spécialiste américain devaient annoncer, jeudi 12
février, sur le site Internet de la revue américaine Science (www.sciencexpress.org), avoir
réussi à créer des embryons humains par clonage.</phr>
    <phr position="5">Ils expliquent, en outre, avoir obtenu, à partir d'un embryon, des
lignées de cellules-souches capables de se différencier.</phr>
    <phr position="6">Cette première relance le débat éthique sur le caractère légitime
ou non de la création d'embryons humains à partir de la technique du clonage par transfert
nucléaire.</phr>
  </par>
```

```
[...]
<par>
  <phr position="27">Pour Mme Cavazzana-Calvo, la publication de la première
  coréenne suffit, s'il en est encore besoin, pour démontrer au gouvernement l'urgence de la
  promulgation du texte de loi de révision du dispositif législatif de bioéthique de 1994 qui, à
  titre exceptionnel, autorisera des recherches sur les seules cellules-souches.</phr>
  <phr position="28">«Le gouvernement avait annoncé que la dernière étape, l'examen
  en seconde lecture du texte par le Sénat, interviendrait en février, dit-elle.</phr>
  <phr position="29">Or nous apprenons que cet examen est reporté au mois de
  mai.</phr>
  <phr position="30">Nous ne cessons malheureusement, en France, de prendre du
  retard sur les travaux de nos collègues étrangers.»</phr>
</par>
</texte>
```

À l'intérieur de chaque phrase, les expressions retenues sont entourées par la balise `<lingNot>`, qui comporte plusieurs attributs. Le premier, `@id`, a comme valeur le numéro d'identification de l'élément repéré, comprenant le numéro de position de la phrase et un numéro assigné selon l'ordre d'annotation, séparés par un point virgule. Puis, dans l'attribut `@lem`, on explicite la forme lemmatisée de l'expression lorsqu'elle se trouve sous une forme fléchie ou comprend des éléments non pertinents. La structure morphosyntaxique est consignée à l'aide de l'attribut `@synt`. S'il y a lieu, le prédicat et les actants réalisés sont indiqués au moyen des attributs `@Pred`, `@act` et `@Act`. L'attribut en minuscule est utilisé lorsque la valeur inscrite est le numéro d'identification d'un élément étiqueté, tandis que l'attribut en majuscule accepte comme valeurs les formes lemmatisées d'éléments non étiquetés. Cette décision résulte d'une contrainte informatique liée à l'interrogation des corpus. Lorsqu'il s'agit d'une reformulation, d'une définition ou d'une reprise diaphorique, le lien avec le numéro d'identification ou la forme lemmatisée des éléments reformulés, définis ou repris est fait par l'intermédiaire des couples d'attributs `@ref/@Ref`, `@definition/@Definition`, et `@anaph/@Anaph`. Le cas échéant, le marqueur de la reformulation, de la définition ou de la reprise diaphorique est indiqué dans l'attribut `@marq`. Quant à la notion exprimée par l'unité ou le syntagme repérée, elle figure comme valeur de l'attribut `@notion` et l'hyper-catégorie sous laquelle elle peut être classée, dans l'attribut `@hyper`. Enfin, l'attribut `@comm` sert à exprimer tout commentaire supplémentaire. La figure 4 offre un exemple de phrase annotée (seulement les balises `<phr>` et `<lingNot>` sont présentées).

Figure 4. Balise `<lingNot>` et ses attributs

```
<phr position="15">Ensuite, ils ont réussi à
  <lingNot id="15; 1" synt="sverb" Pred="maintenir" Act1="ils
  [chercheur]" act2="15; 2" Act3="culture" notion="cultiver une
  cellule" hyper="action"> maintenir en culture ces
  <lingNot id="15; 2" lem="cellule" synt="nom" anaph="14; 6"
  marq="ces" not="cellule souche pluripotente" hyper="entité
  concrète">cellules</lingNot> </lingNot>. </phr>
```

Les segments constituant des explications sont identifiés à l'aide de la balise <expl>, où se trouvent également @id, @lem, @synt et @comm. Le couple d'attributs @idRef/@IdRef assurent le lien avec l'élément objet de l'explication à travers son numéro d'identification ou sa forme lemmatisée et l'attribut @marq indique le marqueur explicatif, s'il y a lieu. La figure 5 présente une phrase comportant une explication.

Figure 5. Balise <expl> et ses attributs

```
<phr position="8">Mais il s'agissait d'
  <lingNot id="8; 1" lem="embryon mal formé" synt="snom"
not="embryon humain mal formé" hyper="entité concrète">embryons mal
formés</lingNot>,
  <expl id="8; 3" synt="sadj" idRef="8; 1" marq=",">possédant moins
d'une dizaine de
  <lingNot id="8; 2" lem="cellule" synt="nom" not="cellule"
hyper="entité concrète">cellules</lingNot></lingNot></expl>. </phr>
```

La balise <concInd> sert à isoler les indices de conceptualisation parmi les éléments linguistiques repérés à l'aide de <lingNot>. En plus des attributs @id, @lem, @synt et @comm, dont le rôle a été expliqué ci-dessus, cette balise possède des attributs permettant de rendre compte de la structure prédicative contrôlée par l'indice de conceptualisation. La valeur de @actn indique les actants réalisés linguistiquement dans la phrase tandis que, dans @metn, on mentionne les actants typiques du prédicat. L'emploi de cette balise est illustré dans la figure 6.

Figure 6. Balise <concInd> et ses attributs

```
<phr position="21">Il faut reconnaître que la «rentabilité» de la
  <lingNot id="21; 1" synt="nom" Anaph="passage : phrases 12-16"
marq="la" notion="obtenir une cellule souche d'un embryon humain créé
par clonage" hyper="action">méthode</lingNot>
est très faible puisqu'il a fallu disposer de 246
  <lingNot id="21; 2" lem="ovocyte" synt="nom" notion="ovocyte"
hyper="entité concrète">ovocytes</lingNot> au départ pour obtenir en fait un seul
  <lingNot id="21; 3" synt="snom" notion="embryon humain viable"
hyper="entité concrète">embryon
  <concInd id="21; 5" synt="adj" Act1="embryon" Act2="produire des
cellules souches embryonnaires" met1="être vivant"
met2="action">capable</concInd> de
  <concInd id="21; 6" synt="verbe" Act1="embryon" act2="21; 4"
met1="être humain" met2="produit manufacturé">produire</concInd> des
  <lingNot id="21; 4" lem="cellule souche embryonnaire" Pred="cellule
souche" synt="snom" notion="cellule souche d'un embryon" hyper="entité
concrète" comm="l'adjectif relationnel 'embrionario' renvoie au
premier actant du prédicat : 'embrión'">cellules souches
embryonnaires</lingNot></lingNot>. </phr>
```


Enfin, la balise <sic> indique les coquilles ou les erreurs repérés dans le texte. Aux attributs @id et @comm vient alors s'ajouter @cor, qui permet d'introduire une correction, comme le montre la figure 7.

Figure 7. Balise <sic> et ses attributs

```
<phr position="24">Il y a de plus la question déontologique :
  <lingNot id="24; 1" synt="prop" not="obtenir un ovule"
hyper="action">ces
  <lingNot id="24; 2" lem="ovocyte" synt="nom" anaph="23.1"
marq="ces" not="ovocyte" hyper="entité concrète">ovocytes</lingNot>
ont été collectés auprès de seize femmes</lingNot> qui ont accepté, «sans rétribution
financière», précisent les auteurs, de
  <lingNot id="24; 3" synt="sverb" not="se soumettre à un traitement"
hyper="action">se soumettre à un
  <lingNot id="24; 4" synt="snom" not="thérapie hormonale"
hyper="entité abstraite"45>traitement hormonal d'hyper-
ovulation</lingNot></lingNot>, sans bénéfice pour
  <sic id="24; 5" cor="elles">elle</sic>, mais non sans risque. </phr>
```

Le tableau 2 regroupe les balises et attributs utilisés pour l'annotation du corpus.

Tableau 2. Balises et attributs

Balises	Attributs	Description
<fichier>		Structure XML, nom du fichier et texte annoté
<texte>		Texte
<surtitre>		Surtitre
<titre>		Titre
<chapeau>		Chapeau
<sous-titre>		Sous-titre
<par>		Paragraphe
<phr>		Phrase
	@position	Numéro d'identification de la phrase
<lingNot>		Séquence linguistique faisant référence à une notion
	@id	Numéro d'identification de la séquence
	@lem	Forme lemmatisée de la séquence
	@synt	Structure syntaxique de la séquence
	@Pred	Unité prédicative
	@actn	Actants de l'unité prédicative lorsqu'ils sont réalisés (numéro d'identification ou forme lemmatisée)
	@Actn	
	@ref	Numéro d'identification ou forme lemmatisée de la séquence reformulé par l'expression repérée
	@Ref	
	@definition	Numéro d'identification ou forme lemmatisée de l'élément défini par l'expression repérée
	@Definition	
	@anaph	Numéro d'identification ou forme lemmatisée de l'élément repris par l'expression repérée
	@Anaph	
	@marq	Marqueur linguistique de la reformulation, de la définition ou de la reprise diaphorique
	@notion	Notion à laquelle fait référence l'expression repérée
	@hyper	Hyper-catégorie sous laquelle on peut ranger la notion
	@comm	Commentaire

⁴⁵ Bien que l'on puisse considérer cette notion comme un processus, nous avons décidé de l'envisager comme *quelque chose* pouvant être utilisé pour guérir une maladie.

<conclnd>		Indice de conceptualisation métaphorique
	@id	Numéro d'identification de l'indice de conceptualisation
	@lem	Forme lemmatisée de l'indice de conceptualisation
	@synt	Partie du discours à laquelle appartient l'indice de conceptualisation
	@actn	Numéro d'identification ou forme lemmatisée des éléments qui réalisent linguistiquement les actants de l'indice de conceptualisation
	@Actn	
	@metn	Étiquette des actants typiquement associés à l'indice de conceptualisation
@comm	Commentaire	
<expl>		Élément linguistique constituant une explication d'une expression repérée
	@id	Numéro d'identification de l'explication
	@lem	Forme lemmatisée de l'explication
	@synt	Structure syntaxique de l'explication
	@idRef	Numéro d'identification ou forme lemmatisée de l'élément expliqué
	@ldRef	
	@marq	Marqueur linguistique de l'explication
	@comm	Commentaire
<sic>		Coquille ou erreur de l'auteur
	@id	Numéro d'identification de la coquille
	@cor	Correction
	@comm	Commentaire

4.2.3. Récupération et traitement des données

Une fois les textes étiquetés, les données relatives aux textes de vulgarisation ont été extraites au moyen d'un formulaire d'interrogation en ligne développé par Sylvie Boudreau. Il s'agit d'une base de données alimentée par les textes annotés. Interrogeable en ligne, cette base fait partie du site Biomettico (www.biomettico.net), mis récemment sur pied par l'équipe de Vandaele (Vandaele et Boudreau à paraître, Vandaele *et al.* à paraître). Puis, les données ont été rapportées dans un tableur (*Excel*) permettant une analyse fine des résultats et la création de graphiques. En raison des objectifs méthodologiques et théoriques de la présente étude, le traitement des données ainsi obtenues vise surtout l'analyse qualitative des résultats.

5. *Résultats*

La méthodologie mise en place nous a permis d'obtenir plusieurs types de données quantitatives et qualitatives à propos des corpus étudiés. Afin d'éviter la confusion entre les corpus et les sous-corpus, nous désignerons les premiers par le nom de la langue dans laquelle sont rédigés les textes (corpus en espagnol et corpus en français), tandis que nous appellerons les seconds par leur lieu de publication (Espagne, Argentine, France et Québec).

Les résultats portent sur les notions présentées dans les textes, les expressions employées pour les exprimer, les modes de conceptualisation ainsi que sur certaines stratégies discursives adoptées. En ce qui concerne les données quantitatives, il faut préciser que, sauf exception, elles n'ont pas fait l'objet d'un traitement statistique. D'une part, plusieurs facteurs limitent la possibilité de réalisation de tests ainsi que la puissance explicative de ceux qui aurait pu être appliqués : la taille réduite des corpus, la variabilité entre les textes pour certaines variables et la faible quantité de données pour quelques-uns des phénomènes étudiés. D'autre part, le caractère prospectif de cette étude – qui vise surtout la validation d'une démarche méthodologique et le recueil d'observations préliminaires – rendait préférable de reporter l'utilisation de tests à une étape ultérieure pour laquelle nous compterons sur un nombre des données permettant l'obtention de résultats significatifs.

La description des résultats est réalisée sur trois paliers⁴⁶. En premier lieu, nous comptons le nombre de *notions* relevées dans le cadre de ce travail : /ovocyte/ en est une et /créer un embryon humain par clonage/ en est une autre. Puis, nous observons la place que ces notions occupent dans les corpus à l'aide du nombre d'*actualisations*, autrement dit, le nombre de fois où l'on y fait référence (que ce soit avec la même expression ou non). Par exemple, la notion /créer un embryon humain par clonage/ est actualisée 22 fois dans le sous-corpus France, au moyen de 15 unités linguistiques et syntagmes différents. Enfin, nous étudions le nombre d'*expressions* employées pour rendre chaque notion. Ainsi, dans le sous-corpus Argentine, la notion /ovocyte/, qui y est actualisée 13 fois, a été exprimée par 4 expressions (unités linguistiques ou syntagmes) différentes : óvulo, óvulo en desarrollo dentro del ovario, ovocito, células que, cuando se desprenden del ovario para deslizarse hacia el útero, pasan a llamarse óvulos.

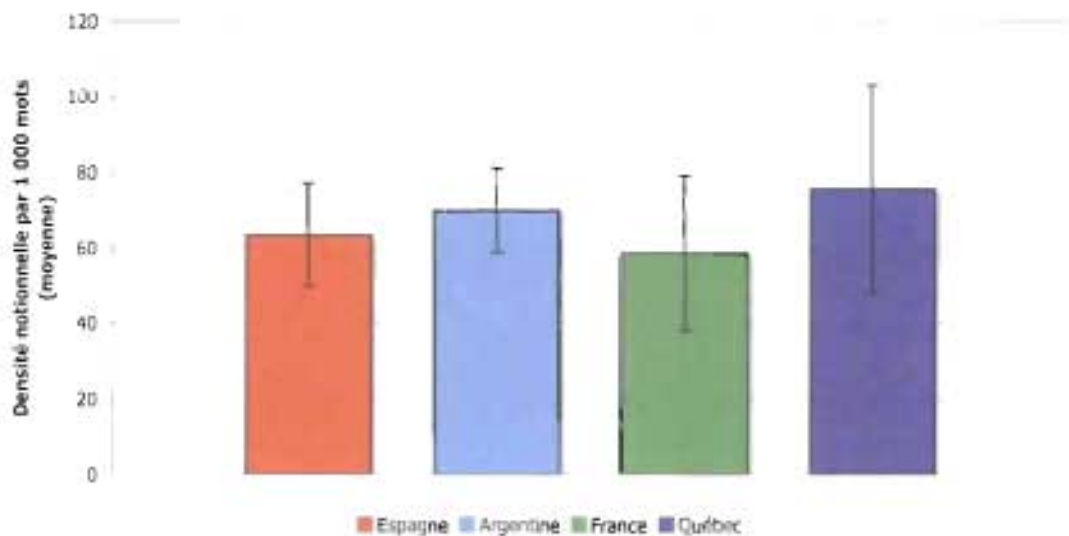
⁴⁶ Ces paliers d'analyse ont été annoncés à la section 2.3. du « Cadre théorique ».

5.1. Notions

5.1.1. Nombre de notions et densité notionnelle

Les unités et syntagmes repérés au cours de l'annotation renvoient à un total de 184 notions, qui ne se trouvent cependant pas toutes dans les quatre sous-corpus, ni même dans les deux corpus. De l'ensemble de 184 notions, 158 apparaissent dans le corpus en espagnol et 117 figurent dans le corpus en français. Nous en trouvons 120 dans Espagne, 124 dans Argentine, 81 dans France et 77 dans Québec. Afin d'éliminer le biais relié à la variabilité du nombre de mots dans les textes, nous avons ramené la densité notionnelle moyenne de chaque sous-corpus à 1 000 mots (Figure 8).

Figure 8. Densité notionnelle moyenne par 1 000 mots dans chaque sous-corpus

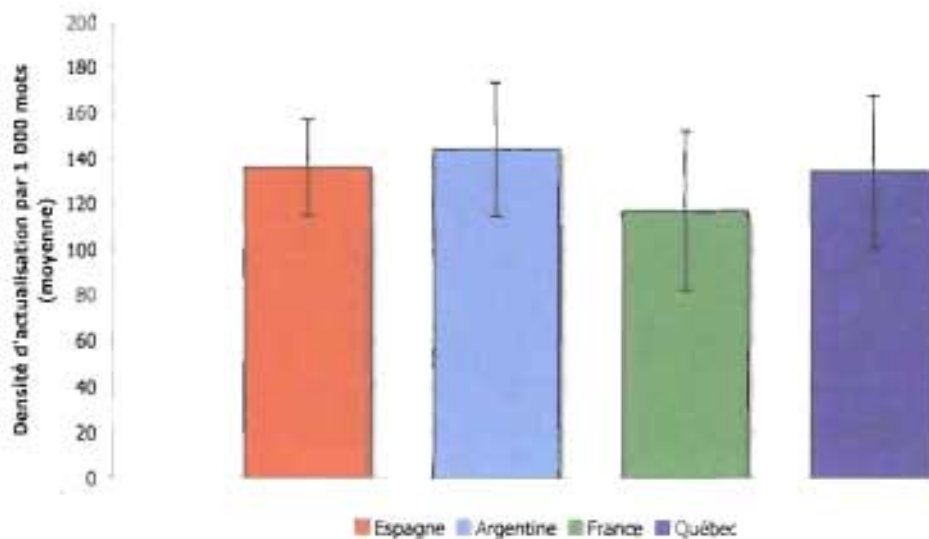


Ainsi, on observe en moyenne 66,46 notions par 1 000 mots en espagnol ($\pm 11,84$) et 65,10 notions par 1 000 mots en français ($\pm 22,14$). Les textes en français semblent donc presque aussi denses que les textes en espagnol, même si on remarque une variabilité plus marquée chez les premiers. En effet, le test t de Student ne révèle pas une différence significative ($p = 0,899$), mais la taille des corpus et les variabilités intertexte assez élevées (plus prononcées dans le corpus en français) exigent d'être prudent dans l'interprétation de ce résultat. Quant aux sous-corpus, ce sont les textes de Québec qui présentent la plus haute densité notionnelle moyenne (75,32 par 1 000 mots [$\pm 27,54$]) tandis que la plus basse correspond aux textes de France (58,28 par 1 000 mots [$\pm 20,66$]).

5.1.2. Actualisation des notions

Globalement, les notions sont actualisées - c'est-à-dire réalisées linguistiquement dans les textes - 1291 fois dans les deux corpus. Les densités moyennes d'actualisation des notions ramenées à 1 000 mots ne montrent pas un grand écart entre le français et l'espagnol : 140,05 fois ($\pm 23,37$) et 123,92 ($\pm 31,38$) fois respectivement. La réalisation d'un test t de Student appuie cette idée : $p = 0,353$. On compte 426 actualisations de notions dans Espagne, 432 dans Argentine, 263 dans France et 170 dans Québec. La plus haute densité moyenne d'actualisation de notions se trouve dans les textes d'Argentine (144,02 par 1 000 mots [$\pm 29,56$]) et la plus basse dans les textes de France (117,07 par 1 000 mots [$\pm 35,03$]), comme le montre la figure 9. La variation intertexte en ce qui concerne les actualisations des notions semble également assez élevée.

Figure 9. Densité moyenne d'actualisation de notions par 1 000 mots dans chaque sous-corpus



5.1.3. Groupes de notions communes et noyau thématique

Comme nous l'avons mentionné, les notions prises en compte sont actualisées en nombre variable au sein de chaque sous-corpus. Espagne et Argentine présentent 86 notions communes, ce qui représente en moyenne 82,97 % ($\pm 10,35$ %) des notions et 88,35 % ($\pm 7,35$ %) des actualisations du premier sous-corpus ainsi que 77,04 % ($\pm 7,74$ %) des notions et 83,70 % ($\pm 4,79$ %) des actualisations du second (Figures 10 et 11).

¹ Les mêmes remarques que pour le résultat statistique précédent s'impesent ici.

Figure 10. Proportion moyenne dans chaque sous-corpus des notions communes à l'ensemble du corpus en espagnol

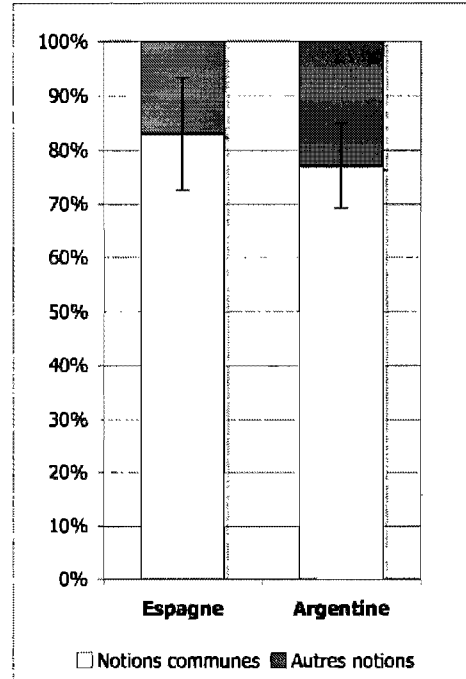
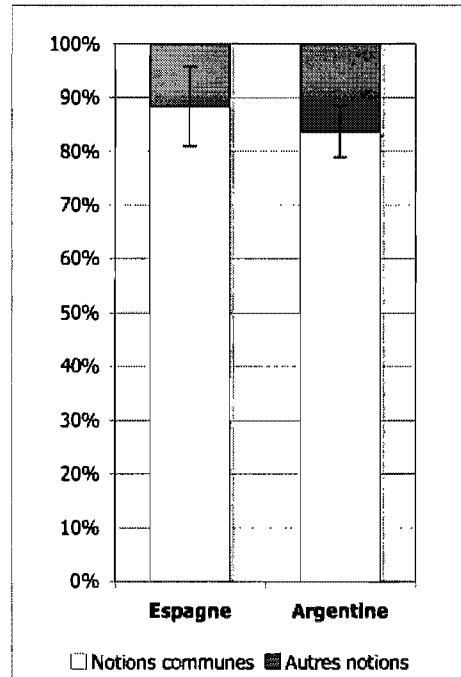


Figure 11. Proportion moyenne dans chaque sous-corpus des actualisations des notions communes à l'ensemble du corpus en espagnol



France et Québec partagent 41 notions, représentant en moyenne 62,80 % ($\pm 2,44$ %) et 60,01 % ($\pm 7,22$ %) des notions, et 78,53 % ($\pm 2,12$ %) et 70 % ($\pm 2,50$ %) des actualisations respectivement (Figures 12 et 13). Nous observons donc une convergence thématique plus marquée dans le corpus en espagnol. Cependant, les notions communes au corpus en français sont aussi celles qui présentent le plus grand nombre d'actualisations.

Figure 12. Proportion moyenne dans chaque sous-corpus des notions communes à l'ensemble du corpus en français

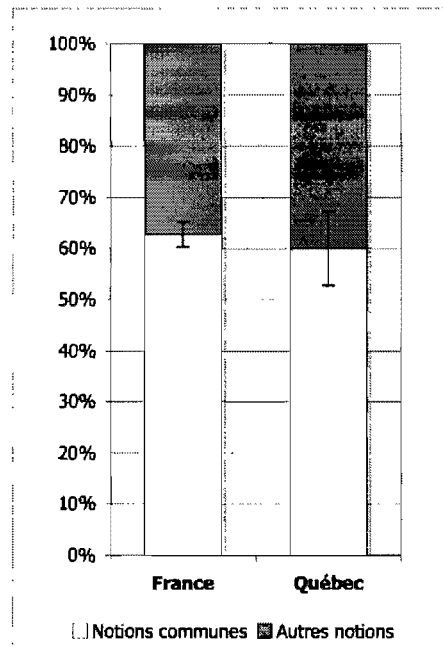
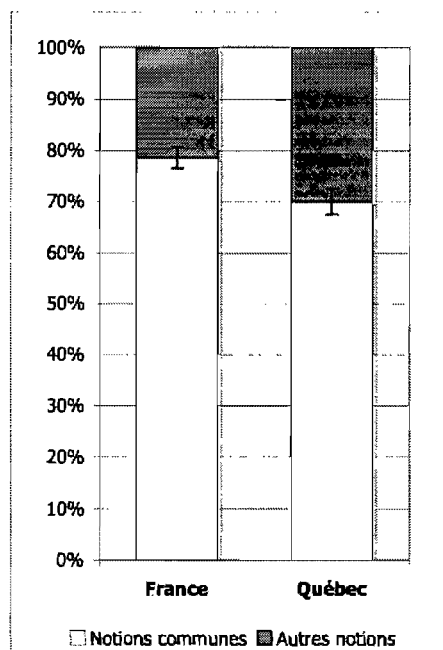


Figure 13. Proportion moyenne dans chaque sous-corpus des actualisations des notions communes à l'ensemble du corpus en français

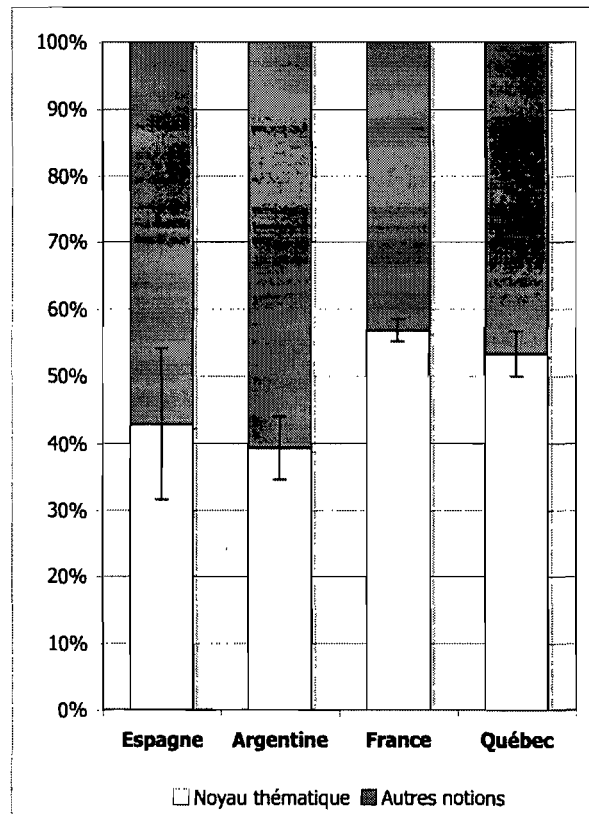


Les deux corpus ont 91 notions en commun - dont certaines s'actualisent dans un sous-corpus plutôt que dans l'autre - tandis que 35 notions sont actualisées dans les quatre sous-corpus (Tableau 3). Les notions de ce dernier groupe, que nous considérons le **noyau thématique** des textes, se répartissent en trois ensembles notionnels principaux : le clonage, l'obtention de cellules souches embryonnaires pluripotentes et les applications thérapeutiques. Si nous classons toutes les notions des corpus selon leur fréquence d'actualisation, les 35 notions du noyau thématique se trouvent dans les 49 premiers rangs. De plus, bien que ce groupe ne constitue que 19,02 % de l'ensemble des notions, il représente en moyenne 47,59 % ($\pm 9,59$ %) des notions de chaque texte. Si nous considérons les sous-corpus individuellement, il correspond en moyenne à 42,79 % ($\pm 11,28$ %) des notions dans chaque texte d'Espagne, à 39,25 % ($\pm 4,73$ %) de celles d'Argentine, à 56,86 % ($\pm 1,62$ %) de celles de France et à 53,40 % ($\pm 3,42$ %) de celles de Québec (Figure 14).

Tableau 3. Notions actualisées dans les quatre sous-corpus

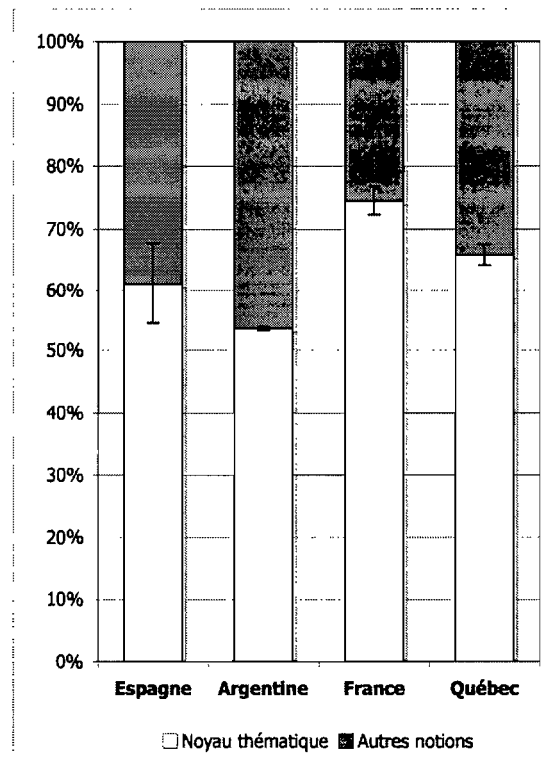
/blastocyste/
 /cellule/
 /cellule différenciée/
 /cellule génétiquement identique à un être humain/
 /cellule souche/
 /cellule souche d'un embryon/
 /cellule souche pluripotente/
 /créer un embryon animal par clonage/
 /créer un embryon humain par clonage/
 /créer un embryon humain par clonage dans un but thérapeutique/
 /créer un embryon par clonage/
 /développement d'un embryon/
 /diabète/
 /différenciation d'une cellule souche en une cellule différenciée/
 /donner un ovocyte/
 /embryon génétiquement identique à un être humain/
 /embryon humain/
 /embryon humain viable/
 /être humain génétiquement identique à un autre/
 /extraire un noyau/
 /greffer une cellule souche/
 /guérir un patient/
 /insérer un noyau d'une cellule dans un ovocyte/
 /lignée de cellules souches/
 /noyau/
 /obtenir une cellule souche d'un embryon humain créé par clonage/
 /organisme/
 /ovocyte/
 /patient/
 /prélever un ovocyte/
 /prélever une cellule/
 /remplacer le noyau d'un ovocyte par le noyau d'une cellule/
 /stade de blastocyste/
 /thérapie/
 /tissu/

Figure 14. Proportion moyenne dans chaque sous-corpus des notions du noyau thématique



Par ailleurs, ces 35 notions comptent à elles seules pour plus de la moitié du total d'actualisations (795, soit 61,58 %). Si nous tenons compte des actualisations de chaque sous-corpus, ce groupe représente en moyenne 61,05 % ($\pm 6,58$ %) dans le sous-corpus Espagne, 53,53 % ($\pm 0,34$ %) dans Argentine, 74,49 % ($\pm 2,24$ %) dans France et 65,88 % ($\pm 1,66$ %) dans Québec (Figure 15).

Figure 15. Proportion moyenne dans chaque sous-corpus des actualisations des notions constituant le noyau thématique



5.1.4. Hyper-catégories

Les notions repérées dans les textes ont été classées au sein de sept hyper-catégories⁴⁸ : entité concrète (/embryon humain/), action (/guérir un patient/), entité abstraite (/maladie/⁴⁹), processus (/développement d'un embryon/), propriété (/pluripotence/), état (/stade de blastocyste/) et phénomène (/présence d'une aberration chromosomique/). Les deux premières regroupent la majorité des notions des quatre sous-corpus; les entités concrètes et les actions constituent en moyenne 77,26 % ($\pm 4,45$ %) des notions d'Espagne, 78,26 % ($\pm 4,59$ %) de celles d'Argentine, 77,06 % ($\pm 4,76$ %) de celles de France et 80,28 % ($\pm 8,16$ %) de celles de Québec. Le tableau 4 présente la proportion moyenne de chaque hyper-catégorie dans les quatre sous-corpus et la figure 16 montre leur distribution.

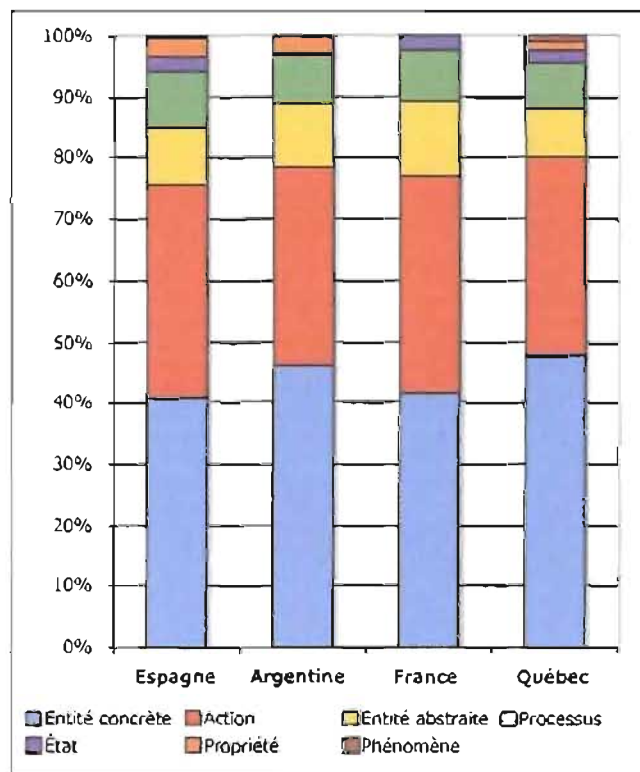
⁴⁸ Cette notion a été introduite à la section 2.3. du « Cadre théorique ».

⁴⁹ La notion /maladie/ pourrait aussi être considérée comme un processus ou un état, mais nous avons privilégié sa représentation en tant que « quelque chose » pouvant faire l'objet d'un traitement ou d'une guérison.

Tableau 4. Proportion moyenne des hyper-catégories dans les quatre sous-corpus³³

Hyper-catégorie	Espagne	Argentine	France	Québec
Entité concrète	41,44 % (± 1,53 %)	46 % (± 2,59 %)	41,58 % (± 5,43 %)	47,70 % (± 14,77 %)
Action	35,82 % (± 5,74 %)	32,26 % (± 2,53 %)	35,49 % (± 1,19 %)	32,58 % (± 6,61 %)
Entité abstraite	9,64 % (± 2,94 %)	10,61 % (± 2,22 %)	12,01 % (± 4,06 %)	8,03 % (± 8,06 %)
Processus	9,05 % (± 2,68 %)	7,83 % (± 3,02 %)	8,57 % (± 1,69 %)	7,23 % (± 3,64 %)
État	2,51 % (± 0,63 %)	0,45 % (± 0,78 %)	2,35 % (± 0,38 %)	2,14 % (± 0,26 %)
Propriété	3,19 % (± 0,90 %)	2,86 % (± 1,22 %)	0 %	1,16 % (± 1,64 %)
Phénomène	0,38 % (± 0,66 %)	0 %	0 %	1,16 % (± 1,64 %)

Figure 16. Distribution moyenne des notions des quatre sous-corpus en fonction de l'hyper-catégorie



Les entités concrètes et les actions sont aussi largement majoritaires dans les ensembles de notions communes de chaque corpus. Elles constituent 75,58 % des notions communes au corpus en espagnol, qui ne comprennent pas de phénomènes (Figure 17), et représentent 87,80 % des notions communes au corpus en français, où il n'y a ni propriétés ni phénomènes (Figure 18).

³³ Un tableau (tableau 13) avec les totaux par texte est présenté à l'Annexe I.

Figure 17. Distribution des notions communes du corpus en espagnol en fonction de l'hyper-catégorie

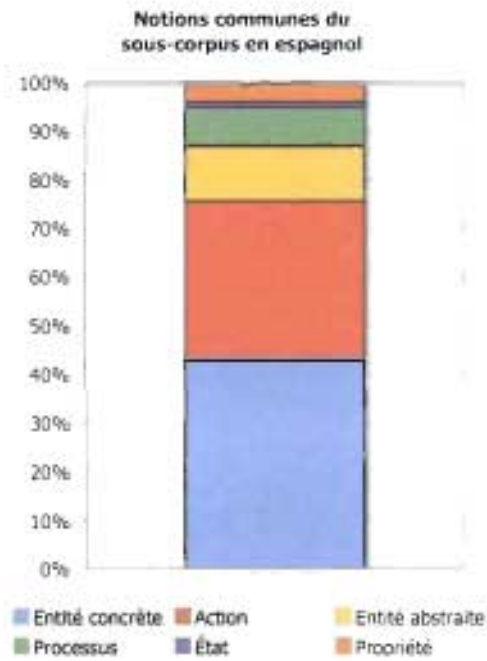
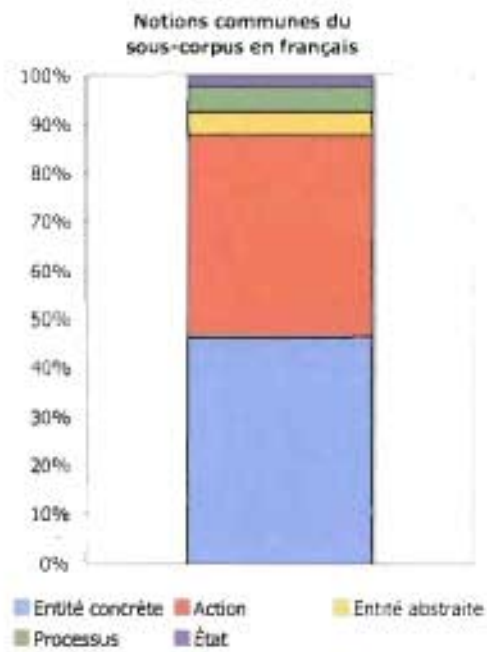
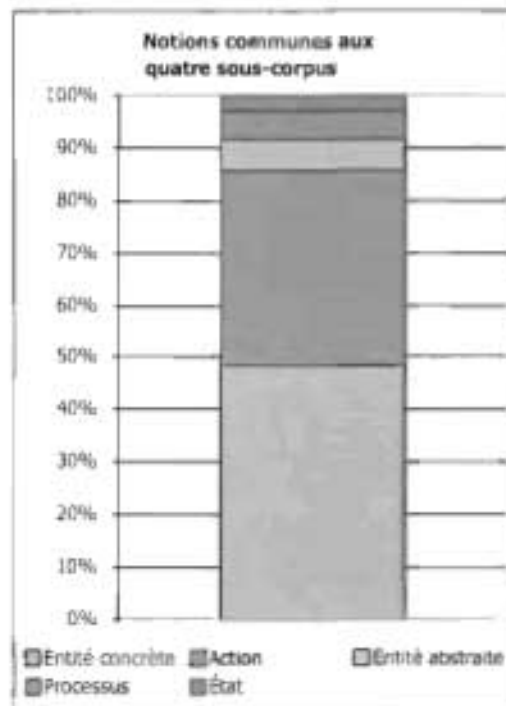


Figure 18. Distribution des notions communes du corpus en français en fonction de l'hyper-catégorie



Dans le noyau thématique, au sein duquel nous ne trouvons ni propriétés ni phénomènes, les entités concrètes et les actions comptent pour 85,71 % (Figure 19).

Figure 19. Distribution des notions en fonction de l'hyper-catégorie dans le noyau thématique



5.2. Expressions linguistiques

5.2.1. Nombre d'expressions et expressions communes

Nous avons relevé 475 unités linguistiques et syntagmes dans le corpus en espagnol, dont 267 apparaissent dans Espagne et 266 dans Argentine. Pour exprimer les 86 notions communes aux deux sous-corpus, le premier en emploie 223 et le second, 217; 58 expressions sont employées dans les deux (Tableau 5).

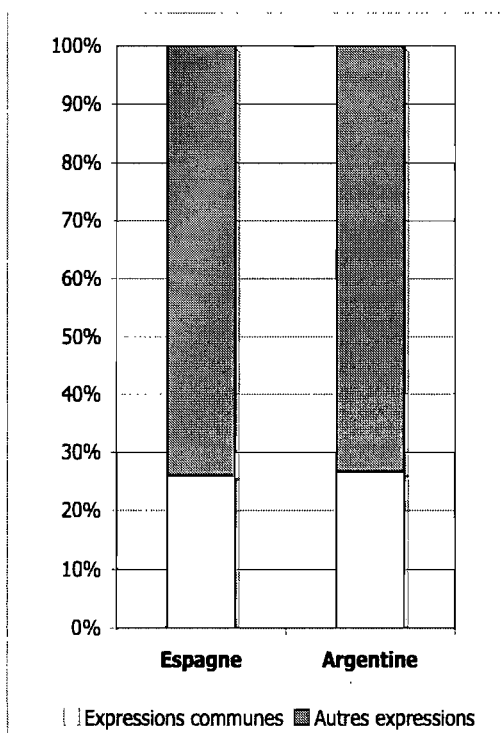
Tableau 5. Expressions communes aux sous-corpus Espagne et Argentine

Notions	Expressions
/blastocyste/	blastocito
/cellule/	celula
/cellule différenciée/	celula somática
/cellule souche/	celula madre
/cellule souche d'un embryon/	celula madre embrionaria
/créer un embryon animal par clonage/	clonación animal
/créer un embryon humain par clonage/	clonación
	clonación humana
	clonar embriones
	técnica

/créer un embryon humain par clonage dans un but reproductif/	<u>clonación reproductiva</u>
/créer un embryon humain par clonage dans un but thérapeutique/	<u>clonación terapéutica</u>
/créer un embryon par clonage/	<u>clonación</u>
/créer un embryon par clonage dans un but reproductif/	<u>clonación reproductiva</u>
/diabète/	<u>diabetes</u>
/différenciation d'une cellule souche en une cellule différenciée/	<u>diferenciación</u>
/division cellulaire/	<u>división</u>
	<u>dividirse</u>
/donner un ovocyte/	<u>donar óvulos</u>
/donner une cellule/	<u>donar células</u>
/embryon/	<u>embrión</u>
/embryon humain/	<u>embrión humano</u>
	<u>embrión</u>
/extraire un noyau/	<u>retirar el núcleo de un óvulo</u>
/foetus/	<u>feto</u>
/gène/	<u>gen</u>
/greffer un tissu/	<u>trasplante</u>
	<u>trasplantar</u>
/guide pour créer un embryon humain par clonage/	<u>receta</u>
/identité génétique/	<u>llevar el material genético del (individuo) (paciente)</u>
/lignée de cellules souches/	<u>línea de células madre</u>
	<u>línea celular</u>
/maladie/	<u>enfermedad</u>
/maladie d'Alzheimer/	<u>mal de Alzheimer</u>
/maladie de Parkinson/	<u>Parkinson</u>
/maladie dégénérative/	<u>enfermedad degenerativa</u>
/matériel génétique/	<u>material genético</u>
/membrane/	<u>membrana</u>
/naissance d'un être vivant/	<u>nacimiento</u>
/noyau/	<u>núcleo</u>
/obtenir une cellule souche d'un embryon humain créé par clonage/	<u>extraer células madre de embriones (humanos clonados) (humanos clónicos) (clones)</u>
	<u>logro</u>
	<u>resultado</u>
/organe/	<u>órgano</u>
/ostéoarthrite/	<u>osteoartritis</u>
/ovocyte/	<u>óvulo</u>
/ovule/	<u>óvulo</u>
/patient/	<u>paciente</u>
/placenta/	<u>placenta</u>
/port du matériel génétique/	<u>contener el material genético</u>
/rejet d'une entité par le système immunitaire/	<u>rechazo inmunológico</u>
/reproduction d'un être humain/	<u>reproducción humana</u>
/thérapie/	<u>terapia</u>
/thérapie à l'aide de greffes/	<u>medicina de trasplantes</u>
/tissu/	<u>tejido</u>
/type de cellule différenciée/	<u>tipo de células</u>
/utérus/	<u>útero</u>
/utiliser un ovocyte/	<u>utilizar óvulos</u>

Si nous mettons ces unités et syntagmes en perspective avec le groupe d'expressions renvoyant aux notions communes aux sous-corpus Espagne et Argentine (section 5.1.3., figure 10), nous constatons que les 58 expressions représentent 26,01 % des expressions du premier et 26,73 % des expressions du second (Figure 20).

Figure 20. Proportion des expressions communes dans l'ensemble d'expressions renvoyant à des notions actualisées dans les deux sous-corpus (Espagne et Argentine)



Dans le corpus français, nous trouvons 280 expressions, dont 177 dans France et 123 dans ceux de Québec; 124 et 78 unités linguistiques et syntagmes sont utilisés respectivement pour exprimer les 41 notions communes. Les deux sous-corpus partagent 20 expressions (Tableau 6).

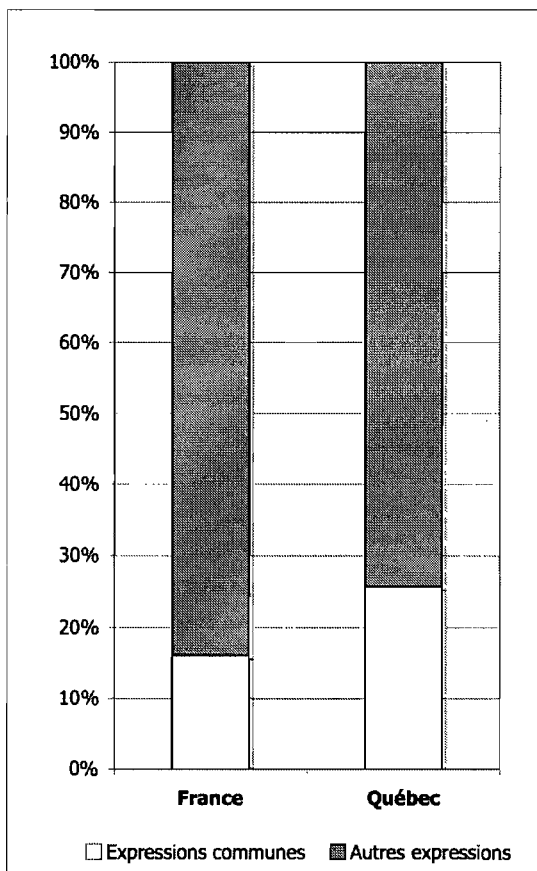
Tableau 6. Expressions communes aux sous-corpus France et Québec

Notions	Expressions
/blastocyste/	blastocyste
/cellule/	cellule
/cellule souche/	cellule souche cellule
/cellule souche d'un embryon/	cellule souche embryonnaire
/créer un embryon humain par clonage/	clonage clonage humain produire des embryons humains par clonage
/créer un embryon humain par clonage dans un but thérapeutique/	clonage thérapeutique

/créer un embryon par clonage/	clonage
/développement d'un embryon	développement se développer jusqu'au stade de blastocyste
/diabète/	diabète
/embryon génétiquement identique à un être humain/	embryon
/embryon humain/	embryon humain
/être humain créant un embryon par clonage/	cloneur
/noyau/	noyau
/organisme/	organisme
/remplacer le noyau d'un ovocyte par le noyau d'une cellule/	remplacer [le noyau de l'ovocyte] par le noyau d'une cellule
/tissu/	tissu

Ce groupe constitue 16,13 % des expressions françaises et 25,64 % des expressions québécoises renvoyant aux notions actualisées dans les deux sous-corpus (Figure 21). Nous remarquons que la proportion des expressions communes à France et à Québec est moins élevée que celle qui a été observée pour Espagne et Argentine.

Figure 21. Proportion des expressions communes dans l'ensemble d'expressions renvoyant à des notions actualisées dans les deux sous-corpus (France et Québec)



5.2.1.1. Expressions différentes

Des résultats précédents, il découle qu'une grande majorité d'expressions ne sont pas communes aux sous-corpus respectifs de chaque corpus. Ne pouvant toutes les présenter, nous fournissons deux tableaux avec les quatre exemples de divergence les plus saillants, qui sont représentatifs de l'ensemble. Ce sont les cas où l'écart entre le nombre d'expressions employées dans chaque sous-corpus pour une même notion est le plus large. Il s'agit, dans le corpus en espagnol, des notions /cellule souche/ et /blastocyste/ (tableau 7), et dans le corpus en français, des notions /blastocyste/ et /cellule souche pluripotente/ (tableau 8).

Tableau 7. Expressions employées dans le corpus en espagnol pour renvoyer aux notions /cellule souche/ et /blastocyste/

Notion	Espagne	Argentine
/cellule souche/	<ul style="list-style-type: none"> • <u>célula madre</u> 	<ul style="list-style-type: none"> • <u>célula madre</u> • <u>célula totipotencial</u> • <u>célula troncal</u> • <u>[célula] indiferenciada</u> • <u>[célula] stem</u> • <u>stem cell</u>
/blastocyste/	<ul style="list-style-type: none"> • <u>blastocito</u> • <u>blastocisto</u> • <u>embrión</u> • <u>embrión humano</u> • <u>embrión del mismo estadio que los que se han clonado ahora</u> • <u>esfera hueca formada por unas 100 células</u> 	<ul style="list-style-type: none"> • <u>blastocito</u>

Tableau 8. Expressions employées dans le corpus en français pour renvoyer aux notions /blastocyste/ et /cellule souche pluripotente/

Notion	France	Québec
/blastocyste/	<ul style="list-style-type: none"> • blastocyste 	<ul style="list-style-type: none"> • blastocyste • blastocyte • embryon aux premiers stades de son développement qui ne contient que 100 cellules
/cellule souche pluripotente/	<ul style="list-style-type: none"> • cellule • cellule souche pluripotente • cellule souche embryonnaire pluripotente • cellule pluripotente • cellule-souche capable de se différencier • cellule-souche dite « pluripotente » • objet de leur [chercheur] quête 	<ul style="list-style-type: none"> • cellule souche qui pourrait permettre d'obtenir des traitements sur mesure contre des maladies comme le diabète ou la maladie de Parkinson

Au sein de ces ensembles paraphrastiques⁵¹, certaines relations entre les expressions sont marquées. Nous trouvons des reprises diaphoriques : des hyperonymes pour l'anaphore (embrión, embrión humano, cellule), et une expression cataphorique dont l'équivalence est créée en contexte (**objet de leur [chercheur] quête**). Puis, pour la notion /cellule souche/, les auteurs proposent quelques variantes lexicales comme des synonymes à l'aide de reformulations (célula madre, célula totipotencial, [célula] indiferenciada, [célula] stem et le syntagme en anglais stem cell).

Le rapport entre les autres expressions n'est pas indiqué. Parmi celles-ci, nous remarquons une variante lexicale (célula troncal), une variante structurale (**cellule pluripotente** par rapport à **cellule souche pluripotente**) et ce qui paraît être une variante orthographique (**blastocyste** et **blastocyte**, en français, blastocisto et blastocito, en espagnol), mais qui se trouve être une confusion de la part des journalistes entre deux notions différentes : le stade du développement embryonnaire (/blastocyste/) et la cellule embryonnaire encore non différenciée (/blastocyte/). Nous trouvons aussi un hyponyme (**cellule souche embryonnaire pluripotente**), qui explicite un trait conceptuel et introduit une précision qui n'a pas été fournie ailleurs dans le corpus, ni même dans ce texte. D'autres paraphrases remplissent une fonction métalinguistique (**cellule souche dite « pluripotente »**) ou de définition sans marques (**embryon aux premiers stades de son développement qui ne contient que 100 cellules**).

Le lien de coréférence se fait par l'intermédiaire de la signification linguistique de l'expression (**cellule souche pluripotente**) ou en contexte (embrión del mismo estadio que los que se han clonado ahora). Cette dernière expression fait le lien avec la situation référentielle. Il faut également noter qu'il n'y a aucune expression commune pour la notion /cellule souche pluripotente/ : le sous-corpus québécois l'exprime par un long syntagme qui ressemble à une définition, mais l'élément défini étant absent.

⁵¹ Notion présentée à la section 2.5. du « Cadre théorique ». Nous examinerons les caractéristiques des ensembles paraphrastiques à la section 5.2.2.2.

5.2.2. Variation linguistique

5.2.2.1. Nombre d'expressions renvoyant à une même notion

Chaque notion est exprimée dans les différents sous-corpus par l'intermédiaire d'une seule expression ou de plusieurs. Or, si certaines notions suscitent peu de variation, d'autres sont à l'origine d'un grand nombre d'unités linguistiques ou syntagmes différents. Les moyennes du nombre d'expressions par notion de chaque texte ne s'avèrent pas représentatives des données : elles sont fortement influencées, d'une part, par la prédominance de notions rendues par une seule expression et, d'autre part, par le grand nombre de notions qui ne sont actualisées dans les textes qu'une seule fois. Autrement dit, la distribution du nombre d'expressions renvoyant à une notion n'est pas de type normal (courbe en forme de cloche). La représentation graphique des notions en fonction du nombre d'expressions employées pour les exprimer permet de mieux rendre compte de la dispersion.

De plus, afin d'éviter le biais introduit par la prise en compte des notions figurant une seule fois - dans ce cas, il est impossible de se prononcer quant à une éventuelle variation dans l'expression -, nous avons jugé pertinent de ne considérer que celles qui sont actualisées au moins à deux reprises (pas nécessairement par la même expression). De cette manière, on peut établir plus précisément la proportion des notions toujours exprimées par la même unité ou syntagme vis-à-vis celle des notions rendues par plusieurs expressions linguistiques. Les figures 22 à 25 montrent cette distribution au sein des textes de chaque sous-corpus. Nous appelons les textes par le nom du journal dans lequel ils ont été publiés; par exemple, dans le sous-corpus Espagne (figure 22), dans le texte publié au journal *El Mundo*, 7 notions actualisées au moins deux fois ont été exprimées par une seule expression et, à l'autre extrême, une notion a été rendue par 6 expressions différentes.

Figure 22. Distribution des notions actualisées au moins deux fois dans Espagne en fonction du nombre d'expressions employées pour les exprimer

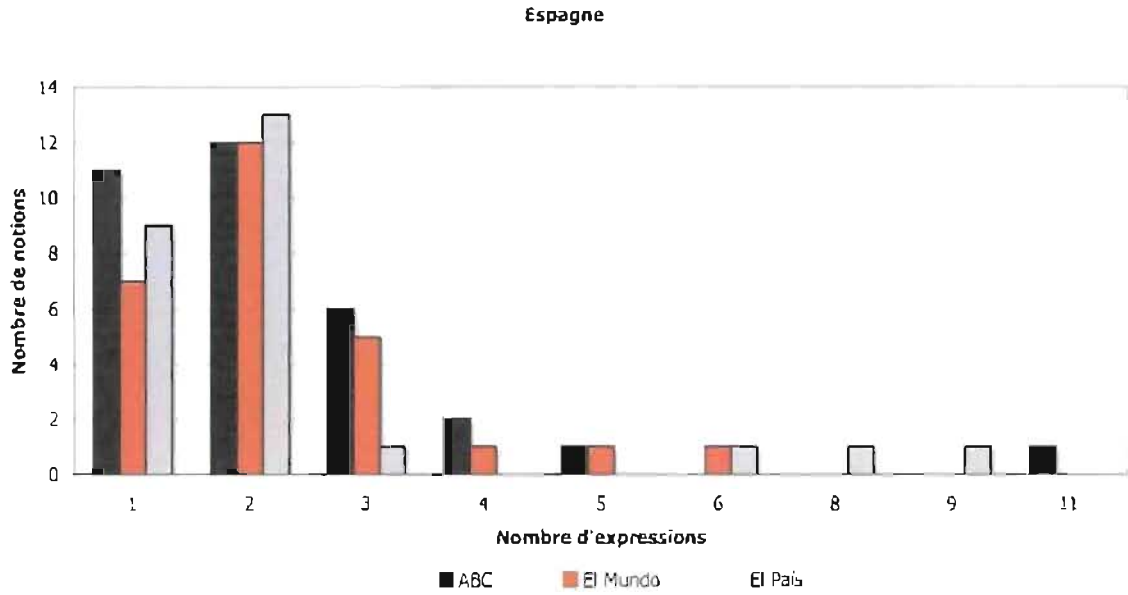


Figure 23. Distribution des notions actualisées au moins deux fois dans Argentine en fonction du nombre d'expressions employées pour les exprimer

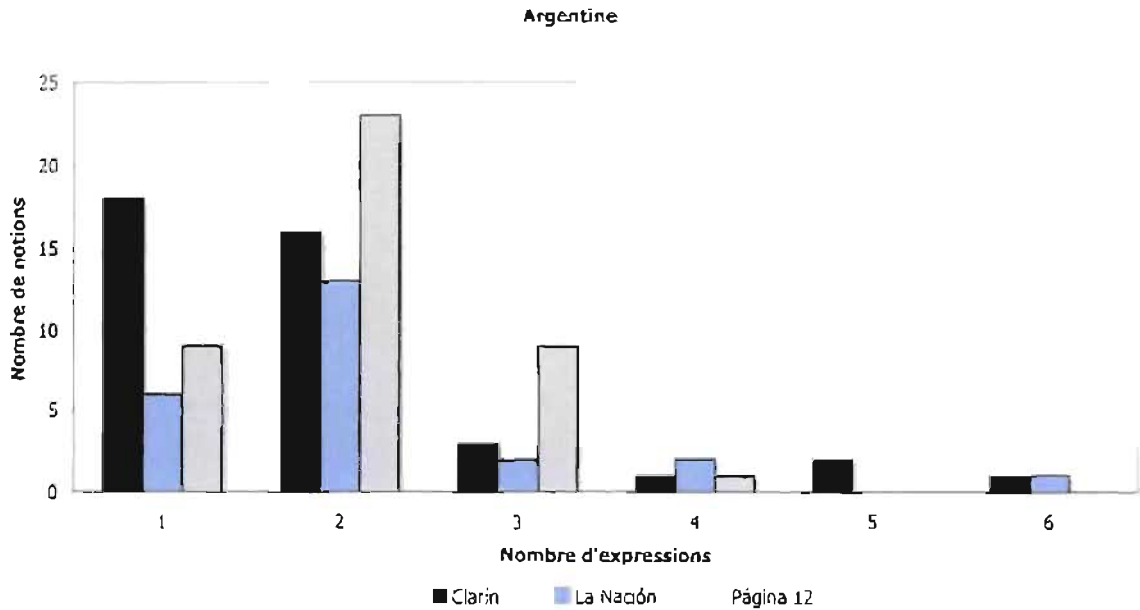


Figure 24. Distribution des notions actualisées au moins deux fois dans France en fonction du nombre d'expressions employées pour les exprimer

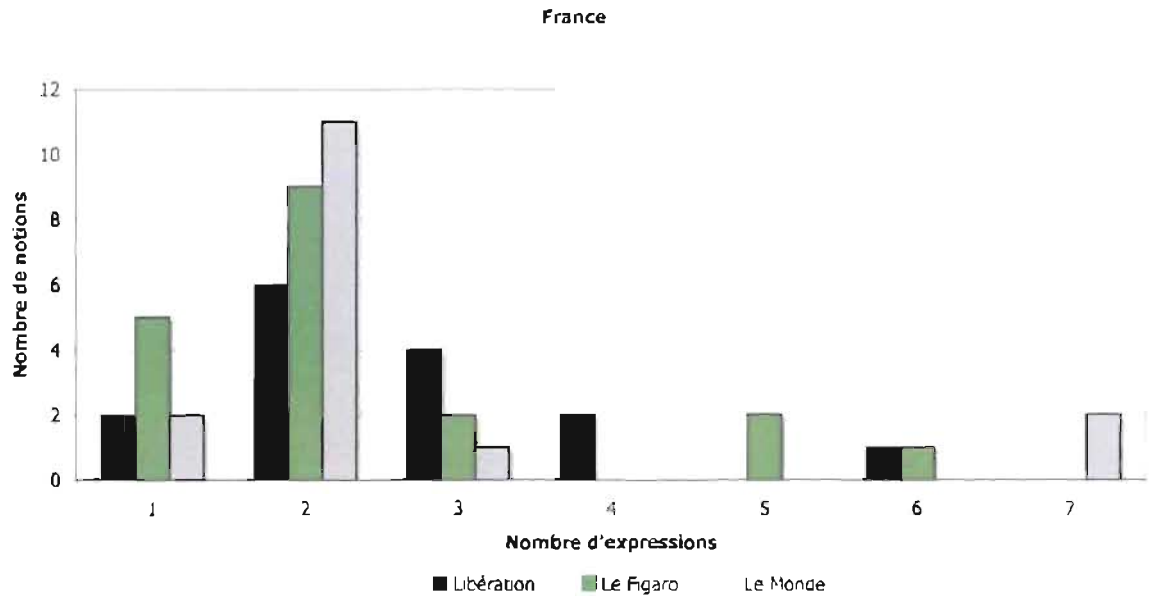
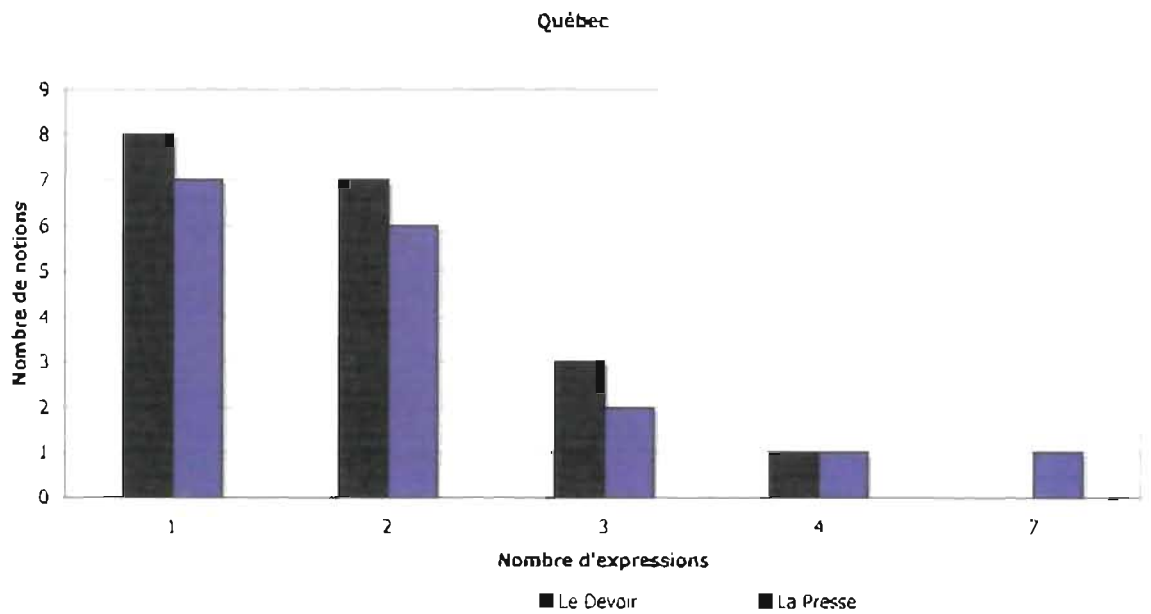


Figure 25. Distribution des notions actualisées au moins deux fois dans Québec en fonction du nombre d'expressions employées pour les exprimer



Dans tous les textes, à l'exception des textes québécois et d'un texte argentin (celui publié dans *Clarín*), nous trouvons un plus grand nombre de notions rendues par 2 expressions différentes que par une même expression. Dans tous les cas, la plupart des notions sont le plus souvent exprimées par 1, 2 ou 3 unités linguistiques ou syntagmes différents; en moyenne, 88,54 % ($\pm 0,71$ %) des notions prises en compte dans Espagne, 91,79 % ($\pm 5,23$ %) de celles Argentine, 83,90 % ($\pm 3,76$ %) de celles de France et 90,16 % ($\pm 3,98$ %) de celles de Québec. En même temps, la dispersion est assez élevée : au sein d'un même texte, jusqu'à 11 expressions ont été utilisées pour renvoyer à une même notion (/obtenir une cellule souche d'un embryon humain créé par clonage/ dans l'article du journal *ABC*).

Une distribution similaire se vérifie lorsque nous regroupons les expressions par sous-corpus et nous comparons le nombre d'expressions utilisées pour renvoyer aux notions communes (au corpus en espagnol, au corpus en français et aux quatre sous-corpus). Encore une fois, nous ne tenons compte que de celles qui y sont actualisées au moins deux fois (Figures 26, 27 et 28).

Figure 26. Distribution des notions communes au corpus en espagnol (actualisées au moins deux fois) en fonction du nombre d'expressions employées pour les exprimer

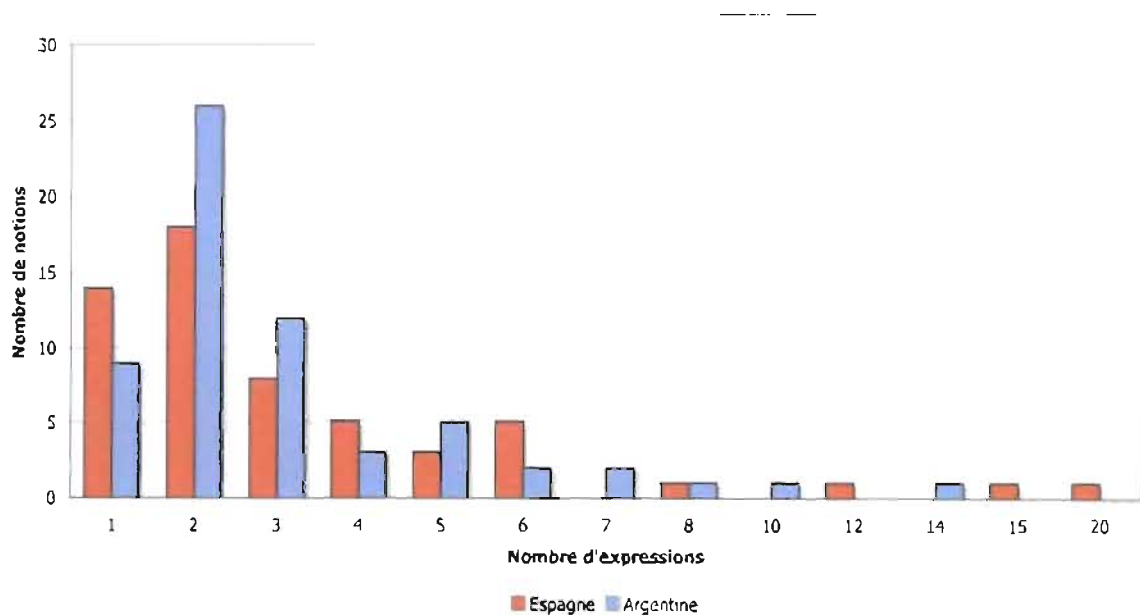


Figure 27. Distribution des notions communes au corpus en français (actualisées au moins deux fois) en fonction du nombre d'expressions employées pour les exprimer

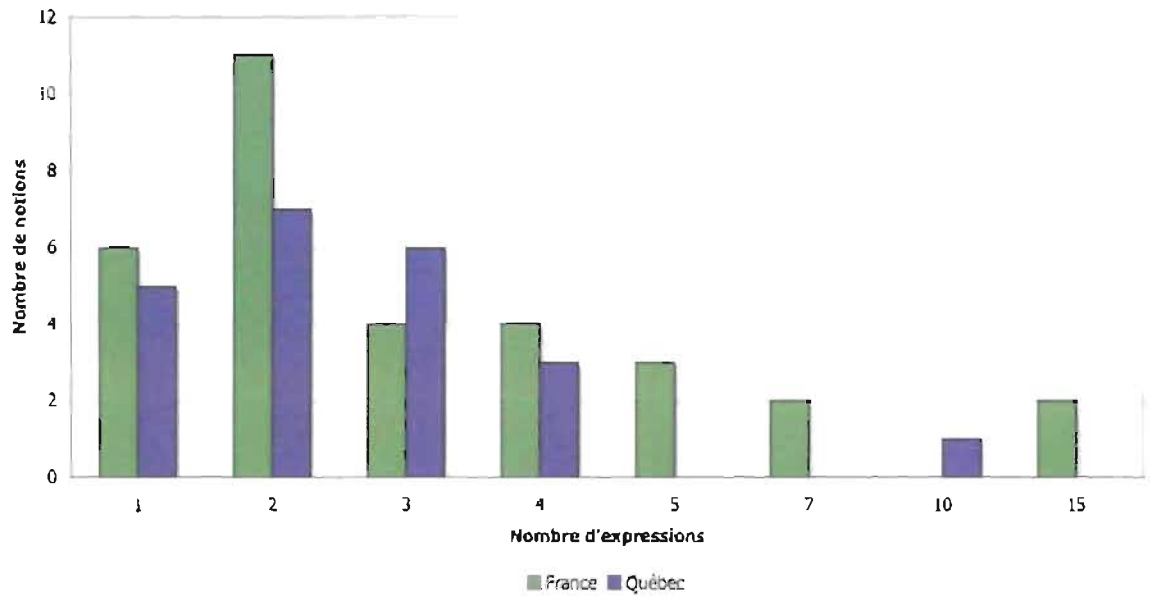
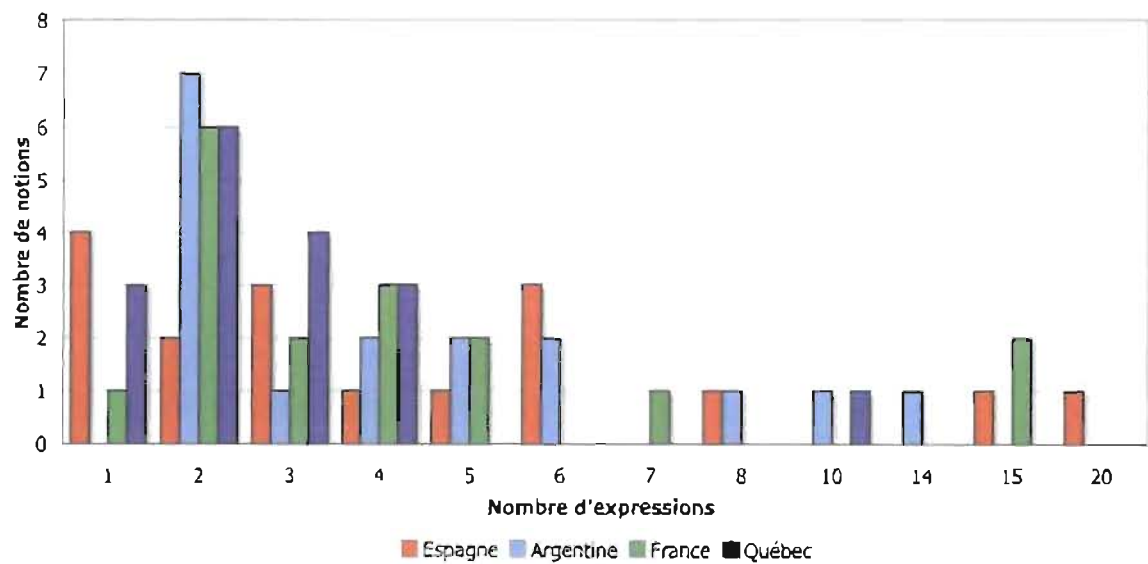


Figure 28. Distribution des notions du noyau thématique (actualisées au moins deux fois dans chaque sous-corpus) en fonction du nombre d'expressions employées pour les exprimer



Nous constatons que, même si la plus grande partie des notions sont exprimées par 1, 2 ou 3 unités linguistiques ou syntagmes, la proportion des notions rendues par 4 expressions ou plus est assez élevée (47,06 % pour Espagne, Argentine et France et 23,53 % pour Québec). La plus grande dispersion se trouve dans le sous-corpus Espagne tandis que le sous-corpus québécois semble présenter moins de dispersion que les autres. Nous trouvons jusqu'à 20 expressions pour une notion dans le sous-corpus Espagne, 14 dans le sous-corpus Argentine (la notion /obtenir une cellule souche d'un embryon humain créé par clonage/ pour les deux sous-corpus), 15 dans le sous-corpus France et 10 dans le sous-corpus Québec (dans ces sous-corpus, il s'agit de la notion /créer un embryon humain par clonage/).

5.2.2.1.1. Rapport avec le nombre d'actualisations

On pourrait être tenté d'expliquer cette dispersion au sein des sous-corpus par la différence dans le nombre d'actualisations des notions. Certes, *a priori* il semble logique de penser qu'une notion revenant plus fréquemment dans les textes donnera lieu à une plus grande variation linguistique. C'est le cas de la notion /créer un embryon humain par clonage/, qui est l'une des notions les plus fortement représentées dans les textes, et qui a été rendue par un grand nombre d'expressions. Cette notion a été actualisée 27 fois au moyen de 15 expressions dans le sous-corpus Espagne, actualisée 15 fois à l'aide de 10 expressions dans le sous-corpus Argentine, actualisée 22 fois par l'intermédiaire de 15 expressions dans le sous-corpus France et actualisée 17 fois par 10 expressions dans le sous-corpus Québec. Pourtant, la notion /ovocyte/, dont on relève 15 actualisations dans les textes du premier, 13 dans ceux du deuxième, 17 dans ceux du troisième et 8 dans ceux du quatrième, n'a été exprimée que par 1, 4, 2 et 1 expressions respectivement. Aucun des deux exemples n'étant un cas isolé, il semble difficile de postuler une corrélation entre la fréquence d'actualisation d'une notion et la variation des expressions qui y renvoient. Lorsque nous représentons à l'aide d'un graphique ce rapport pour les notions actualisées au moins deux fois dans chaque sous-corpus (Figures 29 à 32), nous observons une dispersion de la distribution des couples de valeurs trop grande pour conclure à une dépendance entre ces variables.

Figure 29. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Espagne

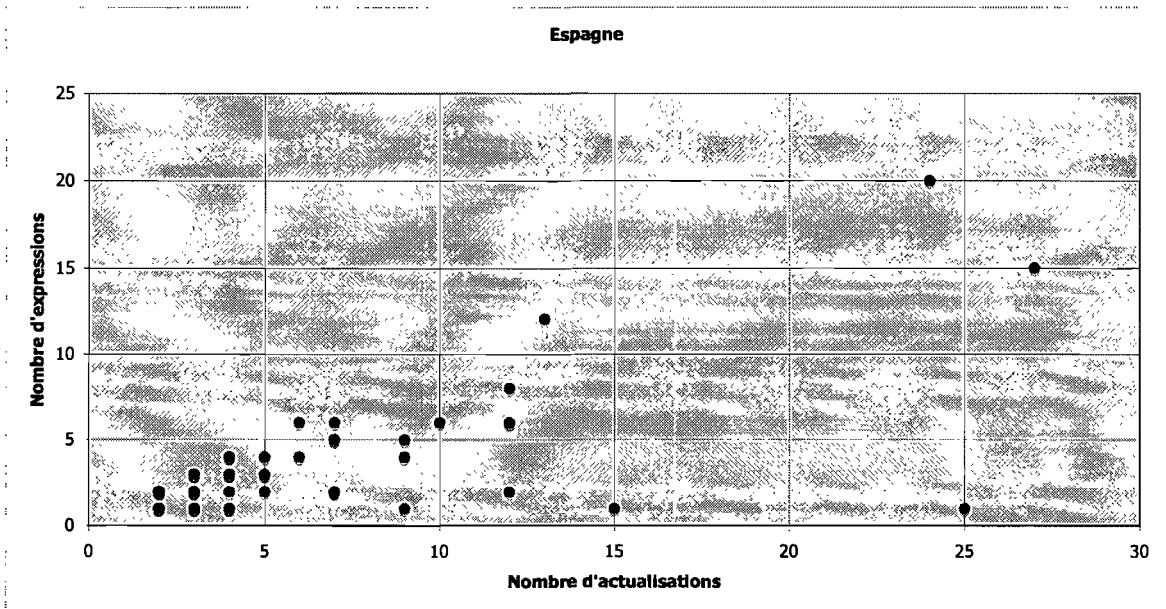


Figure 30. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Argentine

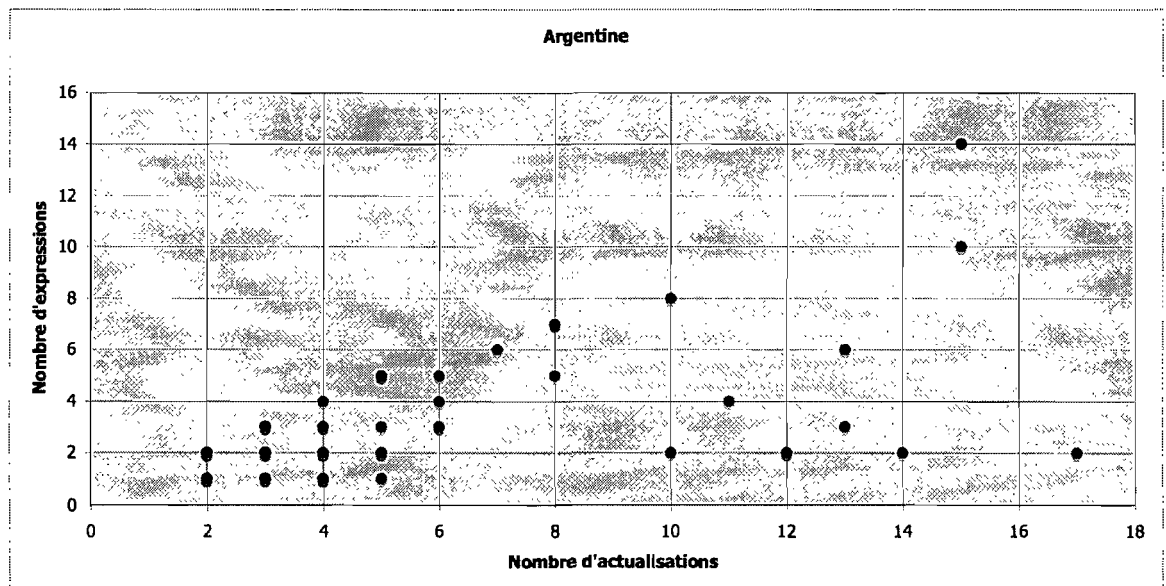


Figure 31. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus France

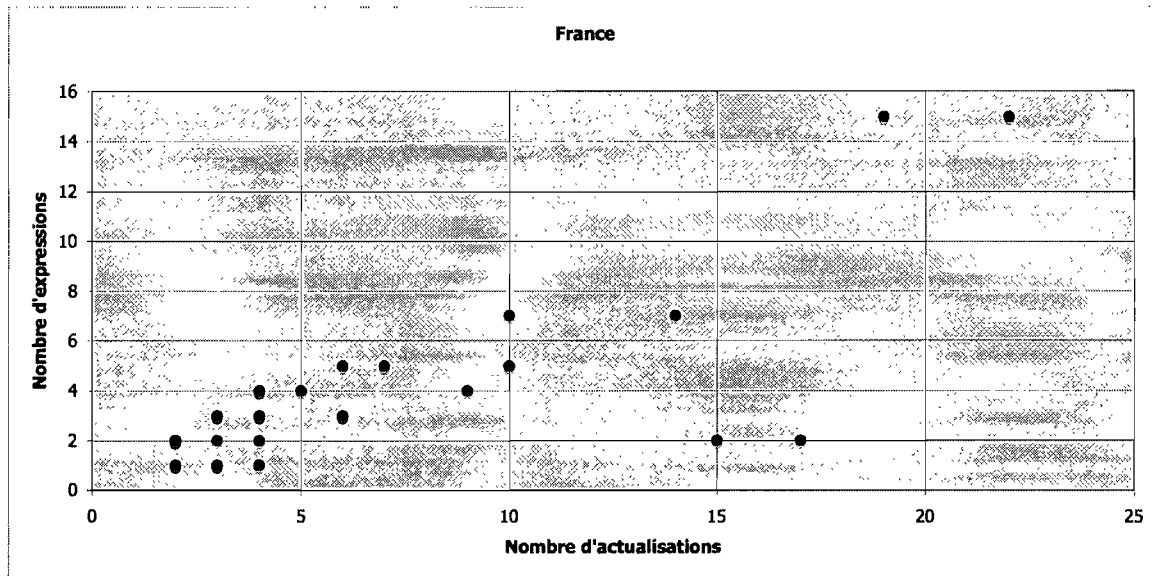
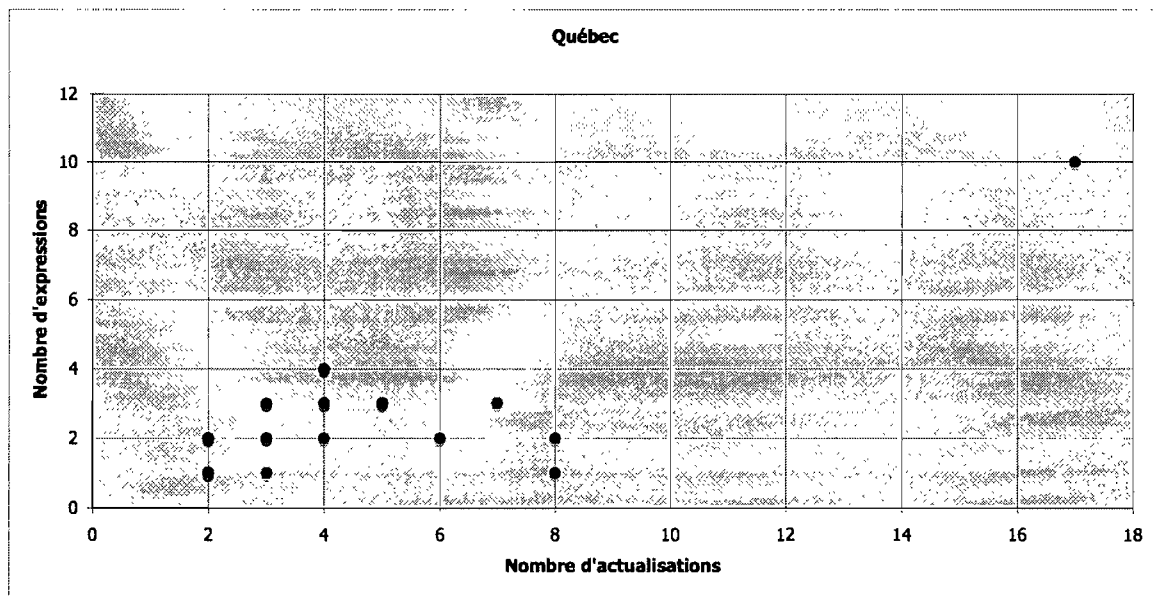


Figure 32. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Québec

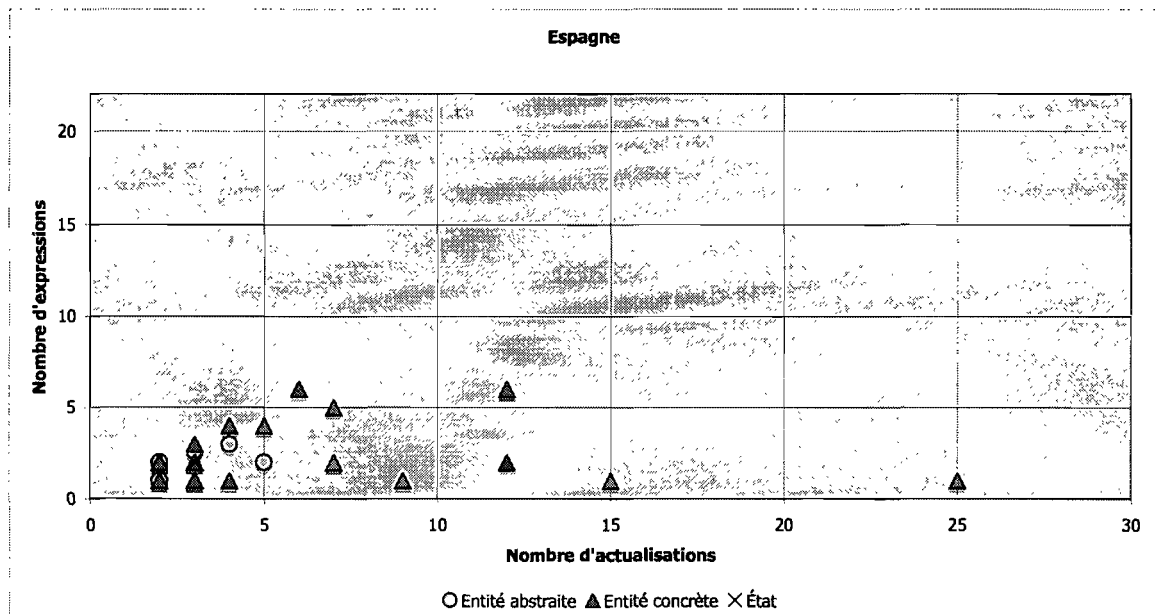
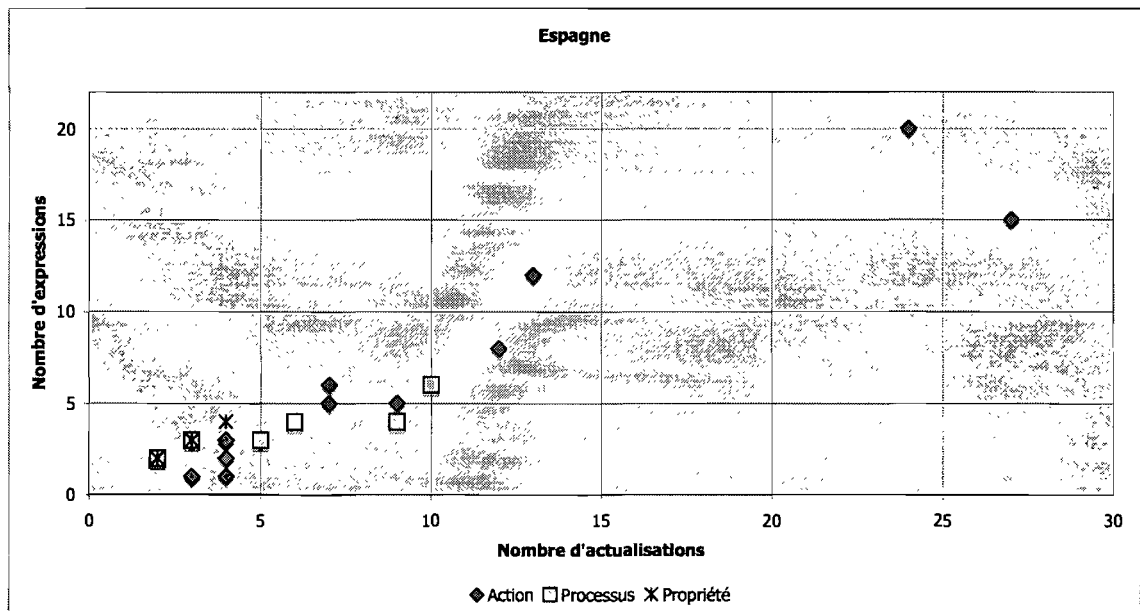


5.2.2.1.2. Rapport avec l'hyper-catégorie

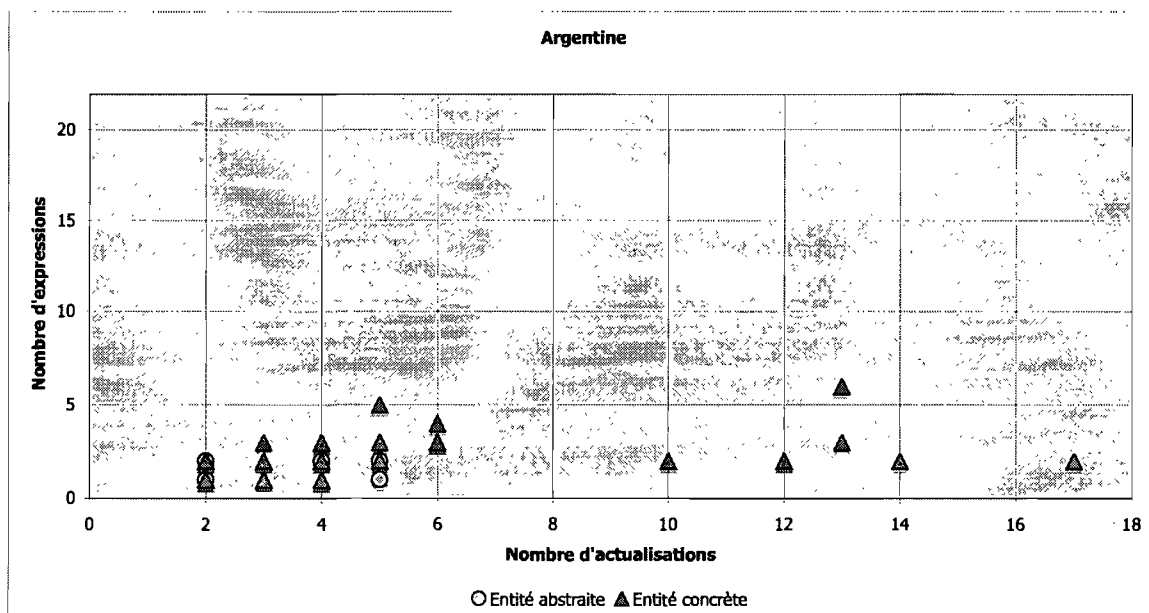
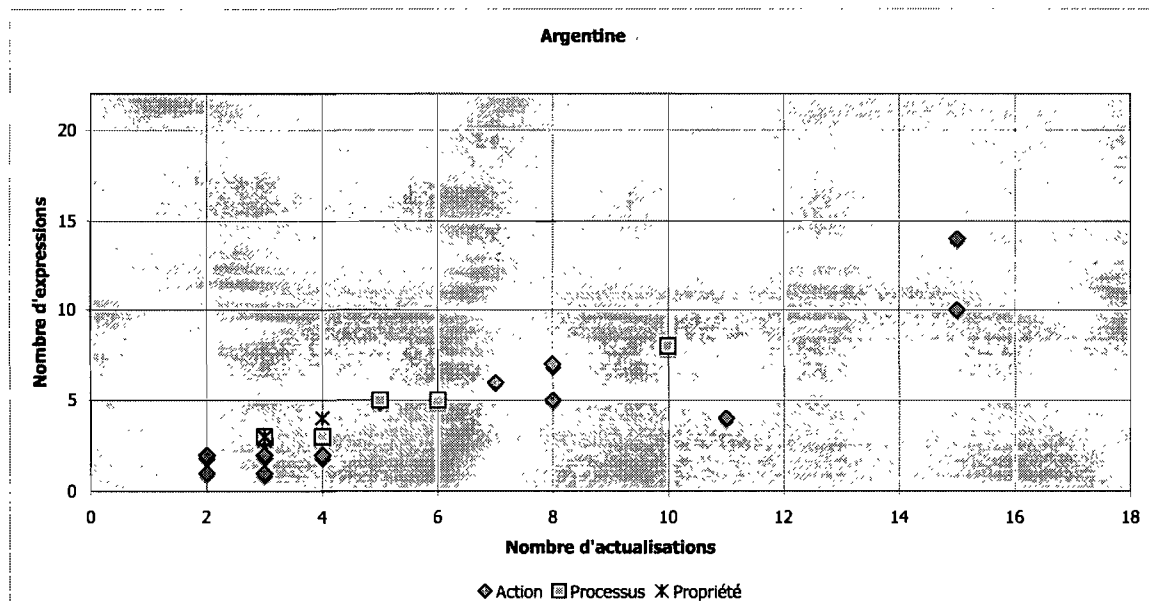
En revanche, les données incitent à penser qu'il peut exister un lien entre l'hyper-catégorie à laquelle appartient la notion et le nombre d'unités ou syntagmes linguistiques employés pour l'exprimer. Les figures 33 à 40 reproduisent les rapports illustrés précédemment, mais regroupent les notions en fonction de l'hyper-catégorie. Nous avons réparti les résultats de chaque sous-corpus dans deux graphiques par souci de clarté. Ce

regroupement par hyper-catégorie permet de dégager des régularités pour chaque groupe. Ainsi, au fur et à mesure des actualisations, les expressions renvoyant à des actions ou à des processus semblent entraîner plus de variation linguistique que celles utilisées pour rendre des entités abstraites ou concrètes. De plus, on remarque que les notions qui sont actualisées plusieurs fois et exprimées au moyen de nombreuses expressions - les points les plus éloignés dans les graphiques - dénotent des actions.

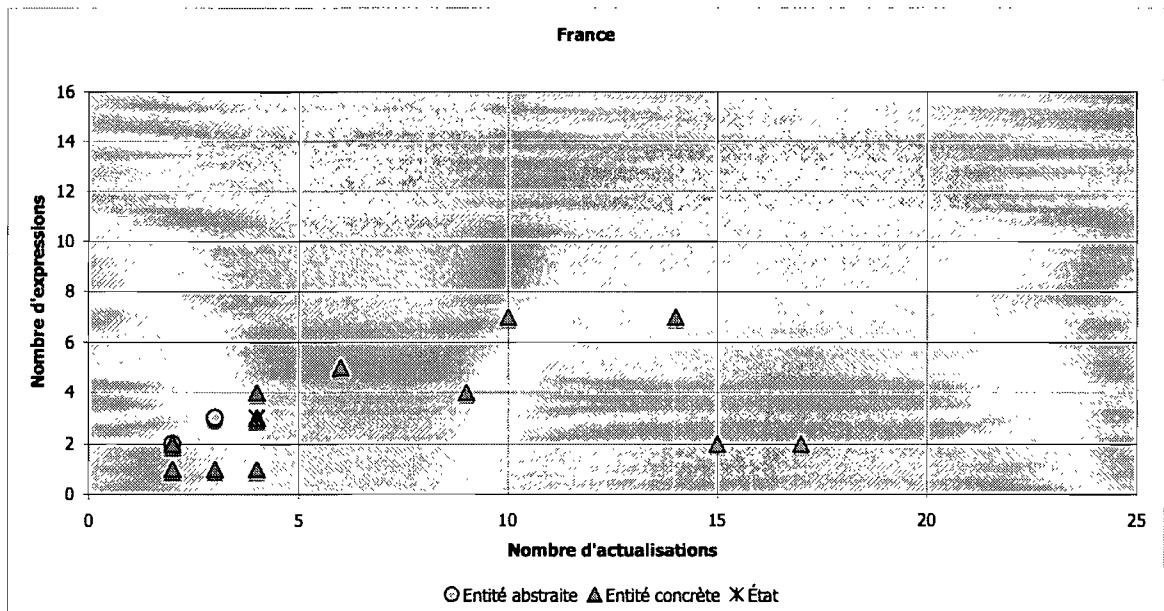
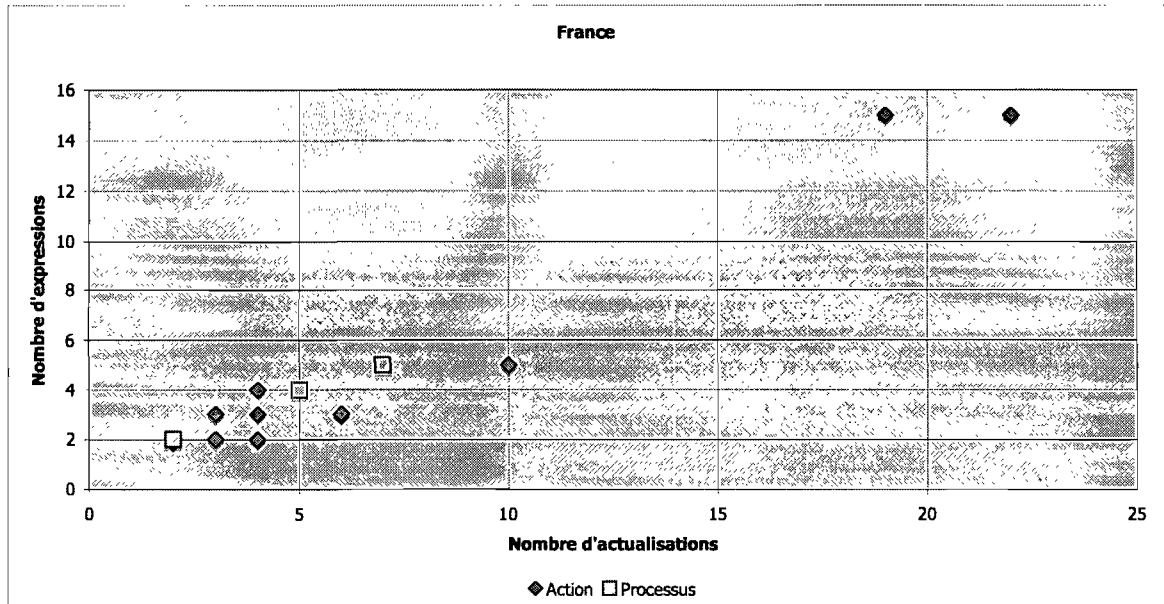
Figures 33 et 34. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Espagne (regroupement en fonction de l'hyper-catégorie)



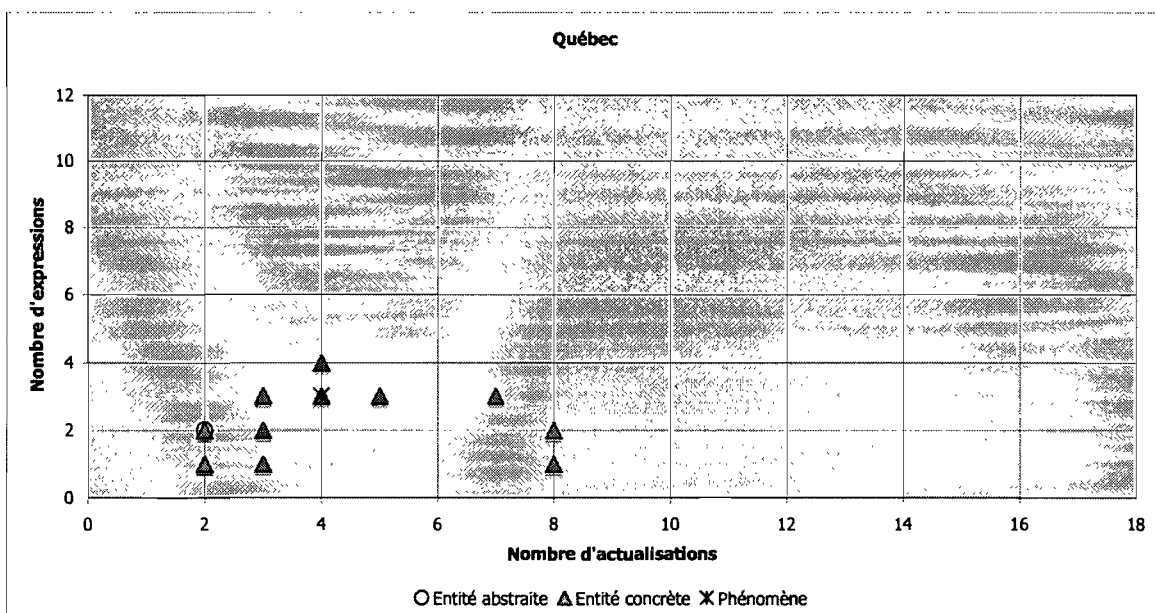
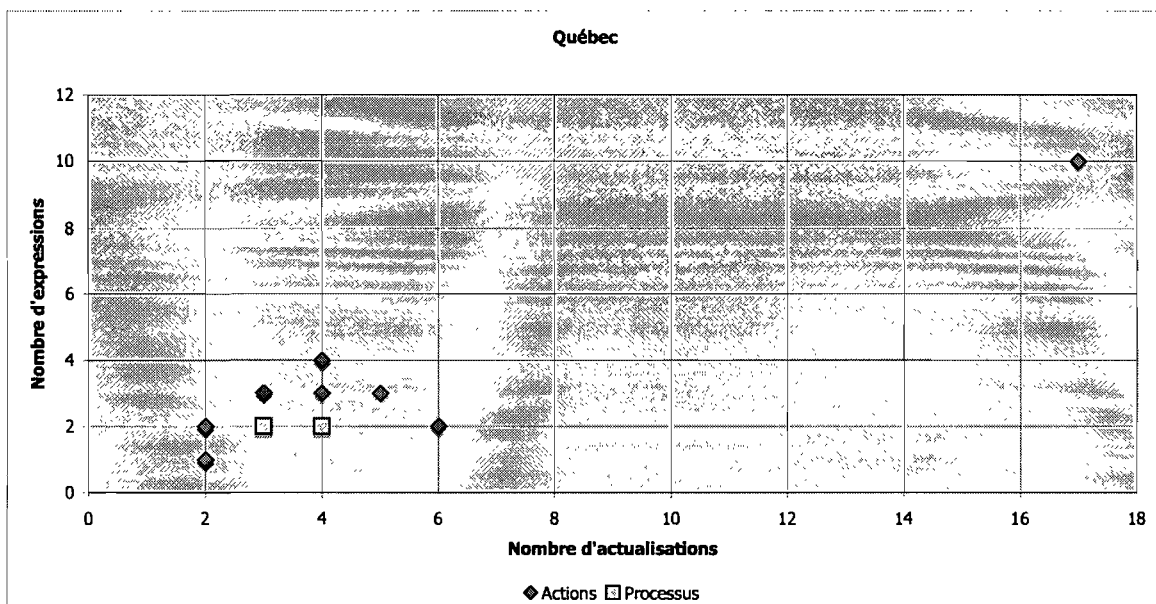
Figures 35 et 36. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Argentine (regroupement en fonction de l'hyper-catégorie)



Figures 37 et 38. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus France (regroupement en fonction de l'hyper-catégorie)



Figures 39 et 40. Rapport entre le nombre d'actualisations et le nombre d'expressions employées pour exprimer les notions actualisées au moins deux fois dans le sous-corpus Québec (regroupement en fonction de l'hyper-catégorie)



5.2.2.2. *Ensembles paraphrastiques*⁵²

Les expressions employées pour rendre une même notion (comme celles présentées à la section 5.2.1.1.) constituent un ensemble des paraphrases liées par une relation de coréférence. Suivant l'analyse proposée par Mortureux (1993), nous discriminons les paradigmes définitionnels, dans lesquels la relation est marquée linguistiquement, des paradigmes désignationnels, où elle ne l'est pas. Nous pouvons distinguer trois types de marqueurs⁵³ mettant en évidence des liens entre les unités et syntagmes renvoyant à une même notion :

- Verbe métalinguistique : **un ovule fécondé est appelé foetus à 14 jours, quand apparaît l'ébauche du système nerveux, un tejido ovárico llamado cumulus**. Ce procédé n'a pas été très utilisé dans les textes étudiés, un verbe apparaît dans les sous-corpus Espagne et Argentine (le verbe llamar(se) avec deux occurrences dans le premier et une occurrence dans le second) et un autre dans le sous-corpus Québec (une occurrence du verbe **appeler**).
- Coordination métalinguistique : **couronne ou cumulus cellulaire, sus [embrión] células, como las de todo embrión, son "totipotenciales", es decir, llevan en sí la potencialidad de transformarse en cualquier tejido del organismo**. Nous trouvons ce procédé dans deux sous-corpus : Argentine (une occurrence de o, une occurrence de es decir et une autre de digamos) et France (deux occurrences de **ou**).
- Typographie : deux points (**l'objet de leur quête : des cellules souches pluripotentes**), dont on observe une occurrence dans le sous-corpus France, une occurrence de tirets (macizo interno -una porción que normalmente da origen al feto-) dans le sous-corpus Argentine, parenthèses (citoplasma (la parte exterior al núcleo)), dont nous avons relevé trois occurrences dans le sous-corpus Argentine et six occurrences de virgules (**blatocystes [sic], un embryon aux premiers stades de son développement qui ne contient que 100 cellules, clonación, una práctica que en el estado de Michigan, donde trabaja, se pena con un millón de dólares y 10 años de cárcel**), présentes dans les quatre sous-corpus.

⁵² Les notions d'*ensemble paraphrastique*, ainsi que celles de *paraphrase*, *paradigme désignationnel* et *paradigme définitionnel*, sous-tendant cette analyse ont été définies à la section 2.5. du « Cadre théorique ».

⁵³ Nous basons cette classification sur la liste des procédés mentionnés dans Mortureux, M.-F. (1993). « Paradigmes désignationnels », *Semen*, 8, p. 117-136.

- Reprise diaphorique (notamment, l'anaphore) : estas células reprend célula primigenia que puede transformarse en cualquier tejido del organismo) tandis que **de tels résultats** évite de répéter la séquence **création, à partir de la même technique de clonage par transfert nucléaire, de trois embryons humains par la firme américaine Advanced Cell Technology**. Ces marqueurs sont les plus fréquents dans les corpus : 19 occurrences dans Espagne, 13 dans Argentine, 23 dans France, 6 dans Québec. À la différence de Mortureux, nous ne considérons ici que les marques de l'anaphore grammaticale (comme les adjectifs démonstratifs), car l'anaphore lexicale implique pour le lecteur un effort cognitif différent. Le procédé diaphorique est décrit à la section 5.3.1.

La présence de ces marques n'est pourtant pas très importante dans nos textes; la plupart des expressions constituent des paradigmes désignationnels. Autrement dit, la coréférence n'est pas marquée linguistiquement et le lecteur doit donc faire appel à ses connaissances afin de la reconnaître. C'est le cas du rapport synonymique entre enfermo et paciente, qui mobilise les savoirs linguistiques du lecteur. Des connaissances encyclopédiques sont nécessaires pour repérer le lien coréférentiel entre les syntagmes **transfert nucléaire** (dont le sens linguistique n'est pas transparent pour un lecteur non expert) et **leur [ovocyte] noyau étant remplacé par celui d'une cellule**. Dans d'autres cas, c'est le contexte qui fournit les clés pour rétablir la relation paraphrastique, sans que l'on puisse pour autant identifier des traces linguistiques spécifiques. Par exemple, il est possible d'inférer la coréférence entre fase de desarrollo adecuada para aislar células madre et fase de blastocisto à partir de la comparaison effectuée dans le texte entre l'expérience réalisée par une équipe américaine où les embryons « *no alcanzaron la fase de desarrollo adecuada para aislar células madre* » et celle menée par les chercheurs coréens où ils ont pu se développer « *hasta alcanzar la fase de blastocisto* ».

En ce qui concerne les relations de coréférence s'établissant entre les expressions, dans tous les sous-corpus nous trouvons des unités et syntagmes liés par leur signification. Il peut s'agir de synonymes, tels que **couronne** et **cumulus** ou des couples d'hyperonymes/hyponymes, comme embrión et embrión humano ainsi que d'une relation d'inclusion, comme la méronymie (línea celular, célula). Nous avons relevé également des cas où la coréférence opère localement, en contexte. Par exemple, les unités **expérience**, **exploit**, **question** et **résultat** ne possèdent pas de lien sémantique avec **clonage humain**, mais dans les textes de nos corpus, elles font partie de l'ensemble paraphrastique renvoyant à la notion /créer un embryon humain par clonage/. Certaines expressions peuvent montrer une appréciation de l'auteur sur le sujet traité (**exploit**, parmi celles citées plus haut, mais nous avons également repéré **percée**, **avance** et **éxito**, qui renvoient à la notion /obtenir une cellule souche d'un embryon humain créé par clonage/) ou

constituer un indice de conceptualisation, comme materia prima qui reprend células del cúmulus et óvulos. La relation coréférentielle en discours est notamment mise en place au moyen de stratégies servant à assurer la cohésion textuelle, telles que la diaphore, la définition et la reformulation⁵⁴.

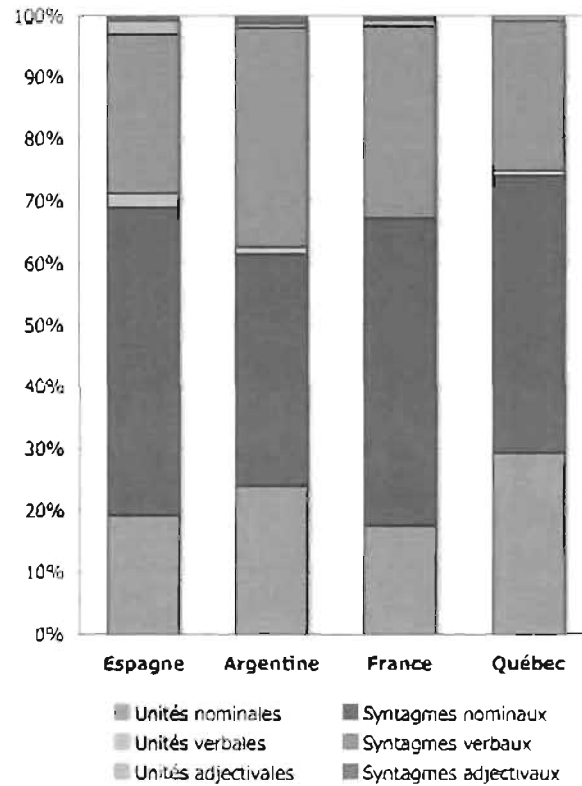
5.2.3. Structure morphosyntaxique

Parmi les expressions relevées dans les corpus, nous trouvons des structures nominales, verbales et adjectivales. En réalité, les constructions adjectivales repérées auraient pu aussi être classées sous la classe nominale ou verbale. D'une part, il s'agit des formes elliptiques : des syntagmes nominaux (nom suivi d'un adjectif qualificatif) dont la tête nominale a été effacée afin d'éviter une répétition. D'autre part, nous trouvons des formes participiales employées en apposition. Les unités et les syntagmes nominaux sont les plus nombreux : ils représentent 68,91 % des expressions d'Espagne, 61,65 % de celles d'Argentine, 67,23 % de celles de France et 73,98 % de celles de Québec. Il y a également un bon nombre de verbes et des syntagmes verbaux à l'infinitif ou conjugués qui comptent pour 28,09 % dans Espagne, pour 36,84 % dans Argentine, pour 31,07 % dans France et pour 25,20 % dans Québec. En revanche, peu de notions sont rendues par des structures adjectivales; seulement 3 % des expressions du premier sous-corpus, 1,88 % des celles du deuxième, 1,69 % des celles du troisième et 0,81 % des celles du quatrième sont des adjectifs ou des syntagmes adjectivaux.

Par ailleurs, dans tous les sous-corpus, les syntagmes nominaux (**embryon humain**) sont plus nombreux que les noms (**embryon**) et la part des syntagmes verbaux (producir 30 embriones clones) est largement majoritaire par rapport à celle des verbes (clonar), qui sont même absents du sous-corpus France. En ce qui concerne les structures adjectivales, les proportions sont variables : les adjectifs l'emportent dans Espagne, France et Québec (on ne trouve pas des syntagmes adjectivaux dans ce sous-corpus), mais sont moins nombreux dans Argentine. La figure 41 montre la distribution des expressions dans chaque sous-corpus en fonction de la structure morphosyntaxique.

⁵⁴ Nous traitons de ces mécanismes discursifs à la section 5.3.

Figure 41. Distribution des expressions des quatre sous-corpus en fonction de la structure morphosyntaxique



Lorsque l'on se penche sur les expressions renvoyant aux notions communes et au noyau thématique, on peut observer des distributions similaires. Les structures nominales prédominent, suivies des structures verbales tandis que la part des structures adjectivales reste mineure. Nous remarquons également que la proportion des syntagmes augmente en détriment des unités lexicales : on constate la diminution, voire la disparition dans certains cas, des unités verbales et adjectivales. Les figures 42 à 44 présentent les distributions en fonction de la structure morphosyntaxique des unités et syntagmes utilisés pour exprimer les notions communes au corpus en espagnol, au corpus en français et aux quatre sous-corpus.

Figure 42. Distribution des expressions renvoyant aux notions communes du corpus en espagnol en fonction de la structure morphosyntaxique

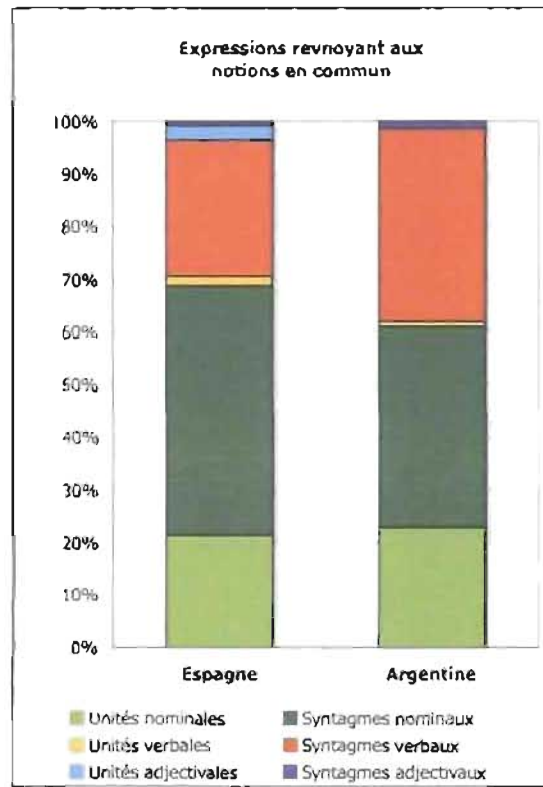


Figure 43. Distribution des expressions renvoyant aux notions communes du corpus en français en fonction de la structure morphosyntaxique

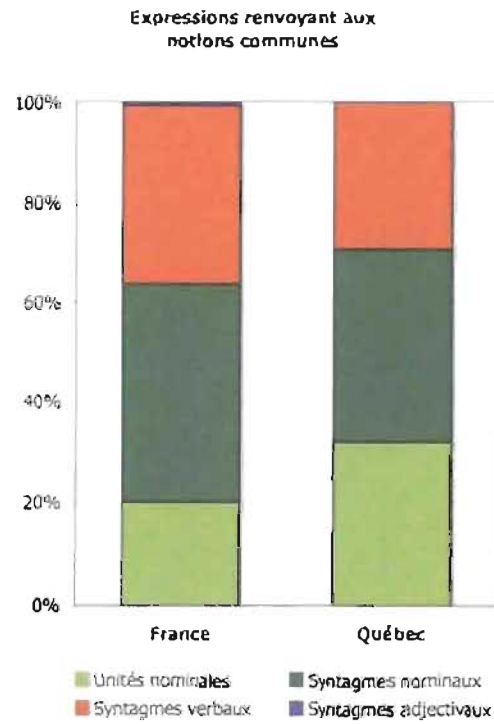
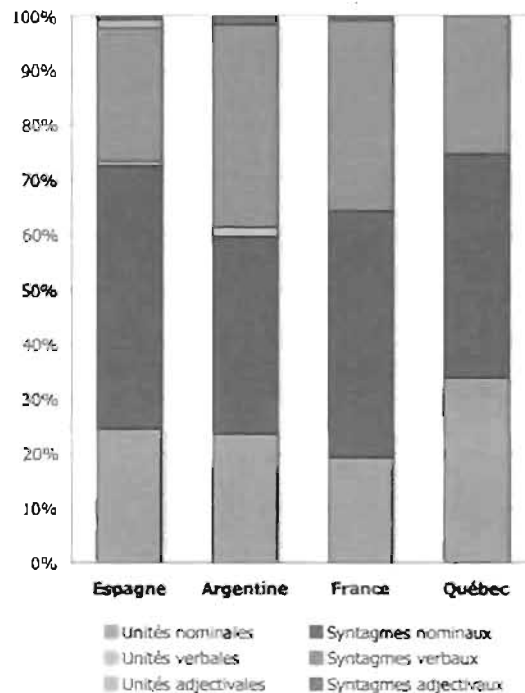


Figure 44. Distribution des expressions renvoyant aux notions communes aux quatre sous-corpus en fonction de la structure morphosyntaxique



5.2.3.1. Noms et syntagmes nominaux

Les unités et syntagmes nominaux présentent des formes diverses :

- Noms : ovaire, celula, blastocyste, técnica, exploit.
- Syntagmes nominaux composés d'un nom suivi d'un ou de plusieurs adjectifs ou d'un syntagme adjectival : insuffisance rénale, celula madre embrionaria humana, embryon capable de produire des cellules souches embryonnaires. Il peut s'agir d'un adjectif qualificatif (niño clónico), d'un adjectif relationnel (noyau cellulaire) ou d'une construction participiale (óvulo reconstruido, embryon produit par clonage).
- Syntagmes nominaux composés d'un nom suivi d'un ou plusieurs syntagmes prépositionnels : etapa de blastocito, traitement contre la stérilité, creación de embriones con fines exclusivamente científicos. En français, parfois le syntagme prépositionnel a été rendu par un adjectif possessif : leur [cellule souche embryonnaire] prélèvement.
- Syntagmes nominaux composés d'un nom suivi d'une subordonnée relative : technique qui vise à soigner un malade avec des cellules souches dérivées de son clone embryonnaire, celula que en un embrion normal está destinada a constituir la placenta.

Ces constructions se trouvent parfois combinées dans des formes plus complexes : célula primigenia que puede transformarse en cualquier tejido del organismo (nom suivi d'un adjectif suivi d'une subordonnée relative), **technique optimisée d'obtention de blastocystes humains clonés** (nom suivi d'un adjectif suivi d'un syntagme prépositionnel), célula del cúmulus donada (nom suivi d'un syntagme prépositionnel suivi d'un adjectif). Par ailleurs, il peut s'agir d'unités non prédicatives, comme **cellule**, ou prédicatives, telles que clonación. Dans les syntagmes, dont la structure est prédicative, le prédicat est un adjectif (*crónica* dans enfermedad crónica) ou un nom (*destruction* dans **destruction de l'embryon**) et, lorsque le ou les actants sont présents, ils sont réalisés par des noms, comme dans les exemples précédents, ou évoqués par un adjectif relationnel (*cellulaire* dans **noyau cellulaire**).

5.2.3.2. *Verbes et syntagmes verbaux*

En ce qui concerne les constructions verbales, si parfois seule la base prédicative verbale est employée pour renvoyer à une notion (clonar, matar), dans la plupart des cas, on utilise un syntagme constitué d'un verbe et d'unités réalisant tous ou certains de ses actants. Ces structures prédicatives comportent entre un et trois actants; plus rarement, quatre. Parmi les classes du premier actant nous trouvons surtout des êtres humains (notamment les chercheurs : **biologiste coréen, investigador**), et des entités biologiques (tejido del páncreas, ovule), mais aussi une action (retirar el núcleo de los óvulos) dans Espagne.

Lorsqu'il appartient à la classe des chercheurs, assez fréquemment ce premier actant n'est pas actualisé. On fait appel à des constructions à l'infinitif (**produire par clonage des embryons humains, obtener seres completos**) ou, particulièrement en français, au passif (**chacun d'entre eux [ovocyte] a été énucléé**) et, en espagnol, à des tournures impersonnelles (el blastocito se implanta en el útero de una mujer) ainsi qu'au sujet implicite (clonaron embriones humanos). Cette classe d'actants, réalisée linguistiquement ou non, restreint le deuxième actant à la classe des entités biologiques (los científicos han utilizado óvulos), des êtres humains (**créer des bébés**), des êtres vivants (clonar primates), des actions (**effectuer toutes sortes de clonages**), des processus (indujeron su [célula del macizo interno] diferenciación) et des entités abstraites (**créer des traitements médicaux**). C'est également lorsque le premier actant appartient à la classe des chercheurs que l'on observe un troisième et jusqu'à un quatrième actant (**les chercheurs coréens ont injecté cette culture de clones à des souris, por clonación, obtienen células madre para curar**) dans Argentine.

Dans les expressions où le premier actant appartient à la classe des entités biologiques, le deuxième actant est réalisé par une unité ou syntagme désignant une entité biologique (transformarse [célula] en cualquier tejido del organismo), un état (**il [embryon] est arrivé au stade du blastocyste**), une propriété (**los embriones eran clónicos**) ou un processus (**des agents chimiques ont lancé la division cellulaire**). En espagnol, où la diversité des classes est encore plus grande, il peut s'agir d'une action (esas células madre se originaron por la técnica de clonación), d'une entité abstraite (que [tejido del páncreas] curarían su [persona diabética] enfermedad) ou d'un être humain (las células provendrían de los mismos pacientes que buscan curas), dans Argentine, ainsi que d'un phénomène (sufrir [embrión] aberraciones cromosómicas) dans Espagne. Enfin, certaines structures prédicatives ne sont constituées que par un prédicat et un premier actant pouvant appartenir à la classe des entités biologiques (se reproducen [célula troncal] indefinidamente) ou, dans Québec, à celle des êtres humains (**le bébé clone est-il né?**).

5.2.3.3. *Adjectifs et syntagmes adjectivaux*

Il convient de rappeler (section 5.2.3.) que les structures adjectivales peuvent être des formes elliptiques (des syntagmes nominaux dont la tête nominale a été élidée, comme [clonación] reproductiva) ou des formes participiales en apposition ([noyau] remplacé par celui d'une cellule issue de la même donneuse). Nous pouvons remarquer un phénomène propre au corpus en espagnol où l'on évite la répétition d'un nom à l'aide d'une construction elliptique ([célula] adulta). À la différence d'une énumération, où une succession d'adjectifs ou de syntagmes adjectivaux partagent une même tête nominale (enfermedades respiratorias, circulatorias e inmunológicas), dans ces cas, la forme elliptique ne suit pas une forme « pleine », comme le montre l'exemple suivant :

« *Las [células] **externas** son las que más tarde formarán la placenta.* »

Si l'on devait traduire ce type de phrases, la syntaxe française exigerait l'ajout d'une tête nominale.

5.2.3.4. *Constructions elliptiques et formes réduites*

Il est nécessaire de distinguer les constructions elliptiques des formes dues au phénomène de la réduction terminologique⁵⁵. Si toutes les deux servent en principe à éviter la répétition d'une unité lexicale, les secondes présentent un plus haut degré d'autonomie et peuvent figurer dans un texte sans que la forme complète s'y trouve nécessairement. C'est le cas des noms de maladies **Alzheimer** et Parkinson. Quant aux formes elliptiques, en plus du phénomène observé dans le corpus en espagnol, circonscrit aux structures adjectivales, nous observons des structures nominales présentant des élisions. La tête nominale du syntagme est élidée lorsque le syntagme immédiatement précédent la comporte déjà, lors des énumérations. Ainsi, **cellules nerveuses** et **cellules rénales** deviennent **nerveuses** et **rénales** dans le passage :

« **Mais ce sont ces fameuses cellules souches embryonnaires pluripotentes qui pourraient dans un milieu de culture adéquat se différencier ultérieurement en cellules cardiaques, nerveuses ou rénales [...]** »

L'élément élidé peut aussi être l'expansion du syntagme nominal, comme dans le passage suivant, où línea de células madre devient línea :

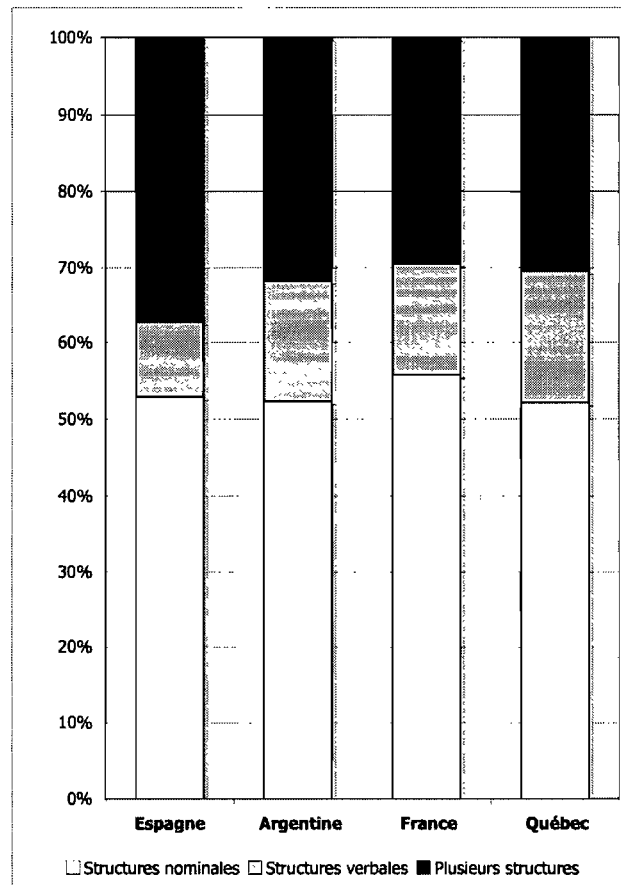
« Al retirar las células del macizo interno, y cultivarlas en el laboratorio, los científicos en Corea del Sur hicieron que se convirtieran en células madre. Formaron una línea que sería capaz de reemplazar hueso, músculo, células cerebrales dañadas, o cualquier célula del cuerpo, [...] »

5.2.3.5. *Structure morphosyntaxique et variation linguistique*

Au sein des ensembles paradigmatiques, il semble exister une certaine stabilité quant à la structure morphosyntaxique des expressions. En effet, chaque notion est très souvent exprimée par des unités ou des syntagmes appartenant à une seule partie du discours (le nom ou le verbe). Afin d'éliminer le biais introduit si nous tenons compte des notions rendues par une seule expression, nous avons ciblé l'analyse sur celles pour lesquelles au moins deux expressions avaient été relevées. La proportion de notions toujours exprimées par des structures d'une même partie du discours est de 62,75 % dans Espagne, de 68,25 % dans Argentine, de 70,59 % dans France et de 69,57 % dans Québec (figure 45).

⁵⁵ Nous référons à la définition de réduction terminologique de Collet, T. (2000). *La réduction des unités terminologiques complexes de type syntagmatique*, Thèse de Doctorat, Université de Montréal, Montréal. : « *Les produits de la reprise anaphorique présentent un degré d'autonomie contextuelle presque nul, à l'inverse des produits de la réduction à caractère lexical qui peuvent être exportés tels quels vers d'autres textes (ou contextes) sans qu'il soit nécessaire de recréer le lien réductionnel* » (p. 6 et 7).

Figure 45. Distribution des notions en fonction de la partie du discours des expressions employées (notions exprimées par au moins deux expressions)



Si nous discriminons les notions selon le nombre d'expressions qui y font référence, nous constatons que la proportion de celles qui sont toujours exprimées par la même partie du discours augmente parmi les notions exprimées par 2 ou 3 unités ou syntagmes (70,59 % dans Espagne, 76,60 % dans Argentine, 73,91 % dans France et 78,95 % dans Québec), tandis qu'elle diminue parmi les notions rendues par 4 expressions ou plus (47,06 % dans Espagne, 43,75 % dans Argentine, 63,64 % dans France et 25 % dans Québec).

5.2.3.6. *Partie du discours et hyper-catégorie*

La partie du discours à laquelle appartiennent les expressions est liée avec l'hyper-catégorie sous laquelle on peut classer les notions⁵⁶. Ainsi, on observe que les structures nominales servent à rendre toutes les hyper-catégories notionnelles à l'exception des phénomènes, principalement les entités concrètes, les entités abstraites et les états. Ce

⁵⁶ Cette relation a été étudiée en profondeur en linguistique. Pour une vision à partir de la sémantique cognitive, consulter Cordier, F. et J. François (dirs.) (2002). *Catégorisation et langage*, Paris: Hermès Science Publications.

sont surtout les déverbaux qui font référence aux actions et aux processus. Les structures verbales sont réservées aux notions dénotant des actions, des processus, des propriétés, des phénomènes et, en moindre mesure, des états (/indifférenciation/ dans le sous-corpus Espagne). Les figures 46 à 49 montrent la distribution dans chaque sous-corpus des expressions rendant compte de chaque hyper-catégorie notionnelle en fonction de la partie du discours. Dans les figures, la proportion des structures adjectivales n'est pas discriminée.

Figure 46. Distribution dans Espagne des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours

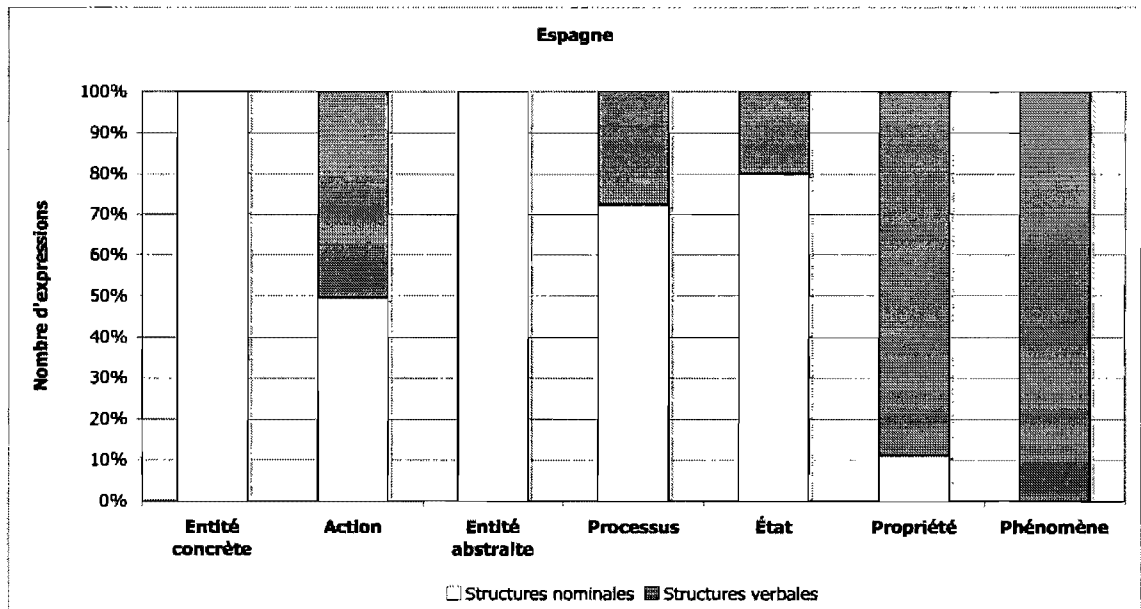


Figure 47. Distribution dans Argentine des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours

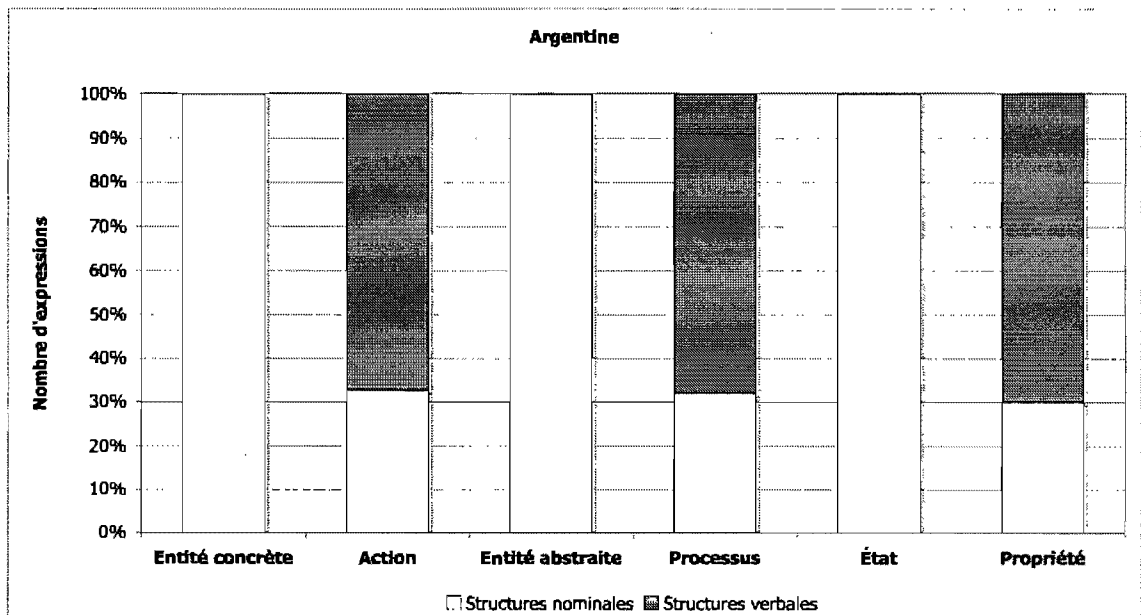


Figure 48. Distribution dans France des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours

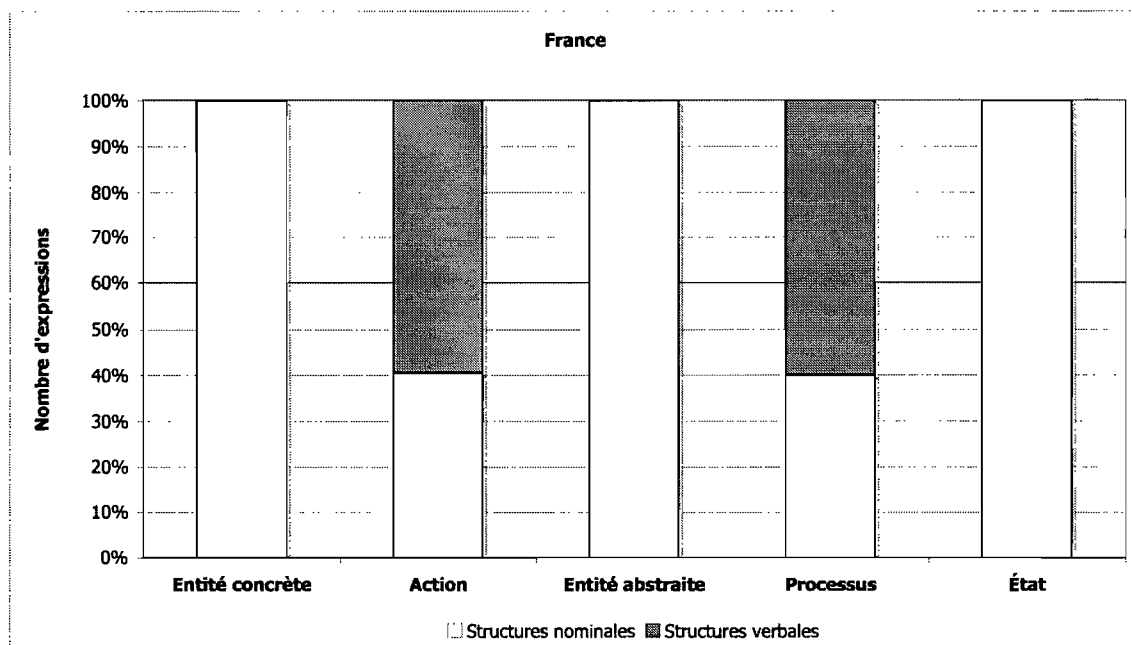
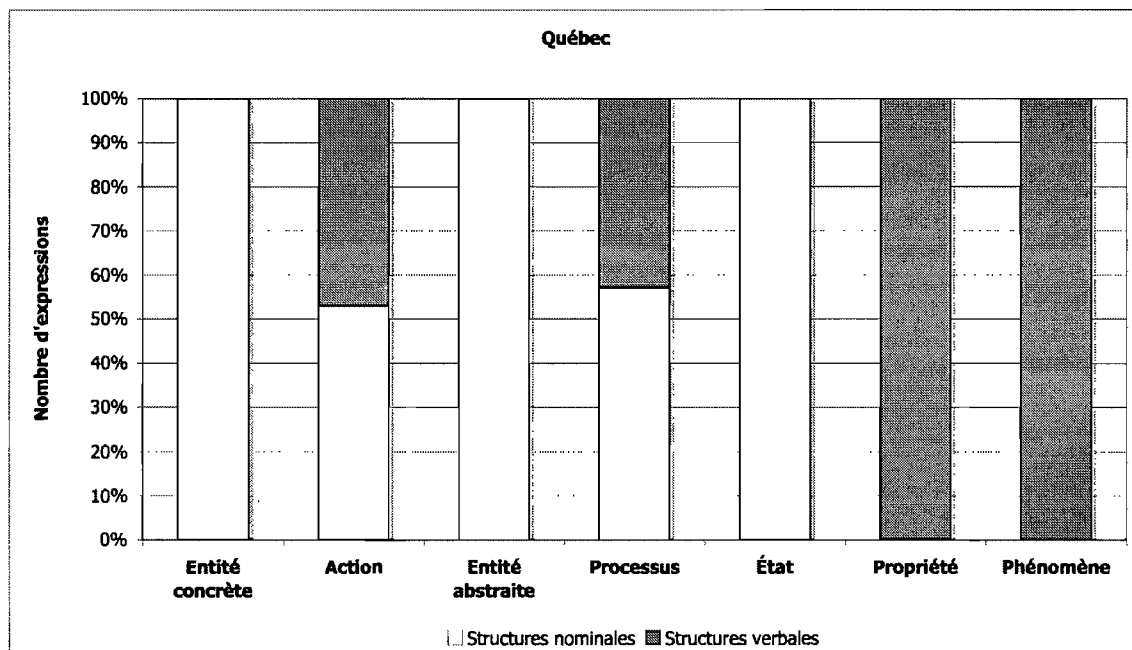


Figure 49. Distribution dans Québec des expressions renvoyant à chaque hyper-catégorie en fonction de la partie du discours



5.3.Stratégies discursives

La présentation de notions spécialisées en discours fait appel à plusieurs mécanismes discursifs. Nous nous intéressons spécifiquement à ceux qui permettent de structurer les connaissances par la mise en relation des expressions employées pour les véhiculer. Dans cette section, nous examinons la place de la diaphore, de la reformulation, de la définition et de l'explication dans nos corpus.

5.3.1.Diaphore

Des procédés considérés, celui de la diaphore est le plus utilisé dans nos corpus. Nous ne trouvons cependant que 32 reprises diaphoriques dans Espagne (ce qui touche 7,51 % des actualisations de ce sous-corpus), 21 dans Argentine (4,86 %), 28 dans France (10,56 %), dont une cataphore, et 12 dans Québec (7,06 %). Le lien coréférentiel entre les deux expressions peut être inscrit en langue ou instauré en discours. Ainsi, les rapports existant entre les expressions diaphoriques et leurs antécédents sont de nature diverse. Nous observons dans les quatre sous-corpus :

- L'hyponymie : L'antécédent peut être repris par un hyperonyme, comme dans l'exemple cité, où **embryon** reprend **embryon humain**, mais aussi par un hyponyme : embrión clon pour embrión (6 cas dans Espagne, 7 dans Argentine, 8 dans France et 1 dans Québec).
- La recatégorisation : Dans ce cas, la reprise est effectuée au moyen de ce qu'on pourrait appeler un « faux hyperonyme »⁵⁷. L'auteur propose ainsi une recatégorisation de la notion. Par exemple, reprendre **clonage humain** par la **question** permet de situer le clonage dans sa dimension sociale (13 cas dans Espagne, 5 dans Argentine, 6 dans France et 5 dans Québec).
- Appréciation : Comme nous avons avancé à la section 5.2.2.2., certaines expressions anaphoriques introduisent un jugement appréciatif. Généralement, il s'agit d'appréciations positives de l'expérience des chercheurs coréens. Ainsi, **cette percée** qualifie positivement l'action exprimée par **produire par clonage des embryons humains** (5 cas dans Espagne, 2 dans Argentine, 2 dans France et 3 dans Québec).

⁵⁷ Nous empruntons cette notion à Vandaele (communication personnelle).

- Résumé : L'anaphore peut également servir à condenser des informations fournies à différents endroits du texte ou au sein d'un passage afin d'assurer la progression thématique. C'est la fonction de la experiencia, qui reprend une description détaillée de la technique de clonage (2 cas dans Espagne, 3 dans Argentine, 5 dans France et 1 dans Québec).

Deux autres relations présentes dans les deux corpus (mais seulement dans trois sous-corpus) sont :

- La répétition : Dans certains cas, l'unité ou le syntagme repris est répété et le rapport anaphorique est signalé par un adjectif démonstratif, comme **ces cellules souches embryonnaires pluripotentes** qui reprend **cellules souches embryonnaires pluripotentes** (1 cas dans Espagne, 1 dans Argentine et 3 dans France).
- La refocalisation : Les expressions anaphoriques parfois focalisent un aspect différent de la notion. On emploie, par exemple, **ces embryons clonés** pour reprendre **embryons humains** qui, dans la phrase précédente, fait référence aux embryons obtenus par clonage par les chercheurs coréens (1 cas dans Espagne, 2 dans France et 2 dans Québec).
- La paraphrase : La reprise peut consister en une paraphrase de la signification ou du sens en contexte de l'antécédent. Un exemple de la première est la clonación de embriones qui reprend obtención de 30 embriones humanos clonados, et un exemple de la seconde est **objet de leur [chercheur] quête** comme reprise de **cellule souche pluripotente** (2 cas dans Espagne, 1 dans Argentine et 1 dans France).

Enfin, nous pouvons distinguer des liens de méronymie et holonymie (entre célula et línea de células), dans Espagne et Argentine, et la reprise par un indice de conceptualisation (materia prima reprend célula del cúmulus), qui est un cas particulier de recatégorisation, dans Argentine.

Dans leur grande majorité, les diaphores sont grammaticales, c'est-à-dire, comportent une marque linguistique explicitant le lien avec l'antécédent. Ces marqueurs sont des adjectifs et des pronoms démonstratifs (**celui**, **esos**), des articles définis (**les**, **la**), des adjectifs possessifs (**leur**, **su**), des adjectifs indéfinis (**tels**, **cada**), un adverbe (**así**), et des éléments typographiques (virgule, deux points). Seules trois reprises anaphoriques se passent de marques : éxito de los coreanos reprend la séquence obtienen de ellos [embrión humano] células madre dans un texte d'Espagne, et meta reprend obtuvieron de ellos [embrión humano] las células primigenias que pueden transformarse en cualquier tejido del organismo dans un texte d'Argentine (les deux renvoient à la notion /obtenir une cellule

souche d'un embryon humain créé par clonage/), tandis que **embryon** est employé pour **embryon humain** (qui fait référence à la notion /embryon humain/) dans un texte de France. Parmi les expressions relevées, nous comptons un seul exemple de cataphore (dans France) où la séquence **objet de leur [chercheur] quête** annonce, par l'intermédiaire de deux points, le syntagme **cellule souche pluripotente**. L'antécédent peut être une unité ou un syntagme de longueur variable, de deux unités à un passage en entier.

5.3.2.Reformulation

Ce procédé n'est pas très fréquent dans nos corpus; il représente 0,93 % dans Argentine et 0,76 % dans France. Nous trouvons trois reformulations dans le premier. Comme synonymes de célula totipotencial, on indique entre parenthèses célula madre (entre guillemets) et le terme anglais *stem cell* (en italique). Deux syntagmes unis par la conjonction o sont proposés pour renvoyer à la notion cellule souche, célula stem et célula indiferenciada. Enfin, enseñarle a ese núcleo celular que ahora tenía que comportarse como la primera de todas las células del organismo est reformulé par atrasar su reloj biológico à l'aide du marqueur digamos. Dans France, il est utilisé à deux reprises pour introduire des synonymes de **couronne**, **cumulus** et **cumulus cellulaire**, toujours au moyen de la conjonction **ou**.

5.3.3.Définition

Nous avons relevé 3 définitions dans Espagne (0,70 % des actualisations), 5 dans Argentine (1,16 %), 2 dans France (0,76 %) et 2 dans Québec (1,18 %). Les unités ou syntagmes définis expriment, pour la plupart, des entités concrètes : macizo interno, ovocito et citoplasma, dans Argentine et cumulus dans Espagne, ainsi que **blastocyste** et **foetus**, dans Québec. Nous trouvons également des définitions de clonación terapéutica (Argentine) et de **clonage thérapeutique** (France), qui renvoient à l'action /créer un embryon humain par clonage dans un but thérapeutique/. De plus, sont définis un syntagme exprimant une propriété (sus [embrión] células son totipotenciales, qui fait référence à la notion /pluripotence/), dans Argentine, et une unité exprimant un processus (partenogénesis, qui renvoi à la notion /parthénogenèse/), dans Espagne.

Les définitions prennent souvent une forme nominale - forme canonique de la définition composée d'un générique et de traits spécifiques : une unité nominale suivie d'une unité ou d'un syntagme adjectival (fenómeno natural, qui définit partenogénesis), d'une subordonnée relative (**stratégie médicale qui consisterait à traiter un malade en lui**

greffant des cellules souches dérivées de son clone embryonnaire est utilisé pour définir clonage thérapeutique) ou d'un syntagme prépositionnel (**embryon aux premiers stades de son développement qui ne contient que 100 cellules** est la définition de **blastocyste**). Elles peuvent aussi se présenter sous la forme de syntagmes verbaux (desarrollar, por clonación, "células madre" définit clonación terapéutica).

Les marques linguistiques de la définition sont les verbes métalinguistiques llamar et **appeler**, la coordination métalinguistique es decir et des marques typographiques (des tirets, des parenthèses et des virgules). À trois reprises les guillemets sont utilisés pour mettre une expression en évidence : **clonage thérapeutique**, citoplasma et totipotenciales.

5.3.4. Explication

La place occupée par les contextes explicatifs dans nos corpus n'est pas très significative, quoique plus importante que celle de la définition. Dans Espagne, nous avons repéré 23 explications (5,40 % des actualisations), dans Argentine, nous en avons relevé 17 (3,94 %), dans France, 14 (5,32 %) et dans Québec, 10 (5,88 %). Le tableau 9 regroupe les notions faisant l'objet d'une explication.

Tableau 9. Notions faisant l'objet d'une explication dans les quatre sous-corpus

Espagne	Argentine	France	Québec
/cellule souche/	/cellule souche/	/cellule souche/	/cellule souche/
/créer un embryon humain par clonage dans un but thérapeutique/	/créer un embryon humain par clonage dans un but thérapeutique/	/créer un embryon humain par clonage dans un but thérapeutique/	/créer un embryon humain par clonage dans un but thérapeutique/
/cellule souche pluripotente/	/cellule souche pluripotente/	/cellule souche pluripotente/	
/blastocyste/	/blastocyste/		/blastocyste/
/ovocyte/		/ovocyte/	/ovocyte/
/remplacer le noyau d'un ovocyte par le noyau d'une cellule/			/remplacer le noyau d'un ovocyte par le noyau d'une cellule/
/noyau/	/noyau/		
/obtenir une cellule souche d'un embryon humain créé par clonage/	/obtenir une cellule souche d'un embryon humain créé par clonage/		
/créer un embryon humain par clonage/			
/indifférenciation/			
/pluripotence/			
	/animal génétiquement identique à un autre/		
	/cellule de la masse interne/		

/cellule du cumulus/	
/feuillet/	
/ovule/	
/reprogrammation du matériel génétique/	
	/créer un embryon par clonage/
	/cumulus/
	/embryon humain mal formé/
	/guérir un patient/
	/stade de blastocyste/
	/traiter un patient/
	/cellule souche d'un embryon/
	/embryon humain surnuméraire/

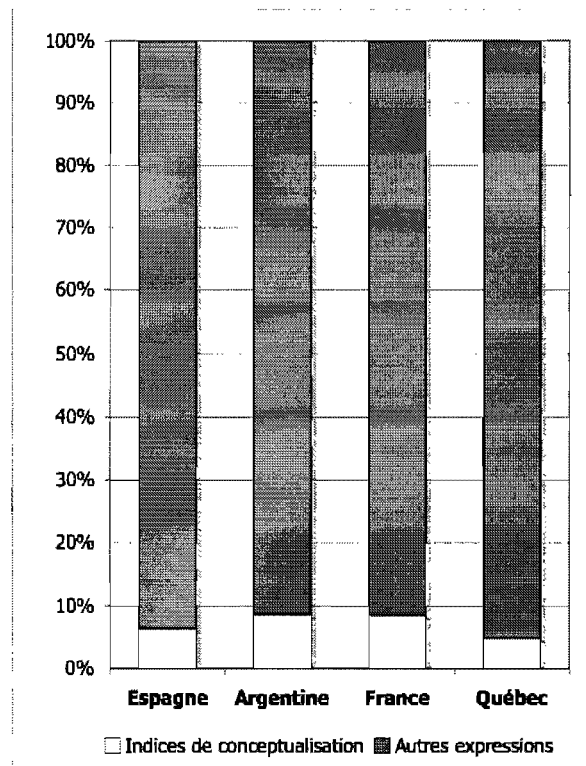
Quant à la structure des explications, nous trouvons des subordonnées relatives (**qui [couronne] entoure naturellement l'ovocyte**), des subordonnées circonstancielles (cuando los embriones son sólo una esfera hueca formada por unas 100 células), des propositions indépendantes (**les cellules souches pourraient devenir la source de thérapies médicales importantes, par exemple des organes de remplacement**) ainsi que des syntagmes nominaux (**matériel humain rare et précieux**), verbaux (silenciar todos los genes correspondientes a célula somática y activar los genes embrionarios) et adjectivaux ([óvulo] necesarios para el proceso de clonación).

L'explication n'est pas toujours marquée (notamment dans le cas des propositions indépendantes), mais ses marques linguistiques sont variées. On emploie la virgule, les parenthèses, les tirets, les guillemets (pour indiquer l'unité ou le syntagme expliqués) et des éléments de coordination métalinguistique tels que es decir et **c'est-à-dire**. Des verbes métalinguistiques sont mis à profit, même s'ils diffèrent de ceux utilisés pour la définition (**parler de, ser, consistir en**) ainsi que la conjonction de subordination porque. Deux structures analogiques servent également à introduire une explication : un blastocito contiene una estructura que recuerda a una mora porque está formada por una masa de células rodeadas de una membrana et **ils [ovule] sont très collants, comme du riz asiatique**.

5.4. Modes de conceptualisation

Parmi les expressions employées pour rendre les notions spécialisées, nous avons identifié des indices de conceptualisation métaphorique. Il s'agit, pour la plupart, d'unités prédicatives dont la structure actancielle établit le lien entre le cadre conceptuel source et le cadre conceptuel cible. Le corpus en espagnol en comporte 31, dont 17 figurent dans le sous-corpus Espagne et 23 dans le sous-corpus Argentine, tandis que l'on en trouve 18 dans les textes en français, 15 dans les articles de France et 6 dans ceux du Québec. Dans tous les sous-corpus, la variabilité d'un texte à l'autre est très grande. Aux fins de la comparaison, il est nécessaire de mettre en rapport ces chiffres avec le nombre d'expressions relevées dans chaque sous-corpus. Ainsi, nous constatons que les indices de conceptualisation représentent 6,37 % des expressions d'Espagne, 8,65 % de celles d'Argentine, 8,47 % de celles de France et 4,88 % de celles de Québec, tel que le montre la figure 50.

Figure 50. Proportion des indices de conceptualisation métaphorique au sein des expressions de chaque sous-corpus



Ces indices rendent compte de nombreux modes de conceptualisation métaphorique, dont plusieurs sont communs aux deux corpus. Nous avons distingué les modes de conceptualisation des différentes entités biologiques car ils ne sont pas toujours les mêmes. Ainsi, dans les quatre sous-corpus, nous trouvons des indices de la conceptualisation des cellules et des embryons comme des êtres doués d'intentionnalité

(un seul embryon *capable de produire des cellules souches embryonnaires*), et des indices de la conceptualisation des cellules comme des végétaux (*estas células se cultivan*).

D'autres modes de conceptualisation sont partagés par les deux corpus, bien que seulement par trois sous-corpus : les cellules sont conçues comme des êtres humains (**on les [cellule souche pluripotente] y [produire tous les types de cellules différenciées de l'organisme humain] pousse**), des produits (*producir una línea de células madre*) et des fournitures pour la production (**les cellules souches embryonnaires constituent la matière première d'où proviennent tous les types de tissu de l'organisme**); le matériel génétique est conceptualisé comme un programme informatique (**reprogrammation génétique**); les êtres humains sont vus comme des produits (*fabricar bebés*); et le cadre conceptuel de la cuisine est évoqué pour parler de la recherche scientifique (*receta de cómo hacer niños clónicos*).

Trois autres modes de conceptualisation apparaissent dans les deux corpus, celui des cellules comme des objets (*manejo de células que formarían el feto*), celui des embryons perçus comme des objets vivants « cultivés » comme peuvent l'être des végétaux (**les chercheurs coréens ont obtenu 30 embryons clonés, qui ont pu être cultivés jusqu'au stade blastocyste**) et celui des thérapies comme des objets ayant une forme physique que l'on peut modifier (*diseñar terapias específicas para enfermedades como la diabetes o el Parkinson, tratamiento sur mesure*).

Nous avons repéré exclusivement dans le corpus en espagnol la conceptualisation du génome comme une empreinte (*huella genética*) et celle du matériel génétique comme un objet (*estas células llevan el material genético nuclear del individuo*). Les embryons sont conçus comme des produits seulement dans le corpus en français (**fabrication d'un embryon de clone humain**). Certains modes de conceptualisation ne se manifestent que dans un seul sous-corpus (ils sont présentés en détail dans le tableau 7).

En ce qui concerne les entités biologiques, nous observons deux tendances qui peuvent paraître opposées. D'une part, elles sont conceptualisées comme des êtres doués d'intentionnalité, voire comme des êtres humains. D'autre part, les indices témoignant d'une réification sont aussi très courants (produit, objet). Par ailleurs, les modes de conceptualisation des êtres humains vont généralement dans le sens d'une réification; ils sont conceptualisés comme des machines ou comme des produits.

Seuls deux indices de conceptualisation se présentent sous la forme d'une analogie : **ils [ovule] sont très collants, comme du riz asiatique** et un blastocite contiene una estructura que recuerda a una masa porque está formada por una masa de células rodeadas de una membrana. Ils ne font pas partie d'expressions renvoyant à des notions, mais se trouvent à l'intérieur de passages explicatifs. Ce type de structure se distingue de

l'expression métaphorique par l'explicitation du mécanisme de conceptualisation à l'aide de marques linguistiques (**comme** et que recuerda a). Dans les deux cas, le trait similaire justifiant l'analogie est mentionné (les ovules ressemblent à du riz asiatique parce qu'ils sont **collants** tandis que le blastocyste et la mûre se ressemblent dans leur structure [estructura]).

Le tableau 10 présente les divers modes et indices de conceptualisation. Certains indices y figurent plus d'une fois car ils témoignent de plusieurs modes de conceptualisation. Les actants réalisés faisant l'objet de la conceptualisation et les actants typiques impliqués sont indiqués en italique.

Tableau 10. Modes de conceptualisation métaphorique et indices de conceptualisation repérés dans les corpus

LA CELLULE EST...							
... UN ÊTRE DOUÉ D'INTENTIONNALITÉ							
Indice	Actants réalisés			Classes typiques d'actants			Sous-corpus
	1 ^{er} actant	2 ^e actant	3 ^e actant	1 ^{er} actant	2 ^e actant	3 ^e actant	
<u>capaz</u>	<ul style="list-style-type: none"> • <u>célula madre embrionaria de origen clónico</u> • <u>célula</u> • <u>línea [de células madre]</u> 	<ul style="list-style-type: none"> • <u>diferenciarse en distintos tejidos del cuerpo humano</u> • <u>formar músculos, huesos, tejidos y neuronas</u> • <u>reemplazar hueso, músculo, células cerebrales dañadas, o cualquier célula del cuerpo</u> 		<u>être doué d'intentionnalité</u>	action		Espagne / Argentine
<u>capacidad</u>	<ul style="list-style-type: none"> • <u>célula madre</u> • <u>célula troncal</u> 	<ul style="list-style-type: none"> • <u>originar las [célula] propias de los tejidos adultos</u> • <u>formar todos los tipos de células que construye el organismo</u> 		<u>être doué d'intentionnalité</u>	action		Espagne / Argentine
<u>guiar</u>	<ul style="list-style-type: none"> • <u>científico</u> • <u>investigador</u> 	<ul style="list-style-type: none"> • <u>célula del embrión</u> • <u>célula</u> 	<ul style="list-style-type: none"> • <u>convertirse en tejidos que podrían aplicarse para distintas terapias</u> • <u>diferenciación en tejidos</u> 	<u>être humain</u>	<u>être doué d'intentionnalité</u>	lieu	Argentine
<u>llevar</u>	<ul style="list-style-type: none"> • <u>célula madre</u> • <u>célula</u> 	<ul style="list-style-type: none"> • <u>material genético</u> • <u>material genético nuclear</u> 		<u>être doué d'intentionnalité</u>	objet		Espagne / Argentine
<u>poder</u>	<ul style="list-style-type: none"> • <u>célula primigenia</u> 	<ul style="list-style-type: none"> • <u>transformarse en cualquier tejido del organismo</u> 		<u>être doué d'intentionnalité</u>	action		Argentine
<u>reproducirse</u>	<ul style="list-style-type: none"> • <u>célula troncal</u> 			<u>être doué d'intentionnalité</u>			Argentine

capable	• <i>cellule-souche dite pluripotente</i>	• produire les différents tissus de l'organisme		être doué d'intentionnalité	action		France
survivre	• <i>culture</i>			être doué d'intentionnalité			Québec

... UN ÊTRE HUMAIN

Indice	Actants réalisés			Classes typiques d'actants			Sous-corpus
	1 ^{er} actant	2 ^e actant	3 ^e actant	1 ^{er} actant	2 ^e actant	3 ^e actant	
<u>cumplir</u>	• <i>célula</i>	• <i>función</i>		<i>être humain</i>	fonction		Argentine
<u>disparar*</u>	• <i>célula pluripotencial</i>	• <i>rechazo fisiológico natural</i>		<i>être humain</i>	coup de feu		Espagne
* Le verbe <i>disparar</i> est employée couramment de façon métaphorique en espagnol dans le sens de « déclencher un processus »							
pousser	• on [chercheur]	• <i>cellule souche pluripotente</i>	• produire tous les types de cellules différenciées de l'organisme humain	être humain	<i>être humain</i>	action	France
produire	• <i>cellule souche pluripotente</i> • <i>cellule souche dite pluripotente</i>	• type de cellules différenciées • tissu		<i>être humain</i>	produit		France

... UN VÉGÉTAL

Indice	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>cultivo</u>	• <i>célula</i> • <i>célula madre obtenida de un embrión clónico</i>		végétal		Espagne / Argentine
<u>cultivar</u>	• <i>científico</i> • <i>investigador</i>	• <i>célula del macizo interno</i> • <i>célula</i> • <i>blastocito</i>	être humain	<i>végétal</i>	Espagne / Argentine
<u>culture</u>	• [cellule]		<i>végétal</i>		Québec
<u>cultivé</u>	• <i>cellule</i> • <i>cellule souche</i>		<i>végétal</i>		France
<u>cultiver</u>	• chercheur sud-coréen	• <i>cellule souche</i> • <i>cellule souche embryonnaire pluripotente</i> • <i>lignée de ces cellules [souches]</i>	être humain	<i>végétal</i>	France

culture	<ul style="list-style-type: none"> • <i>cellule</i> • <i>cellule souche</i> • <i>cellule souche embryonnaire pluripotente</i> 		<i>végétal</i>		France
----------------	--	--	----------------	--	--------

... UN OBJET

Indice	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>construir</u>	• <i>organismo</i>	• <i>tipo de células</i>	être vivant	<i>objet</i>	Argentine
<u>manejo</u>	• <i>célula que formaría el feto</i>		<i>objet</i>		Argentine
fournir*	• <i>embryon</i>	• <i>lignée de cellules pluripotentes</i>	être humain	<i>objet</i>	France

* Le verbe *fournir* est employée couramment de façon métaphorique en français dans le sens de « être à l'origine »

... UN PRODUIT

Indice	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>producción</u>	• <i>célula madre</i>		<i>produit</i>		Espagne
<u>producir(se)</u>	• <i>persona</i>	<ul style="list-style-type: none"> • <i>célula</i> • <i>célula madre</i> • <i>línea de células madre</i> • <i>línea estable de células madre</i> 	être humain	<i>produit</i>	Espagne / Argentine
produire	<ul style="list-style-type: none"> • <i>cellule souche pluripotente</i> • <i>embryon</i> • <i>ils [chercheur]</i> 	<ul style="list-style-type: none"> • <i>type de cellules différenciées</i> • <i>cellule souche embryonnaire</i> • <i>cellule souche</i> 	être humain	<i>produit</i>	France
production	• <i>cellule souche embryonnaire</i>		<i>produit</i>		France

... UNE FOURNITURE

Indice	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>materia prima</u>	Ici, la conceptualisation métaphorique n'opère par la projection des actants du prédicat, mais par une ressemblance de traits entre l'entité évoquée et l'entité conceptualisée.				Argentine
matière première					Québec / France

LES COMPOSANTES DE LA CELLULE SONT...

... LES PARTIES D'UNE MACHINE

Indice	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>maquinaria</u>	• <u>división</u>		fonction d'une machine		Espagne

L'EMBRYON EST...

... UN ÊTRE DOUÉ D'INTENTIONNALITÉ

Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>ir</u>	• <u>embrión</u>	• <u>ocho o diez células</u>	<i>être doué d'intentionnalité</i>	lieu	Argentine
<u>llegar</u>	• <u>embrión</u>	• <u>punto clave de desarrollo en que se diferencia en tres capas</u>	<i>être doué d'intentionnalité</i>	lieu	Argentine
<u>sufrir</u>	• <u>embrión</u>	• <u>aberración cromosómica</u>	<i>être doué d'intentionnalité</i>	douleur; malheur	Espagne
<u>matar</u>		• <u>[embrión]</u>		<i>être doué d'intentionnalité</i>	Argentine
<u>survivre</u>	• <u>embryon</u>		<i>être doué d'intentionnalité</i>		Québec
<u>capable</u>	• <u>embryon</u>	• <u>produire des cellules souches embryonnaires</u>	<i>être doué d'intentionnalité</i>	action	France
<u>fournir*</u>	• <u>embryon</u>	• <u>lignée de cellules pluripotentes</u>	<i>être doué d'intentionnalité</i>	objet	France

* Le verbe *fournir* est employée couramment de façon métaphorique en français dans le sens de « être à l'origine »

... UN ÊTRE HUMAIN

Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>produire</u>	• <u>embryon</u>	• <u>cellule souche embryonnaire</u>	<i>être humain</i>	produit	France

... UN VÉGÉTAL

Indice	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>cultivar</u>	• <u>científico</u> • <u>investigador</u>	• <u>blastocito</u>	être humain	<i>végétal</i>	Espagne / Argentine

... UN PRODUIT					
Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
produire	• chercheur sud-coréen	• embryon humain	être humain	<i>produit</i>	Québec / France
fabrication	• embryon de clone humain		<i>produit</i>		France

LE CYTOPLASME EST...

... UN ÊTRE DOUÉ D'INTENTIONNALITÉ

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>silenciar</u>	• <u>citoplasma</u>	• <u>gen</u>	<i>être doué d'intentionnalité</i>	• être humain	Argentine
<u>reprogramar</u>	• <u>citoplasma</u>	• <u>núcleo</u>	<i>être doué d'intentionnalité</i>	programme informatique	Argentine

LE MATÉRIEL GÉNÉTIQUE EST...

... UN PROGRAMME INFORMATIQUE

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>reprogramación</u>	• <u>reloj biológico</u> • <u>material genético</u>		<i>programme informatique</i>		Espagne / Argentine
<u>reprogramar</u>	• <u>citoplasma</u>	• <u>núcleo</u>		<i>programme informatique</i>	Argentine
<u>reprogrammation</u>			<i>programme informatique</i>		France

... UN OBJET

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>llevar</u>	• <u>célula madre</u>	• <u>material genético</u> • <u>material genético nuclear</u>	être vivant	<i>objet</i>	Espagne / Argentine

L'ÊTRE HUMAIN EST...

... UNE MACHINE

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous- corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
réparer		• être humain		machine	France

... UN PRODUIT

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous- corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
fabricar		• bebé		produit	Espagne
producción	• ser humano		produit		Argentine
produire	• processus	• clone des personnes ayant donné le noyau d'une de leurs cellules	être humain	produit	Québec

LE CHERCHEUR EST...

... UN ATHLÈTE

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous- corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
s'entraîner	• ils [chercheur]		athlète		France

LA THÉRAPIE EST...

... UN VÊTEMENT

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous- corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
diseñar		• terapia específica para enfermedades como la diabetes o el Parkinson		vêtement	Espagne
sur mesure	• traitement		vêtement		Québec

* Le syntagme *sur mesure* est employée couramment de façon métaphorique en français dans le sens de « adapté »

FAIRE DE LA RECHERCHE SCIENTIFIQUE EST...

... FAIRE DE LA CUISINE

Indice (unité prédicative)	Actants réalisés		Classes typiques d'actants		Sous- corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>receta</u>	Ici, la conceptualisation métaphorique n'opère par la projection des actants du prédicat, mais de façon directe.				Espagne / Argentine
<u>recette</u>					France

LE GÈNE EST...

... UN ÊTRE DOUÉ D'INTENTIONNALITÉ

Indice	Actants réalisés		Catégories typiques d'actants		Sous- corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>huella</u>	• <u>[gen]</u>		être doué d'intentionnalité		Espagne
<u>silenciar</u>	• <u>citoplasma</u>	• <u>gen</u>	être humain	être doué d'intentionnalité	Argentine

... UN OBJET

Indice	Actants réalisés		Catégories typiques d'actants		Sous- corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>herencia</u>	• <u>[padre]</u>	• <u>[gen]</u>	être vivant	objet	Espagne

LE NOYAU EST...

... UN ÊTRE VIVANT

Indice	Actants réalisés			Classes typiques d'actants			Sous- corpus
	1 ^{er} actant	2 ^e actant	3 ^e actant	1 ^{er} actant	2 ^e actant	3 ^e actant	
<u>enseñarle</u>		• <u>núcleo celular</u>	• <u>comportarse como la primera de todas las células del organismo</u>		être vivant	action	Argentine

L'ORGANISME EST...					
... UN ÊTRE DOUÉ D'INTENTIONNALITÉ					
Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>construir</u>	• <i>organismo</i>	• <u>tipo de células</u>	<i>être doué d'intentionnalité</i>	objet	Argentine
... UN PRODUIT					
Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>producir</u>		<i>organismo</i>	être humain	<i>objet</i>	Argentine
L'OVULE EST...					
... UN OBJET					
Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>emplear</u>	• <u>investigación</u>	• <i>óvulo</i>	être vivant	• <i>objet</i>	Espagne
... UN PRODUIT					
Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>producción</u>	• <i>ovulo</i>		<i>produit</i>		Espagne
... UNE FOURNITURE					
Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>materia prima</u>	Ici, la conceptualisation métaphorique n'opère par la projection des actants du prédicat, mais par une ressemblance de traits entre l'entité évoquée et l'entité conceptualisée.				Argentine
LE TISSU EST...					
... UN PRODUIT					
Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>produire</u>	• <u>cellule souche dite pluripotente</u>	• <i>tissu</i>	être humain	<i>produit</i>	France

LE SYSTÈME NERVEUX EST...

... UNE CRÉATION

Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
ébauche	• <i>système nerveux</i>		<i>création</i>		Québec

L'ORGANE EST...

... UNE MACHINE

Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
reparar	• <i>órgano</i>		<i>machine</i>		Argentine

LA DIFFÉRENCIATION CELLULAIRE EST...

... UN LIEU

Indice	Actants réalisés			Classes typiques d'actants			Sous-corpus
	1 ^{er} actant	2 ^e actant	3 ^e actant	1 ^{er} actant	2 ^e actant	3 ^e actant	
guiar	• <i>científico</i> • <i>investigador</i>	• <i>célula del embrión</i> • <i>célula</i>	• <i>convertirse en tejidos que podrían aplicarse para distintas terapias</i> • <i>diferenciación en tejidos</i>	être vivant	être vivant	<i>lieu</i>	Argentine

LE REJET IMMUNOLOGIQUE EST...

... UN COUP DE FEU

Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
disparar*	• <i>célula pluripotencial</i>	• <i>rechazo fisiológico natural</i>	être humain	<i>coup de feu</i>	Argentine

* Le verbe *disparar* est employée couramment de façon métaphorique en espagnol dans le sens de « déclencher un processus »

LE STADE DE BLASTOCYSTE EST...

... UN LIEU

Indice	Actants réalisés		Catégories typiques d'actants		Sous-corpus
	1 ^{er} actant	2 ^e actant	1 ^{er} actant	2 ^e actant	
<u>ir</u>	• <u>embrión</u>	• <i>ocho o diez células</i>	être vivant	<i>lieu</i>	Argentine
<u>llegar</u>	• <u>embrión</u>	• <i>punto clave de desarrollo en que se diferencia en tres capas</i>	être vivant	<i>lieu</i>	Argentine

6. Discussion

Au cours de cette recherche, une quantité considérable de données ont été recueillies. Il est nécessaire maintenant de les examiner en regard de notre problématique et des travaux d'autres chercheurs. Dans ce chapitre, nous tenterons de déterminer dans quelle mesure les résultats obtenus permettent de fournir une réponse aux questions et si les hypothèses de départ rendent compte des phénomènes observés. En outre, nous discuterons de certains points soulevés par la réalisation de ce travail. Cependant, avant de procéder à l'analyse des résultats, il est nécessaire de faire un retour sur la démarche méthodologique adoptée.

6.1. Démarche méthodologique

Notre problématique naît d'une vision particulière du texte, de la traduction et de la communication spécialisée. Ainsi, pour tenter de répondre aux questions soulevées, nous avons dû articuler des principes méthodologiques appartenant à divers horizons théoriques (principalement, l'analyse du discours, la sémantique et les sciences cognitives). Cependant, les démarches méthodologiques proposées dans la littérature ne pouvaient être appliquées telles quelles pour notre travail; d'une part, parce que les travaux des autres chercheurs poursuivent des buts différents des nôtres et, d'autre part, en raison de certains choix, à notre avis, discutables (par exemple, le fait que Ciapuscio [2003a] ait traduit elle-même des textes pour les besoins de son étude)⁵⁸.

Il nous a donc fallu concevoir une méthodologie adaptée aux objectifs spécifiques de cette recherche, ce qui comporte des avantages et des inconvénients. Parmi ces derniers, nous pouvons mentionner le temps que demande l'élaboration d'une démarche propre et les désavantages de ne pas compter sur une méthode éprouvée. En revanche, les avantages sont considérables, car nous avons pu éviter d'adopter des principes d'autres travaux qui semblaient inadéquats pour le nôtre (nous songeons notamment aux critères formels souvent employés pour le repérage des séquences linguistiques analysées).

⁵⁸ Nous l'avons signalé lors de la présentation de ces travaux à l'« État de la question ».

6.1.1. Le corpus

La compilation des corpus a fait appel à quatre critères : le lieu de publication (la France, le Québec, L'Espagne et l'Argentine), le moment discursif (l'annonce des résultats de la recherche d'une équipe coréenne), le type de journal (journaux nationaux et provinciaux) et le genre textuel (exclusion des textes d'opinion). Ce jeu de critères – notamment l'interaction des trois derniers – a limité la taille des corpus et causé un déséquilibre au détriment du sous-corpus Québec, qui comporte un texte de moins que les autres. De fait, un test t de Student montre une différence significative entre le nombre des mots du corpus en français et celui du corpus en espagnol ($p=0,04$). Lors d'une recherche future, il sera nécessaire de prévoir le recours à des statistiques permettant de surmonter cet obstacle ou de réviser les paramètres afin de monter des corpus d'analyse plus vastes et plus équilibrés.

Malgré ces faiblesses, le choix des textes en fonction d'un moment discursif commun et d'un même genre discursif comporte le grand avantage méthodologique que représente le contrôle d'une variable. Ainsi, nos corpus présentent une certaine homogénéité thématique (autant au niveau des notions que des hyper-catégories) autorisant une analyse comparative sur plusieurs points. Nous voyons là une amélioration par rapport au travail de Freixa (2002), qui rapporte une grande disparité notionnelle dans son corpus. De plus, l'appartenance de tous les textes au genre des « articles journalistiques informatifs » rend possible, en fonction des résultats obtenus, de tirer des observations préliminaires et d'avancer des hypothèses quant à la relation entre certaines caractéristiques stylistiques de la vulgarisation et le genre discursif. Il faut également mentionner que, si l'observation de textes traduits réalisée par Carreño (2004) et par Suárez (2004) fournit des données quant aux conséquences des choix des traducteurs, la réflexion sur des corpus comparables s'avère plus pertinente pour la caractérisation du discours spécialisé.

6.1.2. La collecte des données

Notre objectif étant d'examiner toutes les expressions renvoyant aux notions spécialisées, et pas seulement les unités terminologiques, nous avons adopté pour le repérage un principe d'ordre cognitif. Ce choix nous a évité certains écueils, mais nous en a fait rencontrer d'autres. La première difficulté se trouvait dans la délimitation des expressions : nous avons été conduits parfois à relever de longues séquences et des balises concrètes étaient essentielles pour minimiser le « bruit ». Puis, afin de déterminer quelle était la notion exprimée par l'expression relevée, il était nécessaire de faire la

distinction entre la signification linguistique et le sens référentiel en discours⁵⁹, ce qui demandait un examen systématique du contexte. Cerner la notion à laquelle on faisait référence n'a pas toujours été chose facile et les décisions prises dans certains cas peuvent donner lieu à discussion. C'est le cas de certaines expressions imbriquées. Par exemple, dans la séquence **des biologistes coréens ont créé des embryons humains par clonage**, qui renvoie à la notion /créer un embryon humain par clonage/, **clonage** n'a pas été repéré de façon autonome, car nous avons considéré que cette unité renvoie à la même notion que l'expression de laquelle elle fait partie. Pourtant, on pourrait argumenter qu'il s'agit de deux notions distinctes.

Enfin, l'identification des indices de conceptualisation métaphorique par dissonance cognitive⁶⁰ comporte une part d'introspection qui peut être discutable à certains égards. Ce processus cognitif est pourtant inévitable, étant donné l'impossibilité de recueillir ces données à l'aide de critères formels ou d'une analyse exclusivement linguistique. La confrontation des résultats obtenus avec ceux d'autres chercheurs devient dès lors essentielle.

Nous constatons donc que l'approche cognitive exige un travail de réflexion assez ardu et la mise en place de plusieurs critères secondaires afin d'assurer la cohérence, la reproductibilité et la pertinence du repérage. Là réside cependant à nos yeux l'originalité de notre démarche au regard d'autres études. Les travaux menés sous l'angle terminologique mettent l'accent sur la structure morphosyntaxique (Carreño 2004, Freixa 2002, Suárez 2004), tandis que les méthodes de collecte des données de l'analyse du discours restent toujours assez floues (Cassany et Martí 1998, Cataldi 2004). Seules Ciapuscio (2003a) et Vandaele (plusieurs travaux) se servent de modèles cognitifs pour étudier les réalisations linguistiques, bien que leurs objectifs soient différents des nôtres. Du point de vue traductologique, le recours à un critère cognitif était fondamental afin de pouvoir comparer des corpus dans deux langues différentes. Ce choix a aussi été crucial pour l'analyse dans une perspective discursive. Nous avons pu ainsi regrouper les expressions dans des ensembles paraphrastiques, analyser les liens de coréférence s'établissant entre elles et mettre à jour les diverses stratégies déployées dans la mise en discours des notions.

En outre, le repérage des expressions a nécessité l'élaboration d'un système complexe d'annotation des corpus. En dépit d'avoir fait appel à une ressource d'étiquetage déjà développée pour d'autres projets (Vandaele et Boudreau 2006), le jeu de balises et

⁵⁹ La distinction entre notion et signifié est l'un des principes théoriques sous-tendant notre démarche (voir la section 2.3. du « Cadre théorique »).

⁶⁰ Cette notion a été présentée à la section 2.6. du « Cadre théorique ».

d'attributs a dû être créé spécifiquement pour ce travail. Cela a eu pour corollaire d'allonger le processus même d'annotation, qui a comporté plusieurs étapes de correction. Cette méthode offrait pourtant la souplesse nécessaire à notre analyse, car elle permet une annotation fine des textes quels que soient les critères choisis. La commodité et la rapidité dans la récupération des données constitue également un atout du système.

6.1.3. Validité des données obtenues

En regard de l'interprétation des résultats, il faut admettre des limites à la démarche méthodologique. Les critères utilisés pour le repérage de toutes les expressions renvoyant à des notions spécialisées (repérage essentiel pour cette étude exploratoire) ne pouvaient en même temps viser une analyse approfondie des phénomènes observés. Par ailleurs, la faible taille du corpus et la dispersion des résultats réduisent l'applicabilité et le pouvoir explicatif des tests statistiques. Ainsi, une prochaine étape de la recherche se devrait de raffiner les critères méthodologiques afin d'isoler des phénomènes spécifiques et compter sur un corpus, non seulement plus vaste, mais autorisant un traitement statistique solide.

Le mérite de la démarche est d'avoir servi à un premier examen - inédit - des variables qui nous intéressaient et à tirer un nombre considérable d'observations préliminaires ouvrant des pistes de recherche. Nous considérons donc qu'elle est pertinente, compte tenu des objectifs de ce travail. Bien évidemment, les critères qui la constituent doivent être revus à la lumière des nouvelles hypothèses, mais nous avons pu ainsi jeter des bases méthodologiques pour la suite de notre programme de recherche.

6.2. Expression linguistique des notions

La première question que nous nous posions concernait la façon dont les notions spécialisées sont exprimées linguistiquement dans les articles journalistiques de vulgarisation. D'une part, nous nous attendions à l'emploi d'un nombre assez élevée d'unités et syntagmes pour chaque notion - autrement dit, à de riches ensembles paraphrastiques -, ainsi qu'à trouver des structures morphosyntaxiques très variées. D'autre part, nous prévoyions que les textes accorderaient une part importante à certaines stratégies textuelles : la diaphore, la reformulation, la définition et l'explication. Par ailleurs, nous nous demandions quel était l'impact de la langue dans laquelle le texte est rédigé et du contexte socioculturel (le lieu de publication) sur la mise en discours des notions spécialisées. Nous envisagions une possible influence du contexte socioculturel sur le choix des unités et des syntagmes employés et sur le recours aux stratégies textuelles analysées.

6.2.1. *Variation linguistique*

6.2.1.1. *Nombre d'expressions renvoyant à une même notion*

Pour l'analyse du nombre d'expressions faisant partie d'un même ensemble paraphrastique, nous nous penchons exclusivement sur les notions qui ont été actualisées au moins deux fois dans les corpus (section 5.2.2.1.). En effet, tenir compte des notions actualisées une seule fois introduirait un biais. Nous observons dans les quatre sous-corpus une grande dispersion dans le nombre d'expressions renvoyant à une même notion, même si, le plus souvent, il se situe entre 1 et 3. Si certaines notions sont exprimées par un grand nombre d'unités ou syntagmes (jusqu'à 11), d'autres suscitent peu de variation. Les résultats de Freixa (2002)⁶¹, qui a ciblé seulement les dénominations, sont similaires aux nôtres : une majorité de notions sont exprimées par 1 ou 2 dénominations, mais jusqu'à 14 sont employées pour rendre une même notion. Il faut noter qu'elle aussi a décidé de ne pas prendre en compte les notions figurant une seule fois dans ses corpus. Nous retrouvons donc une distribution et une dispersion semblables à celle de nos corpus.

Freixa utilise toutefois la moyenne du nombre des dénominations par notion (« *indice de dénomination* ») comme indicateur du taux de variation de ses corpus, et conclut à une variation plus élevée des discours de vulgarisation. Contrairement à elle, nous ne croyons pas que la moyenne puisse représenter la distribution observée, car elle n'est pas normale (gaussienne). La dispersion des données devrait plutôt nous amener à nuancer l'hypothèse quant à la variation dans la vulgarisation, et à chercher les variables susceptibles d'avoir une incidence sur le nombre d'expressions employées pour renvoyer à une même notion.

À l'instar de Freixa, nous ne trouvons pas une corrélation entre ce nombre et le nombre d'actualisations de la notion. Nous pouvons, par contre, remarquer une certaine régularité dans la dispersion : les actions et les processus semblent être à l'origine d'un plus grand nombre d'expressions que les entités concrètes et abstraites (section 5.2.2.1.2., figures 33 à 40). Quant aux propriétés, états et phénomènes, nous n'avons pas assez de données pour dégager une tendance claire. Ces résultats nous font penser que les hyper-catégories constituent un paramètre essentiel pour rendre compte de la variation. De cette façon pourrait s'expliquer la différence dans la variation de l'expression entre la notion /créer un embryon par clonage/, dénotant une action, et la notion /ovocyte/, qui dénote une entité concrète. Rappelons que la première a été actualisée 27 fois dans Espagne au moyen

⁶¹ Nous avons présenté son travail de recherche à la section 3.1. de l'« État de la question ».

de 15 expressions, 15 fois dans Argentine par 10 expressions, 22 fois dans France par 15 expressions et 17 fois dans Québec par 10 expressions. Pour la seconde 1 expression a été employée au cours de 15 actualisations dans Espagne, 4 expressions pour 13 actualisations dans Argentine, 2 expressions pour 17 actualisations dans France et 1 expression pour 8 actualisations dans Québec. On pourrait donc envisager une étude plus approfondie du phénomène de la variation en rapport avec les hyper-catégories.

6.2.1.2. Ensembles paraphrastiques

6.2.1.2.1. Expressions communes

Au sein du corpus de chaque langue, quelques expressions sont communes aux sous-corpus respectifs (section 5.2.1., tableaux 5 et 6). Il est intéressant de noter que la plupart de ces expressions communes sont des unités (nominales), ce qui est à l'opposé de la tendance de l'ensemble des expressions repérées, où les syntagmes (nominaux et verbaux) sont les plus nombreux (figure 41). Nos résultats rejoignent ceux de Freixa (2002), qui remarque une fréquence élevée d'unités nominales parmi les dénominations coïncidant dans ses corpus. Par ailleurs, la plupart des constructions verbales (**se développer**, clonar) et des reprises anaphoriques discursives (técnica), ainsi que les hyperonymes (**cellule**, embrión) sont utilisés dans les deux sous-corpus respectifs de chaque langue lorsqu'il existe déjà une autre expression nominale commune (**développement**, clonación humana, clonación, **cellule souche**, embrión humano).

Ces observations nous incitent à penser que les expressions communes à chaque corpus sont surtout des unités et syntagmes possédant un rapport assez stable et récurrent avec la notion exprimée, c'est-à-dire, des dénominations⁶². Nous nous demandons surtout s'il s'agit des dénominations spécifiques au discours de vulgarisation ou si elles appartiennent au domaine du clonage et des cellules souches. Afin de répondre à cette question, il serait nécessaire de réaliser une étude comparative qui vérifierait la présence de ces expressions dans des textes scientifiques. Ciapuscio (2003a) et Freixa (2002) ont réalisé cette comparaison, mais leurs résultats sont contradictoires. Ciapuscio ne trouve aucune coïncidence entre les textes de vulgarisation et scientifique observés, mais elle admet l'existence d'une terminologie se trouvant dans une « *zona intermedia entre el llamado léxico especial y el general* »⁶³ (Ciapuscio 2003a, p. 90). En revanche, Freixa signale que 35 % des notions sont exprimées exactement de la même façon dans ses deux corpus.

⁶² Cette notion a été présentée à la section 2.5. du « Cadre théorique ».

⁶³ [Notre traduction :] « *zone intermédiaire entre ce que l'on appelle le lexique spécialisé et le lexique général* ».

Une étude portant sur l'ensemble des notions que nous avons identifiées dépasse les limites du présent travail et fera l'objet de travaux ultérieurs. Nous avons cependant effectué une exploration préliminaire en ayant recours au vaste corpus scientifique compilé par le puissant moteur de recherche *Google Scholar*, dans ses versions française et espagnole (*Google Académico*). En raison des difficultés soulevées par la recherche des structures verbales en raison de la structure actancielle fluctuante, nous avons limité l'interrogation aux constructions nominales. Nous avons également exclu les trois éléments anaphoriques discursifs (*técnica*, *logro* et *resultado*) et l'indice de conceptualisation métaphorique (*receta*)⁶⁴, car il aurait été nécessaire de faire un tri laborieux dans les résultats pour s'assurer de leur pertinence. L'analyse a été menée sur 75,86 % des expressions communes au corpus en espagnol et sur 85 % de celles partagées dans le corpus en français. Les tableaux 11 et 12 affichent les expressions retrouvées et le nombre de résultats obtenus.

Tableau 11. Expressions communes aux sous-corpus Espagne et Argentine retrouvées dans le moteur de recherche *Google Académico*

Expression	Résultats
<u>blastocito</u>	224
<u>célula</u>	84 700
<u>célula somática</u>	658
<u>célula madre</u>	780
<u>célula madre embrionaria</u>	24
<u>clonación animal</u>	83
<u>clonación</u>	9 940
<u>clonación humana</u>	729
<u>clonación reproductiva</u>	233
<u>clonación terapéutica</u>	302
<u>diabetes</u>	33 800
<u>diferenciación*</u>	17 600
<u>división*</u>	19 800
<u>embrión</u>	12 000
<u>embrión humano</u>	921
<u>feto</u>	14 600
<u>gen</u>	49 200
<u>trasplante</u>	9 600
<u>línea de células madre</u>	20
<u>línea celular</u>	1 920
<u>enfermedad</u>	260 000
<u>mal de Alzheimer</u>	177
<u>Parkinson**</u>	5 720
<u>enfermedad degenerativa</u>	865
<u>material genético</u>	6 180
<u>membrana*</u>	19 100
<u>nacimiento*</u>	9 710
<u>núcleo*</u>	19 100
<u>órgano</u>	48 300

⁶⁴ Cet indice de conceptualisation a été exclu, car on sait pertinemment qu'il n'appartient pas au discours scientifique.

<u>osteoartritis</u>	1 640
<u>óvulo</u>	3 080
<u>paciente</u>	207 000
<u>placenta</u>	8 040
<u>rechazo inmunológico</u>	220
<u>reproducción humana</u>	3 090
<u>terapia</u>	74 300
<u>medicina de trasplantes</u>	17
<u>tejido</u>	37 300
<u>tipo de células</u>	1 130
<u>útero</u>	14 000

* Pour que la notion à laquelle renvoie ce nom prédicatif soit précisée, il doit être accompagné de son actant (célula ou bebé). Nous l'avons donc ajouté comme mot clé pour l'interrogation.

** Nous avons ajouté le mot clé « enfermedad » afin d'exclure les résultats non pertinents.

Tableau 12. Expressions communes aux sous-corpus France et Québec retrouvées dans le moteur de recherche *Google Scholar* en français

Expression	Résultats
blastocyste	463
cellule	63 900
cellule souche	1 130
cellule souche embryonnaire	34
clonage	6 540
clonage humain	381
clonage thérapeutique	217
développement*	17 500
diabète	22 500
embryon	15 500
embryon humain	742
cloneur	10
noyau	52 800
organisme	54 600
tissu	46 900

* Pour que la notion à laquelle renvoie ce nom prédicatif soit précisée, il doit être accompagné de son actant (**embryon**). Nous l'avons donc ajouté comme mot clé pour l'interrogation.

Toutes les expressions que nous avons cherchées se retrouvent dans le corpus scientifique exploré. Le nombre élevé de résultats obtenu pour un grand nombre d'expressions indique avec un bon degré de certitude qu'elles sont bel et bien employées dans le discours scientifique du domaine, même si l'on tient compte du fait que toutes les occurrences de ces unités et syntagmes ne renvoient pas nécessairement à la notion considérée. Autrement dit, les vulgarisateurs se servent de la « langue des experts », ce qui étaye notre vision d'un continuum dans les discours visant la diffusion des savoirs spécialisés⁶⁵.

⁶⁵ Cette position a été discutée à la section 2.2. du « Cadre théorique ».

6.2.1.2.2. *Expressions différentes*

Il reste à examiner les expressions qui ne sont pas communes aux sous-corpus respectifs de chaque corpus. Pour cette analyse, nous prendrons comme échantillon l'extrait présenté à la section 5.2.1.1. Nous y observons des paradigmes définitionnels⁶⁶ (reprises diaphoriques, comme l'élément anaphorique **embryon**, et reformulations : célula totipotencial, stem cell) et désignationnels. Les expressions qui font partie de ces derniers présentent des variations lexicales (célula madre et célula troncal), orthographiques (dans ce cas, la variation implique une confusion entre les notions /blastocyste/ et /blastocyte/, mais nous avons relevé **cellule souche** et **cellule-souche** dans le corpus en français pour renvoyer à /cellule souche/) et structurales (**cellule souche pluripotente**, **cellule pluripotente**). Elles peuvent prendre la forme d'unités (**blastocyste**) ou de longs syntagmes pouvant fonctionner comme une définition ou une explication (esfera hueca formada por unas 100 células). Nous trouvons également une expression introduisant la dénomination par l'intermédiaire d'un élément métalinguistique (**cellule-souche dite « pluripotente »**). Certaines paraphrases entretiennent des liens linguistiques entre elles (hyperonymes, hyponymes), tandis que d'autres ne sont reconnues qu'en discours (embrión del mismo estadio que los que se han clonado ahora).

Notre démarche méthodologique a favorisé la collecte d'expressions ayant des rapports de coréférence et de référence variés. Cependant, les résultats des autres auteurs montrent aussi de la diversité. Ciapuscio (2003a) relève, parmi les expressions faisant référence à une même notion (/protéine SPARC/), des dénominations (« *glicoproteína segregada SPARC* »), des hyperonymes (« *protéine* »), des paraphrases définitoires (« *una proteína que favorece la metástasis* ») et métalinguistiques (« *una proteína conocida como SPARC* »). Des cas d'équivalence en contexte (hyperonymes, hyponymes, méronymes et cohyponymes) sont également signalés par Freixa (2002), en plus d'un grand nombre de transformations graphiques, morphosyntaxiques et lexicales. Les résultats des chercheurs qui s'inscrivent dans les courants de l'analyse du discours, comme Cataldi (2004), sont très variés, mais ils sont difficilement comparables en raison du manque de précision sur les éléments linguistiques repérés.

Chaque type de paraphrase remplit des fonctions différentes à l'égard de la recontextualisation de connaissances effectuée par la vulgarisation, même si toutes visent la structuration des savoirs et offrent des points d'ancrage pour la compréhension. Les paraphrases discursives peuvent mettre en relief le lien avec la situation référentielle ou

⁶⁶ Nous rappelons que les paradigmes définitionnels sont ceux qui s'appuient sur des marques linguistiques du rapport entre les expressions tandis que les paradigmes désignationnels n'en comportent pas (sections 2.5. du « Cadre théorique » et 5.2.2.2. des « Résultats »).

proposer un point de vue différent sur la notion. Les définitions et les explications aident à cerner la notion en explicitant certains traits conceptuels. Les reprises diaphoriques assurent la cohérence textuelle et la progression thématique. Les variantes lexicales et orthographiques, et les constructions métalinguistiques introduisent les dénominations du domaine. Le cas de /cellule souche/ s'avère donc intéressant du point de vue de la comparaison des sous-corpus. Le fait que, dans Argentine, 6 variantes lexicales soient employées pour renvoyer à /cellule souche/ peut témoigner d'une spécificité régionale, car nous ne trouvons qu'une dénomination dans Espagne. Il se peut qu'une activité de recherche et un débat de société sur les cellules souches plus développés en Espagne aient contribué à fixer la dénomination. Par contre, en Argentine, cette notion semble encore *nouvelle* et indéterminée (il reste à déterminer si les synonymes proposés le sont vraiment), au point où l'on fait appel à un emprunt pour la nommer (*stem cell*). Cette particularité mérite d'être explorée dans un corpus plus vaste.

N'ayant pas pu mettre en évidence d'autres traits socioculturels propres à certaines unités ou syntagmes ou à certaines notions, nous ne pouvons conclure davantage quant à l'influence du lieu de publication sur le choix des expressions. Comme Freixa, nous voyons la cause de la diversité d'expressions - et de la divergence entre les sous-corpus - surtout dans l'abondance de moyens utilisés par les locuteurs pour bâtir leurs textes (reprises diaphoriques, reformulations, définitions, explications, paraphrases linguistiques et discursives). À notre avis, ces moyens sont employés en fonction des tâches cognitives exigées par la recontextualisation des notions⁶⁷ : la réélaboration des réseaux notionnels, la dénomination et la « *textualisation* ». Notre recherche future devrait ajuster la démarche méthodologique de façon à décrire cette interaction plus précisément et dégager les causes des choix des locuteurs.

6.2.2. Structure morphosyntaxique des expressions

Malgré une grande diversité en ce qui concerne les constructions syntagmatiques (notamment dans les structures actancielles des syntagmes verbaux), les corpus présentent une certaine homogénéité quant aux parties du discours. En effet, la distribution est semblable dans les quatre sous-corpus (section 5.2.3., figure 41) et seulement deux parties du discours sont véritablement représentées : le nom et le verbe; les structures adjectivales pouvant presque toujours être liées à l'un (constructions elliptiques) ou à l'autre (constructions participiales). On observe également dans tous les sous-corpus une présence prédominante des formes nominales. Curieusement, ces

⁶⁷ Voir section 2.5. du « Cadre théorique ».

résultats sont similaires à ceux obtenus par Freixa (2002) à l'aide de critères de repérage assez différents des nôtres (repérage de certaines structures morphosyntaxiques). Par ailleurs, nous pouvons remarquer une tendance à l'uniformité au sein des ensembles paraphrastiques : chaque notion est, le plus souvent, exprimée par des unités et des syntagmes appartenant à la même partie du discours, tout particulièrement le nom (section 5.2.3.5.). Dans l'ensemble des expressions, ce sont les syntagmes nominaux qui prédominent, tandis que les expressions communes aux sous-corpus respectifs de chaque corpus sont plus souvent des unités nominales.

Les régularités observées dans la partie du discours sont cohérentes avec le type de notions véhiculées (hyper-catégories). Rappelons qu'il existe un rapport étroit entre l'hyper-catégorie sous laquelle on peut classer la notion et la partie du discours des unités et syntagmes employés pour l'exprimer (section 5.2.3.6.). Par exemple, les entités concrètes sont exprimées principalement par des structures nominales tandis que l'expression des phénomènes fait toujours appel aux structures verbales. Le fait qu'une majorité des notions actualisées dans nos textes soient des entités concrètes et des actions (rendues surtout par des structures verbales, mais aussi par des structures nominales dérivées) permet d'expliquer l'homogénéité des corpus et la prédominance des noms. Nous pouvons constater une correspondance entre la distribution des expressions en fonction de la structure morphosyntaxique (figure 41) et celle des notions en fonction de l'hyper-catégorie (figure 16).

6.2.3. Stratégies discursives

Nous avons relevé dans les corpus la mise en œuvre des quatre stratégies discursives considérées : la diaphore, la reformulation, la définition et l'explication (section 5.3.). La place de ces mécanismes n'est cependant pas aussi importante que l'affirment certains auteurs (Calsamiglia et Van Dijk 2004, Cassany *et al.* 2000, Ciapuscio 2003a). L'anaphore et l'explication ont été utilisées plus fréquemment que la définition et la reformulation, mais en proportion, peu d'expressions sont liées au moyen de ces procédés. Ainsi, nos résultats contredisent l'idée communément admise qui leur prête un rôle fondamental dans la vulgarisation. Il y a lieu de se demander s'il ne s'agit pas de caractéristiques propres au genre discursif, car la recontextualisation des notions dans le cadre des textes journalistiques obéit à des contraintes spécifiques quant au type d'information et à l'ordre dans lequel elle est présentée⁶⁸.

⁶⁸ Cassany et ses collaborateurs rappellent les six questions auxquelles est censé répondre tout article journalistique (qui, quoi, où, comment, pourquoi et comment) et la loi de la « pyramide inversée » selon laquelle

Les données recueillies permettent cependant de dégager diverses fonctions de ces stratégies et de mettre en perspective leur emploi avec la visée communicative des textes. Les reprises anaphoriques (section 5.3.1.) assurent la progression thématique au moyen du résumé (**cette méthode** reprend l'explication de la façon dont les chercheurs ont obtenu les cellules souches) et de la répétition (estas células madre, célula madre). Elles contribuent également à structurer les connaissances et à les relier à celles du lecteur par l'utilisation d'hypéronymes (esos 20 embriones humanos pour 20 blastocistos), de paraphrases (**transfert cellulaire** reprend [noyau] **remplacé par celui d'une cellule issue de la même donneuse**) et en refocalisant la notion (**ces embryons clonés, embryons humains**). La recatégorisation de la notion à l'aide de « faux hyperonymes » (el tema et clonación), la manifestation d'une appréciation (**leur [chercheur] réussite** reprend **produire par clonage des embryons humains**) et les paraphrases discursives (la cataphore **objet de leur [chercheur] quête** pour annoncer **cellules souches pluripotentes**) situent ces savoirs spécialisés dans un contexte social. Dans les exemples mentionnés, on reconnaît le caractère polémique du clonage et l'image de la science comme une quête de savoir menant au progrès.

La reformulation (section 5.3.2.) a surtout une fonction métalinguistique, car elle sert à introduire des synonymes (célula madre, célula totipotencial), mais elle est aussi employée pour donner un point de vue différent de la notion (atrasar su reloj biológico par rapport à enseñarle a ese núcleo celular que ahora tenía que comportarse como la primera de todas las células del organismo).

En dépit de leurs différences, la définition et l'explication (sections 5.3.3. et 5.3.4.) témoignent toutes les deux des hypothèses faites par l'auteur quant aux savoirs du lecteur. On juge pertinent de définir ou d'expliquer certaines notions, tandis que d'autres sont supposées suffisamment connues. S'il ne nous est pas possible de déterminer avec certitude le rôle joué par le lieu de publication, nous pouvons imaginer que les notions faisant l'objet d'un tel traitement sont considérées comme clés par rapport au sujet principal. En effet, nous trouvons dans les quatre sous-corpus, des explications de /cellule souche/, /cellule souche pluripotente/ et /créer un embryon humain par clonage dans un but thérapeutique/, qui en plus a été définie dans Argentine et France. Les notions /blastocyste/ et /ovocyte/ ont également été expliquées et définies dans les deux corpus.

D'autres expressions au sein des ensembles paraphrastiques fonctionnent comme des définitions et des explications, mais font partie de paradigmes désignationnels (comme

l'information essentielle apparaît en premier lieu et les détails sont apportés dans la suite du texte (Cassany, D. *et al.* (2000). « La transformación divulgativa de redes conceptuales científicas. Hipótesis, modelo y estrategias », *Discurso y Sociedad*, 2(2), p. 73-103.).

cellule souche qui pourrait permettre d'obtenir des traitements sur mesure contre des maladies comme le diabète ou la maladie de Parkinson, qui appartient à l'ensemble paraphrastique renvoyant à /cellule souche pluripotente/). Dans tous les cas, il y a une focalisation de certains traits conceptuels : dans l'exemple, il s'agit de l'application thérapeutique de ces cellules souches, mais on peut aussi décrire la fonction (célula capaz de formar músculos, huesos, tejidos y neuronas renvoie à /cellule souche pluripotente/) et la forme ou situation spatiale (parte exterior al núcleo définit /cytoplasme/).

Bien qu'ils n'aient pas abordé ces stratégies ou qu'ils ne les définissent pas de la même manière, Cassany et ses collaborateurs (2000) signalent le recours à la modalisation (appréciation) et aux définitions basées sur la fonction ou le but, tandis que Calsamiglia et Van Dijk (2004) analysent les différentes classes conceptuelles servant à la description des notions. Les stratégies discursives n'opèrent donc pas seulement au niveau textuel afin de renforcer la cohérence. Elles ont aussi un impact au niveau cognitif, dans la structuration des connaissances et dans l'instauration de points d'ancrage pour le lecteur. De plus, ces mécanismes permettent de faire le lien entre ces savoirs et le contexte extralinguistique en recatégorisant et refocalisant les notions (recontextualisation). Cibler spécifiquement ces mécanismes dans un prochain travail permettrait de mieux cerner les deux aspects.

6.3. Conceptualisation métaphorique des notions

Le deuxième volet de notre travail porte sur la conceptualisation métaphorique des notions spécialisées. Nous sommes partie de l'hypothèse d'une forte présence d'indices de conceptualisation dans les textes du corpus. En plus des expressions métaphoriques, ces indices se retrouveraient au sein de constructions analogiques. Par ailleurs, nous avons postulé une certaine continuité entre la conceptualisation véhiculée dans la vulgarisation et les discours plus spécialisés. Nous avons aussi anticipé que les modes de conceptualisation seraient influencés par la langue employée et le lieu de publication du texte.

D'emblée, nous pouvons remarquer que la proportion d'indices n'est pas très élevée par rapport à l'ensemble des expressions (section 5.4., figure 50). La place prédominante que différents auteurs accordent à la métaphore dans le discours de vulgarisation (Calsamiglia et Van Dijk 2004, Knudsen 2003) ne se voit donc pas reflétée quantitativement dans les résultats. Comme dans le cas des stratégies discursives étudiées (section 6.2.3.), nous croyons que l'on devrait examiner de plus près l'influence des contraintes posées par le genre discursif sur le recours au mécanisme métaphorique.

Nous comptons toutefois sur des données suffisantes pour constater que le corpus en espagnol et le corpus en français possèdent des caractéristiques similaires autant sur le

plan linguistique (les indices) que sur le plan conceptuel (les modes de conceptualisation). Il est possible également d'identifier des points de contact et des éléments distincts en regard des travaux d'autres chercheurs.

6.3.1. Indices de conceptualisation métaphorique

La plus grande partie des indices de conceptualisation relevés dans les deux corpus (tableau 10) sont des unités prédicatives évoquant un cadre conceptuel source qui se projette sur le cadre cible : celui de la biologie cellulaire, plus précisément celui de la recherche sur le clonage et les cellules souches. Dans la plupart des cas, cette projection se produit par l'intermédiaire de la structure actancielle du prédicat, c'est-à-dire, de la classe d'actants typique sur l'actant réalisé dans le texte⁶⁹. Par exemple, dans l'expression células capaces de formar músculos, huesos, tejidos y neuronas, l'indice capaces témoigne de la conceptualisation des cellules (actant réalisé par células) comme des êtres vivants ayant une volonté propre (classe actancielle typique). On peut remarquer que ce ne sont pas tous les actants qui sont impliqués dans la projection (en italique dans le tableau); dans l'exemple, le deuxième actant de capaces, réalisé par formar músculos, huesos, tejidos y neuronas, appartient à la classe des actions, typique de ce prédicat. Nos résultats sont donc compatibles avec le mécanisme de la projection des actants proposé par Vandaele et Lubin à partir de l'étude de textes scientifiques en français et en anglais (Vandaele et Lubin 2005).

Parfois, cependant, la conceptualisation s'appuie sur une ressemblance de traits entre la notion du cadre source dénotée par l'indice et celle à laquelle on fait référence dans le texte (materia prima et **matière première** pour renvoyer aux cellules employées dans le clonage). Vandaele a relevé ce fonctionnement dans le discours scientifique pour des indices non prédicatifs comme **cellule**, né par ressemblance avec la cellule du moine (Vandaele à paraître).

Autant en espagnol qu'en français, les indices de conceptualisation forment des réseaux lexicaux plus ou moins vastes, selon la saillance cognitive du mode de conceptualisation véhiculé. Divers indices témoignent de la conceptualisation des cellules comme des êtres doués d'intentionnalité (capaz, capacidad, guiar, llevar, poder et reproducirse, en espagnol, et **capable** et **survivre**, en français). De plus, certains indices, comme disparar et **fournir**, sont déjà couramment employés de façon métaphorique; leur présence dans la vulgarisation constitue donc une extension de cet usage qui reste cohérente avec les

⁶⁹ Ce mécanisme de projection a été détaillé à la section 2.6. du « Cadre théorique ».

modes de conceptualisation du domaine (celui des cellules, pour le premier indice, et des embryons, pour le second, comme des êtres humains), mais ne témoigne pas de conceptualisations particulièrement saillantes.

Les textes scientifiques présentent également des réseaux lexicaux d'indices (Vandaele *et al.* 2006) et nous pouvons remarquer certaines correspondances à partir de l'analyse réalisée dans Vandaele et Raffo (2007)⁷⁰. Dans nos corpus, nous trouvons des indices présents dans les réseaux des discours plus spécialisés, comme reparar :

(8) [Texte scientifique :] « *Este mecanismo podría agotarse, por lo tanto las SC exógenas circulantes podrían ser las encargadas de **reparar** el tejido dañado.* »

(9) [Texte de vulgarisation :] « [...] *estas células llamadas stem o indiferenciadas que actualmente se contemplan como un camino para **reparar** órganos dañados y curar desde la diabetes hasta el mal de Parkinson.* »;

des indices qui prolongent le réseau lexical des textes scientifiques par un hyponyme comme fabricar :

(10) [Texte scientifique :] « *Una de las posibles formas de obtener embriones para la utilización de sus células madres es a través de la clonación no reproductiva que supone la transferencia de núcleos en cultivos celulares o en embriones preimplantatorios sin la intención de **producir** un individuo clónico.* »

(11) [Texte de vulgarisation :] « *“La era de la clonación humana aparentemente ya ha llegado: hoy, blastocitos clonados para investigación, mañana blastocitos clonados para **fabricar** bebés”, ha comentado el Dr. Kass al diario estadounidense The New York Times.* »;

et des indices témoignant de modes de conceptualisation absents du discours scientifique, comme receta :

(12) [Texte de vulgarisation :] « [...] *parece claro que, paso a paso, los científicos van a lograr realmente escribir la **receta** completa y detallada de cómo hacer niños clónicos [...].* »

Enfin, la majorité des indices de conceptualisation figurent dans des expressions métaphoriques, mais nous avons pu aussi relever deux constructions analogiques (l'une en français et l'autre en espagnol) qui possèdent des caractéristiques particulières⁷¹. Les indices introduits par des analogies ne sont pas prédicatifs; la conceptualisation n'opère pas par l'intermédiaire d'une projection d'actants, mais grâce à une similitude de traits

⁷⁰ Nous avons présenté en collaboration avec Sylvie Vandaele une comparaison des indices de conceptualisation de textes scientifiques et de vulgarisation dans Vandaele, S. et M. Raffo (2007). *Conceptualización metafórica en el discurso científico y en el de divulgación*, I Congreso Internacional sobre Lenguaje y Asistencia Sanitaria, Universidad de Alicante. Les exemples, tirés de ce travail, reprennent des extraits de notre corpus en espagnol.

⁷¹ Le cas des analogies a également été discuté dans l'analyse mentionnée dans la note précédente (*Ibid.*).

entre ce qui est évoqué par l'indice et la notion conceptualisée. Par exemple, les ovules sont comparés à du riz asiatique, parce qu'ils sont tous les deux « *collants* ». De plus, ces indices ne sont compatibles avec aucun des réseaux lexicaux présents dans les corpus et le mode de conceptualisation véhiculé ne semble fonctionner que localement. À l'opposé des expressions métaphoriques insérées dans des réseaux lexicaux et véhiculant une conceptualisation dont la saillance est reliée à l'étendue de ces réseaux, l'analogie semble servir de mécanisme de création de *nouvelles* conceptualisations et remplir surtout une fonction pédagogique.

Il reste maintenant à se demander si la projection directe impliquant des indices prédicatifs, les indices d'emploi courant métaphorique et les constructions analogiques sont des phénomènes propres au discours de vulgarisation. Nous penchons pour une réponse négative, mais envisageons une étude comparative avec des textes scientifiques qui pourrait nous autoriser à tirer des conclusions fondées.

6.3.2. Modes de conceptualisation

La plupart des modes de conceptualisation qui parcourent nos textes font partie de systèmes cohérents qui témoignent de trois modes plus généraux (section 5.4., tableau 10). D'une part, nous pouvons identifier la conceptualisation des entités biologiques (cellule, embryon, organisme, noyau) comme des êtres doués d'une intentionnalité (des êtres vivants ou des êtres humains). Un grand nombre d'expressions métaphoriques manifestent ce mode de conceptualisation dans les quatre sous-corpus⁷² :

(13) [...] *células madre embrionarias de origen clónico* capaces *de diferenciarse en distintos tejidos del cuerpo humano.* (Espagne)

(14) Digamos, "enseñarle" a ese *núcleo celular* que ahora tenía que comportarse como la primera de todas las células del organismo. (Argentine)

(15) [...] *cellule-souche dite "pluripotente"*, capables *de produire les différents tissus de l'organisme.* (France)

(16) Leur *embryon*, dont le développement a été arrêté à six jours, survit⁷³ depuis maintenant un an sous forme de culture cellulaire. » (Québec)

⁷² Les indices de conceptualisation sont encadrés et les actants renvoyant à la notion conceptualisée sont indiqués en italique.

⁷³ Nous sommes ici dans une zone grise où l'identification de la métaphore dépend du classement des embryons à titre d'êtres vivants ou comme des entités biologiques. C'est là, d'ailleurs, le fondement des débats sur le clonage humain. Voir plus loin.

D'autre part, les indices de la réification sont très courants. Les entités biologiques et les êtres humains sont conçus comme des objets, des machines, des produits manufacturés ou des fournitures :

- (17) [...] hoy, blastocitos clonados para investigación, mañana blastocitos clonados para fabricar bebés [...]. (Espagne)
- (18) [...] dejaron el manejo de células que formarían el feto. (Argentine)
- (19) Cette technique [...] pourrait permettre un jour de réparer l'*être humain* atteint d'une maladie incurable [...]. (France)
- (20) Un premier *embryon humain* est produit par clonage. » (Québec)

Ces deux modes de conceptualisation qui semblent contradictoires coexistent cependant dans les textes grâce à un phénomène d'intégration conceptuelle (Vandaele 2006). Nous pouvons toutefois voir une certaine tension entre la conceptualisation des embryons comme des produits manufacturés et celle qui les conçoit comme des êtres vivants, car elles comportent une dimension éthique et idéologique. En effet, à la lumière des débats acharnés sur le clonage humain et l'expérimentation sur les cellules souches, il n'est pas innocent de faire appel à l'une ou à l'autre, comme on peut le voir dans ce passage :

- (21) Sin embargo, no todos aceptan que, con el fin de curar enfermos, se puedan descartar embriones. Ayer, grupos como la organización cristiana antiabortista Operación Rescate, de los Estados Unidos, condenaron el estudio y expresaron: "Clonar para matar es moralmente inaceptable". (Argentine)

La projection du cadre conceptuel de l'agriculture sur la manipulation des cellules et des embryons se manifeste également dans les quatre sous-corpus :

- (22) [...] a partir de ellos se cultivaron 30 *blastocitos* [...]. (Espagne)
- (23) Cuando estas células se cultivan [...]. (Argentine)
- (24) [...] **produire ses cellules souches en culture**. (France)
- (25) [...] **culture cellulaire**⁷⁴. » (Québec)

D'autres modes de conceptualisation spécifiques figurent dans les différents sous-corpus, mais sont beaucoup moins saillants. Mentionnons seulement que, dans Argentine, il est possible de remarquer que le processus de la différenciation cellulaire et le stade de blastocyste sont conceptualisés comme des lieux : nous retrouvons ici le raisonnement d'une notion abstraite en termes d'une expérience physique concrète. La projection du coup de feu sur le rejet immunologique (aussi dans Argentine) met l'accent sur le début soudain plutôt que sur le processus lui-même.

⁷⁴ L'adjectif relationnel « cellulaire » renvoie à l'actant « cellule ».

Les travaux de Vandaele et ses collaborateurs (notamment Vandaele *et al.* 2006), nous permettent de contraster nos résultats avec ceux obtenus à partir de textes du domaine biomédical. Le mode de conceptualisation prédominant identifié dans le discours scientifique est celui des molécules et des cellules comme des personnes ou des êtres vivants, manifesté par des indices comme *capable* et *survie*. Dans nos corpus, cette conceptualisation est également la plus représentée et nous avons repéré des indices faisant partie du même réseau lexical : **capable** et **survivre**. Un autre mode de conceptualisation commun concerne les composantes de la cellule, vues comme les parties d'une machine remplissant une fonction. D'ailleurs, l'indice relevé dans le corpus Espagne (maquinaria) serait un équivalent de celui employé dans des textes scientifiques (*machinerie*). Enfin, Vandaele et ses collaborateurs rendent compte de la conceptualisation des processus biologiques - une suite d'événements - comme un chemin (*voie*) ou comme un mouvement d'entités dans l'espace (*cascade*). Ces modes de conceptualisation seraient liés à la vision spatiale du processus de différenciation cellulaire que nous avons mentionné plus haut et dont témoigne guiar dans Argentine.

Ainsi, nous pouvons voir une continuité entre les textes de nos corpus et un corpus de textes scientifiques, non seulement au niveau des modes de conceptualisation, mais aussi en ce qui concerne les indices (section précédente). Cela semble vouloir dire que les frontières entre les deux discours ne sont pas étanches, ce qui est cohérent avec les résultats de Knudsen (2003), qui montre que le même mode de conceptualisation est véhiculé dans des corpus de textes scientifiques et de vulgarisation traitant du code génétique. Elle suggère que la différence entre les deux corpus réside dans le but dans lequel le mécanisme métaphorique est utilisé. Nous constatons également que les réseaux d'indices scientifiques, sous-jacents à la phraséologie et à la terminologie du domaine, se retrouvent également dans les discours moins spécialisés. Pour notre part, nous croyons que la métaphore, en tant que mécanisme de structuration de connaissances, n'est pas employée dans la vulgarisation uniquement dans un but pédagogique, mais qu'elle peut servir diverses intentions communicatives, comme la refocalisation de la notion à des fins idéologiques. Une étude comparative permettrait de vérifier ces hypothèses en déterminant la coïncidence des divers modes de conceptualisation et en examinant la constitution des réseaux lexicaux.

6.4. Structures notionnelles

Notre travail ne vise pas une analyse de contenu des corpus, cependant, nous tenons à attirer l'attention sur certains aspects liés aux structures notionnelles fournies par les textes. En ce qui concerne la quantité d'information véhiculée, il est possible de remarquer que le corpus en espagnol introduit un plus grand nombre de notions que le corpus en français (section 5.1.1.). Cet écart s'explique par la différence dans le nombre de mots de chaque corpus, car malgré une variabilité intertexte élevée, nous trouvons de chiffres similaires autant pour la densité notionnelle que pour la densité d'actualisations (sections 5.1.1. et 5.1.2.). On pourrait se demander s'il existe une tendance dans la presse en espagnol à présenter des textes de vulgarisation plus longs offrant un contenu thématique plus vaste. Pour répondre à cette question, il faudrait compter sur un corpus plus vaste. En outre, il serait intéressant de comparer ces densités à celles des articles scientifiques afin de mesurer l'influence du degré de spécialisation du texte sur la quantité d'information véhiculée.

Quant aux notions véhiculées, nous constatons que le contenu des différents textes est assez homogène quoiqu'ils ne présentent pas tous les mêmes notions (section 5.1.3.)⁷⁵. En effet, les textes de chacun des corpus partagent une bonne partie des notions - entre 60 % et 80 % en moyenne - et, qui plus est, ces notions communes sont à l'origine de la plupart des actualisations - entre 70 % et presque 90 % en moyenne. De plus, ce que nous avons appelé le noyau thématique (les 35 notions actualisées dans les quatre sous-corpus) comprend une portion réduite de l'ensemble des notions (19,02 %), mais représente plus de la moitié (61,58 %) des actualisations. Il reflète le sujet principal des textes : l'obtention par clonage de cellules souches embryonnaires humaines dans un but thérapeutique. Cette homogénéité thématique résulte de la décision de compiler les textes en fonction du moment discursif⁷⁶.

En même temps, beaucoup de notions sont spécifiques à l'un des corpus, voire à l'un des sous-corpus. L'écart thématique est plus prononcé dans le corpus en français que dans le corpus en espagnol (figures 10 à 13). Pourtant, à première vue, il n'est pas possible de corrélérer le choix des notions à la langue employée ou au contexte socioculturel. Un grand nombre de notions ne sont actualisées qu'une ou deux fois et nous ne pouvons pas juger des motivations des auteurs pour les inclure dans le texte. Faute d'avoir dégagé une tendance forte, déterminer l'existence d'une contrainte culturelle demanderait une

⁷⁵ Nous discutons de la signification méthodologique de l'homogénéité des corpus et nous la mettons en perspective avec les travaux d'autres chercheurs à la section 6.1.1. de la « Discussion ».

⁷⁶ La notion de moment discursif a été introduite à la section 4.1. de la « Méthodologie ».

laborieuse analyse de chaque notion. Par ailleurs, il est logique d'espérer qu'un corpus comportant plus de textes traitant du même sujet, serait plus homogène.

Enfin, il faut souligner la grande similitude des proportions de chaque hyper-catégorie dans les différents sous-corpus. Les actions, les entités et les processus sont privilégiés par rapport aux phénomènes, aux états et aux propriétés. Étant donné que les articles journalistiques traitent d'un événement (dans ce cas, l'expérience des chercheurs coréens), on pourrait y voir une préférence pour la narration au détriment de la description. Ces données et des observations préliminaires nous portent à croire que le genre discursif aurait un impact sur le type de notions présentées. D'ailleurs, Ciapuscio (2003a) signale pour son corpus une disparité de contenu entre des textes de types différents qu'elle met en relation avec la fonction et la situation de communication.

Pour une analyse plus fine des réseaux notionnels, il serait envisageable de prolonger la présente recherche en réalisant une étude plus exhaustive du contenu des textes. Le travail de Ciapuscio (2003a) mentionné propose une démarche méthodologique dont on pourrait s'inspirer. Par ailleurs, une étude comparative entre des textes de vulgarisation et des textes scientifiques serait nécessaire afin d'examiner les différences reliées au degré de spécialisation.

7. Conclusion

Comme nous l'avons constaté dans l'introduction, la caractérisation des discours spécialisés en vue de la traduction constitue un défi en raison du nombre de variables à considérer. Un travail exploratoire était indispensable afin de dégager les pistes pour une recherche approfondie. Dans ce dessein, il nous a fallu forger une méthodologie adaptée à l'étude de la mise en discours des notions spécialisées dans une optique bilingue. Autrement dit, autorisant la mise en correspondance de l'expression des notions spécialisées dans deux langues (le français et l'espagnol) tout en tenant compte des paramètres liés à une situation de communication spécifique : la vulgarisation d'une découverte scientifique dans la presse. À la fin de cette étape préliminaire, nous pouvons dresser un bilan positif. La démarche méthodologique mise au point nous a permis de recueillir des données non seulement nombreuses, mais pertinentes au regard des objectifs de ce travail. Ainsi, nous avons pu mettre en évidence des différences et des similitudes au sein de nos corpus ainsi qu'entre nos textes de vulgarisation et les textes scientifiques.

Les lignes de divergence se situent premièrement dans la grande dispersion du nombre d'expressions employées pour renvoyer à une même notion. Pourtant, nous avons décelé une régularité liant ce nombre à l'hyper-catégorie sous laquelle la notion peut être rangée : la variation semble être plus élevée pour les actions et les processus que pour les entités. Voici donc une première hypothèse méritant d'être examinée lors d'une recherche ultérieure qui devrait s'appuyer sur un plus gros corpus et faire appel à une analyse statistique appropriée. Un deuxième point de dissimilitude concerne les unités et syntagmes faisant référence à la même notion. En discours, les ensembles paraphrastiques sont variés et des rapports divers se tissent entre les unités et syntagmes qui en font partie. Il reste maintenant à déterminer le rôle joué par les paramètres textuels (cohérence, structuration des connaissances dans le texte) et situationnels (le lieu de parution du texte, par exemple). Les critères de compilation du corpus et de repérage des expressions devraient être ajustés en conséquence. Quant aux spécificités propres à la langue et au contexte socioculturel, elles pourront être caractérisées par des études permettant de mieux isoler les variables en jeu.

Parallèlement, nous trouvons des lignes de convergence très marquées qui vont à l'encontre de quelques idées généralement admises sur le discours de vulgarisation. D'une part, nos résultats contredisent les affirmations élevant certaines stratégies discursives que l'on pourrait regrouper sous le nom d'« explicatives » (définition, explication, etc.) et

l'emploi d'expressions métaphoriques au rang de traits caractéristiques des textes de vulgarisation. Nous croyons qu'une étude visant spécifiquement ces mécanismes pourrait dresser un portrait plus précis de ce type de discours, notamment en tenant compte des différents genres discursifs. D'autre part, nous avons pu remarquer une continuité entre le discours à vocation vulgarisatrice et le discours scientifique autant sur le plan linguistique, par le recours dans les textes de vulgarisation à des dénominations propres au domaine de spécialité, que conceptuel, dans le continuum des modes de conceptualisation et des réseaux lexicaux qui en témoignent. Ces points de contact doivent être l'objet d'une étude comparative visant une analyse fine des particularités pouvant être reliées à la langue et au lieu de parution des textes.

Bibliographie

- ATILF (2004) *Le Trésor de la Langue Française informatisé (TLFi)*. [en ligne]. Disponible à <http://atilf.atilf.fr/tlf.htm>. Consulté le 10 novembre 2007.
- Balliu, C. (2005). « La didactique de la traduction médicale, deux ou trois choses que je sais d'elle », *Meta*, 50(1), p. 67-77.
- Béjoint, H. et P. Thoiron (dirs.) (2000). *Le sens en terminologie*, Lyon: Presses Universitaires de Lyon.
- Boyd, R. (1993). « Metaphor and theory change: What is "metaphor" a metaphor for? ». Dans A. Ortony (dir.), *Metaphor and Thought* (2^e ed.), Cambridge: Cambridge University Press.
- Cabré, M. T. (1999). *La terminología, representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos*, Barcelona: IULA.
- (2000). « Sur la représentation mentale des concepts : bases pour une tentative de modélisation ». Dans H. Béjoint et P. Thoiron (dirs.), *Le sens en terminologie*, Lyon: Presses Universitaires de Lyon.
- (2003). « Theories of terminology: their description, prescription and explanation », *Terminology*, 9(2), p. 163-199.
- Calsamiglia, H. et al. (1998). *Análisis discursivo de la divulgación científica*, I Simposio Internacional de Análisis del Discurso, Madrid.
- Calsamiglia, H. et D. Cassany (1999). « Voces y conceptos en la divulgación científica », *Revista Argentina de Lingüística*, 15, p. 173-208.
- Calsamiglia, H. et T. A. Van Dijk (2004). « Popularization Discourse and Knowledge about the Genome », *Discourse Studies*, 15(4), p. 369-389.
- Carreño, I. (2004). *Analyse de la variation terminologique en corpus parallèle anglais-espagnol et de son incidence sur l'extraction de termes bilingue*, Mémoire de Maîtrise, Université de Montréal, Montréal.
- Cassany, D. et J. Martí (1998). « Estrategias divulgativas del concepto prión », *Quark*, 12, p. 58-66.
- Cassany, D. et al. (2000). « La transformación divulgativa de redes conceptuales científicas. Hipótesis, modelo y estrategias », *Discurso y Sociedad*, 2(2), p. 73-103.
- Cataldi, C. (2004). « El debate sobre los transgénicos en la prensa española: cómo los actores sociales denominan esta biotecnología », *Quark*, 33, p. 57-68.
- Charaudeau, P. (1983). *Langage et discours. Éléments de sémiolinguistique*, Paris: Hachette Université.
- Charaudeau, P. et D. Maingueneau (dirs.) (2002). *Dictionnaire d'analyse du discours*, Paris: Seuil.

- Ciapuscio, G. E. (2003a). *Textos especializados y terminología*, Barcelona: Institut Universitari de Lingüística Aplicada.
- (2003b). « Formulation and reformulation procedures in verbal interactions between experts and (semi-)laypersons », *Discourse Studies*, 5(2), p. 207-233.
- Collet, T. (2000). *La réduction des unités terminologiques complexes de type syntagmatique*, Thèse de Doctorat, Université de Montréal, Montréal.
- Cordier, F. et J. François (dirs.) (2002). *Catégorisation et langage*, Paris: Hermès Science Publications.
- de Beaugrande, R.-A. et W. Dressler (1981). *Introduction to Text Linguistics* (R.-A. de Beaugrande), London: Longman.
- Depecker, L. (2005). « Contribution de la terminologie à la linguistique », *Langages*, 157, p. 6-12.
- Ducrot, O. et T. Todorov (1972). *Dictionnaire encyclopédique des sciences du langage*, Paris: Seuil.
- Ducrot, O. (1980). *Les mots du discours*, Paris: Éditions de Minuit.
- Fauconnier, G. et M. Turner (1998/2002). *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, New York: Basic Books.
- Fillmore, C. J. (1977). « Scenes-and-frames semantics ». Dans A. Zampolli (dir.), *Linguistic Structures Processing*, Amsterdam: North Holland.
- Fox Keller, E. (1995). *Refiguring Life: Metaphors of Twentieth Century Biology*, New York: Columbia University Press.
- (1999). *Le rôle des métaphores dans les progrès de la biologie*, Paris: Institut Sanofi-Synthélabo.
- François, J. et S. Mejri (dirs.) (2006). *Composition syntaxique et figement lexical*, Caen: Presses universitaires de Caen.
- Freixa, J. (2002). *La variació terminològica. Anàlisi de la variació denominativa en textos de diferent grau d'especialització de l'àrea de medi ambient*, Thèse de Doctorat, Universitat de Barcelona, Barcelona.
- Fuchs, C. (1982). *La paraphrase*, Paris: Presses Universitaires de France.
- Gambier, Y. (1991). « Travail et vocabulaire spécialisés : prolégomènes à une socio-terminologie », *Meta*, 36(1), p. 9-15.
- Gaudin, F. (1993). *Pour une socioterminologie : Des problèmes pratiques aux pratiques institutionnelles*, Rouen: Publications de l'Université de Rouen.
- Gläser, R. (1993). « A Multi-level Model for a Typology of LSP Genres », *Fachsprache*, 15(1-2), p. 18-26.
- Gülich, E. et T. Kotschi (1983). « Les marqueurs de la reformulation paraphrastique », *Cahiers de linguistique française*, 5, p. 305-351.

- Hall, C. *et al.* (1999). « Speech Representation in Social Work Discourse », *Text*, 19(4), p. 539-570.
- Heinemann, W. et D. Viehweger (1991). *Textlinguistik: eine Einführung*, Tübingen: Niemeyer.
- Jacobi, D. et B. Schiele (1988). « La vulgarisation scientifique. Thèmes de recherche ». Dans D. Jacobi et B. Schiele (dirs.), *Vulgariser la science. Le procès de l'ignorance*, Seyssel: Champ Vallon.
- Kleiber, G. (1984). « Dénomination et relations dénominatives », *Langages*, 76, p. 77-94.
- (1993). *Anaphores et pronoms*, Louvain-la-Neuve: Duculot.
- Knudsen, S. (2003). « Scientific metaphors going public », *Journal of Pragmatics*, 35(8), p. 1247-1263.
- Kocourek, R. (1991). *La langue française de la technique et de la science* (2^e éd.), Wiesbaden: Brandstetter.
- Koller, V. (2004). *Metaphor and Gender in Business Media Discourse (A Critical Cognitive Study)*, New York: Palgrave MacMillan.
- Kordon, C. (1991). *Le langage des cellules*, Paris: Hachette.
- L'Homme, M.-C. (2005). « Sur la notion de "terme" », *Meta*, 50(3), p. 1112-1132.
- Lakoff, G. et M. Johnson (1980/2003). *Metaphors we live by* (2nd éd.), Chicago: The University of Chicago Press.
- Le Ny, F. (2005). *Comment l'esprit produit du sens*, Paris: Odile Jacob.
- Liakopoulos, M. (2002). « Pandora's Box or panacea? Using metaphors to create the public representations of biotechnology », *Public Understanding of Science*, 11, p. 5-32.
- Loffler-Laurian, A.-M. (1983). « Typologie des discours scientifiques : deux approches », *Études de Linguistique Appliquée*, 51, p. 8-20.
- (1994). « Réflexions sur la métaphore dans les discours scientifiques de vulgarisation », *Langue française*, 101, p. 72-79.
- Maillard, M. (1974). « Essai de typologie des substituts diaphoriques », *Langue française*, 21, p. 55-71.
- Mel'čuk, I. *et al.* (1995). *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve: Duculot.
- Méndez Cendón, B. et B. López Arroyo (2003). « Intralinguistic analysis of medical research papers and abstracts », *Terminology*, 9(2), p. 247-268.
- Meyer, I. *et al.* (1998). *Metaphorical Internet Terminology in English and French*, Proceedings of the 9th EURALEX International Congress, Liège.
- Moirand, S. (2003). « Communicative and cognitive dimensions of discourse on science in the French mass media », *Discourse Studies*, 5(2), p. 175-206.
- Mortureux, M.-F. (1993). « Paradigmes désignationnels », *Semen*, 8, p. 117-136.

- Nabokov, V. (1955). « Problems of Translation: *Onegin* in English », *Partisan Review*, 22(4), p. 496-512.
- Nord, C. (1997). *Translating as a Purposeful Activity*, Manchester: St Jerome Publishing.
- Pavel, S. (1993). « La phraséologie en langue de spécialité. Méthodologie de consignation dans les vocabulaires terminologiques », *Terminologies nouvelles*, 10, p. 67-82.
- Phal, A. (1971). *Vocabulaire général d'orientation scientifique (V.G.O.S.) : part du lexique commun dans l'expression scientifique*, Paris: CREDIF.
- Reiss, K. et H. J. Vermeer (1984/1991). *Grundlegung einer allgemeinen Translationstheorie*, Tübingen: Niemeyer.
- Rondeau, G. (1984). *Introduction à la terminologie* (2^e éd.), Chicoutimi: G. Morin.
- Sager, J. C. (1990). *A Practical Course in Terminology Processing*, Amsterdam/Philadelphia: John Benjamins.
- (2000). « Pour une approche fonctionnelle de la terminologie ». Dans H. Béjoint et P. Thoiron (dirs.), *Le sens en terminologie*, Lyon: Presses Universitaires de Lyon.
- Saint Jérôme (1953). « À Pammachius. La meilleure méthode de traduction ». Dans *Saint Jérôme. Lettres* (t. é. e. t. p. J. Labourt) (Vol. 3), Paris: Les Belles Lettres.
- Schleiermacher, F. (1985). « Des différentes méthodes du traduire ». Dans A. Berman (dir.), *Les Tours de Babel: Essais sur la traduction*, Mauvezin: Trans-Europ-Repress.
- Slodzian, M. (2000). « L'émergence d'une terminologie textuelle et le retour du sens ». Dans H. Béjoint et P. Thoiron (dirs.), *Le sens en terminologie*, Lyon: Presses Universitaires de Lyon.
- Suárez, M. M. (2004). *Análisis contrastivo de la variación denominativa en textos especializados: del texto original al texto meta*, Doctorado, Universitat Pompeu Fabra, Barcelona.
- Talmy, L. (2001). *Toward a cognitive semantics. Volume I: Concept structuring systems.*, Cambridge: The MIT Press.
- Temmerman, R. (2000). *Towards New Ways of Terminology Description, the Sociocognitive Approach*, Amsterdam/Philadelphia: John Benjamins.
- (2002a). « Metaphorical models and the translator's approach to scientific texts », *Linguistica Antverpiensia*, 1, p. 211-226.
- (2002b). « Metaphorical Models and the Translation of Scientific Texts », *Linguistica Antverpiensia*, 1, p. 211-226.
- Tesnière, L. (1959/1965). *Éléments de syntaxe structurale*, Paris: Klincksieck.
- Van Dijk, T. A. (2003). « Specialized discourse and knowledge: A case study of the discourse of modern genetics », *Cadernos de Estudos Lingüísticos* 44, p. 21-55.
- Van Rijn-Van Tongeren, G. W. (1997). *Metaphors in medical texts*, Amsterdam/Atlanta: Rodopi.

- Vandaele, S. (2004). *Analyse et représentation de la conceptualisation métaphorique en langue de spécialité à l'aide des bases de données relationnelles*, 11th EURALEX International Congress, EURALEX 2004, Lorient.
- (2005). *Métaphores conceptuelles et fonctions lexicales : des outils pour la traduction médicale et scientifique*, III^e Congreso Internacional de Traducción Especializada, Barcelona.
- Vandaele, S. et L. Lubin (2005). « Approche cognitive de la traduction dans les langues de spécialité : vers une systématisation de la description de la conceptualisation métaphorique », *Meta*, 50(2), p. 415-431.
- Vandaele, S. (2006). « Quelques repères épistémologiques pour une approche cognitive de la traduction spécialisée - Application à la biomédecine », *Meta*, 52(1), p. 129-141.
- Vandaele, S. et S. Boudreau (2006). *Annotation XML et interrogation de corpus pour l'étude de la conceptualisation métaphorique*, JADT 2006 : 8^{es} Journées internationales d'analyse statistique des données textuelles, Besançon.
- Vandaele, S. et al. (2006). « La conceptualisation métaphorique en biomédecine : indices de conceptualisation et réseaux lexicaux », *Glottopol*, 8, p. 73-94.
- Vandaele, S. et M. Raffo (2007). *Conceptualización metafórica en el discurso científico y en el de divulgación*, I Congreso Internacional sobre Lenguaje y Asistencia Sanitaria, Universidad de Alicante.
- Vandaele, S. (à paraître). « Les modes de conceptualisation du vivant : une approche linguistique »
- Vandaele, S. et S. Boudreau (à paraître). « Un multi-outil adapté au parcours cognitif de l'étudiant en traduction spécialisée : application à la biomédecine »
- Vandaele, S. et M. Raffo (à paraître). « Entre sens et notion : unités lexicales prédicatives et syntagmes terminologiques »
- Vandaele, S. et al. (à paraître). « Les défis de la traduction spécialisée : mise en œuvre d'un site de référence en biomédecine »
- Weise, G. (1993). « Criteria for the Classification of ESP Texts », *Fachsprache*, 15(1-2), p. 26-31.
- Wichter, S. (1994). *Experten-und Laienwortschätze. Umriss einer Lexikologie der Vertikalität*, Tübingen: Niemeyer.

