



**HAL**  
open science

# Valorisation des analogies lexicales entre l'anglais et les langues romanes : étude prospective pour un dispositif plurilingue d'apprentissage du FLE dans le domaine de la santé

Fabrice Gilles

## ► To cite this version:

Fabrice Gilles. Valorisation des analogies lexicales entre l'anglais et les langues romanes : étude prospective pour un dispositif plurilingue d'apprentissage du FLE dans le domaine de la santé. Linguistique. Université Grenoble Alpes, 2017. Français. NNT : 2017GREAL011 . tel-01719853

**HAL Id: tel-01719853**

**<https://theses.hal.science/tel-01719853>**

Submitted on 14 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

### **DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES**

en **Sciences du langage**, spécialité : **Didactique et Linguistique**

Arrêté ministériel : 25 mai 2016

Présentée par

**Fabrice GILLES**

Thèse dirigée par **Christian Degache**

codirigée par **Olivier Kraif**

préparée au sein du **Laboratoire LIDILEM – EA 609**

dans l'**École Doctorale Langues, Littérature et Sciences Humaines**

## **Valorisation des analogies lexicales entre l'anglais et les langues romanes : étude prospective pour un dispositif plurilingue d'apprentissage du FLE dans le domaine de la santé**

Thèse soutenue publiquement le **29 septembre 2017** devant le jury  
composé de :

**Mme Cristelle Cavalla**

Maitre de conférences à l'Université Paris 3, Rapporteur

**M. François Maniez**

Professeur à l'Université Lyon 2, Président et Rapporteur

**Mme M<sup>a</sup> Isabel González Rey**

Professeure à l'Universidade de Santiago de Compostela, Membre

**M. Franz-Joseph Meißner**

Professeur émérite à la Justus-Liebig-Universität, Examineur

**Mme Agnès Tutin**

Professeure à l'Université Grenoble Alpes, Examinatrice

**M. Christian Degache**

Professeur à l'Université Grenoble Alpes et à l'Universidade Federal de  
Minas Gerais, Directeur de thèse

**M. Olivier Kraif**

Maitre de conférence à l'Université Grenoble Alpes, Codirecteur de  
thèse





# Remerciements

J'aimerais exprimer toute ma reconnaissance à Christian Degache et à Olivier Kraif, qui se sont partagé la direction de cette thèse. Grâce à vous, je suis entré dans la pédagogie numérique en apprenant à utiliser des corpus qui se prêtent à l'élaboration d'activités d'apprentissage des langues, et en intervenant lors de formations plurilingues en ligne. C'est un bouleversement méthodologique pour moi qui n'avais pas été formé dans ce domaine.

Je remercie les rapporteurs, Cristelle Cavalla et François Maniez, ainsi que les autres membres du jury, Maria Isabel González Rey, Franz-Joseph Meißner et Agnès Tutin, qui ont accepté de lire cette thèse et de participer à son évaluation.

Merci aux collègues du programme de recherche MIRIADI (Mutualisation et Innovation dans un Réseau d'Intercompréhension À Distance) pour la réflexion et l'action communes.

Merci à Zohra Bouhania et Isabelle Rousset, du laboratoire LIDILEM, merci à mes collègues enseignants ainsi qu'aux doctorant.e.s Amadou, Carmenne, Hee-Kyung, Jorge, Nahla, Rui et Wenjun.

Merci à M. Cenefels, mon professeur d'allemand au lycée dans les années 1980, qui par son enseignement m'a donné envie de comparer les langues.

Merci aux relecteurs : Aurélien, Isabelle, Jacqueline, Jean-Pierre, Nathalie, Pierre, Roxane et Rui. Merci aussi à Björn, Tamara, Hélène et Odile.

Je remercie chaleureusement Guillaume Théret, chercheur à l'Institut des Mathématiques de Bourgogne, pour les discussions qui ont inspiré le projet de cette thèse.

Enfin, je sais tout ce que je dois à des proches dont le soutien a été constant. Mille mercis à Cyril, Christophe et Sandrine.

# Résumé

Cette étude lexicologique prospective s'inscrit dans la didactique des L3. L'objectif est d'élaborer un interlexique anglais-espagnol-français-italien-portugais composé des adjectifs, noms et verbes anglais fréquents dans les écrits scientifiques de la santé, et de leurs équivalents de traduction analogues en espagnol, français, italien et portugais. Deux mots sont analogues s'ils ont le même sens et une forme similaire.

Les rapports entre les concepts d'analogie, de similarité et d'identité sont examinés, les types d'analogies intralinguistiques et interlinguistiques illustrés et les principales analogies et dissemblances entre l'anglais, le français et les langues romanes exposées. L'existence de celles-ci est justifiée par les origines indoeuropéennes et surtout d'intenses contacts de langues. Après avoir rappelé l'importance de l'analogie dans l'apprentissage, nous montrons le lien entre notre recherche et deux types d'approches didactiques des langues : l'intercompréhension, qui développe la compréhension de langues voisines, et les approches sur corpus qui permettent de mieux connaître et faire connaître la phraséologie scientifique.

Les 2000 lemmes anglais les plus fréquents ont été extraits du corpus scientifique anglais de ScienText, leurs 2208 acceptions fréquentes délimitées sur la base du profil combinatoire et triées en deux catégories sémantiques : lexique de spécialité et lexique scientifique transdisciplinaire. Les lemmes anglais ont été traduits dans les quatre langues romanes, et la similarité mesurée en fonction de la sous-chaine commune maximale (SCM).

L'interlexique contient des mots qui expriment 47 % des acceptions fréquentes. Par couples de langues, l'analogie est encore plus élevée : anglais – français, 66 %, anglais-italien, 65 %, anglais-espagnol, 63 %, anglais-portugais, 58 %. Ce lexique analogue pourrait donc servir comme base de transfert dans des activités de FLE L3 pour des professionnels de la santé, et l'anglais L2 semble être une passerelle possible vers les langues romanes. Des activités plurilingues sont construites sur des concordances extraites des corpus multilingues alignés EMEA et Europarl. Un questionnement métalinguistique en anglais sensibilise à des traits (morpho)syntaxiques du français ; les analogies des deux langues sont systématiquement mises en relief, et dans les cas d'opacité, celles des autres langues romanes avec l'anglais.

Mots clés : analogies lexicales anglais-langues romanes, FOS / français de la santé, corpus multilingues alignés, profil combinatoire, sous-chaine commune maximale (SCM), langue passerelle

# Abstract

This prospective lexicological investigation belongs to the field of L3 French didactics. The purpose is to elaborate a French-Italian-Portuguese-Spanish interlexicon out of the frequent adjectives, nouns and verbs of the healthcare scientific writings, and their analogue translation equivalents in French, Italian, Portuguese and Spanish. Two words are analogue if they have the same meaning and a similar form.

Related concepts of analogy, similarity and identity are discussed, types of intralinguistic and cross-linguistic analogies reviewed, and the main analogies and differences between English, French and Romance languages detailed. Their many analogies are justified by Indo-European origins and mostly by intense language contacts. Once the importance of analogy in learning procedures has been highlighted, we show how this research and two types of didactic approaches connect together: intercomprehension, which develops comprehension skills in neighbor languages, and corpus approaches which enable to get a closer insight into scientific phraseology.

The 2000 most frequent English lemmas were extracted from the ScienText English scientific corpus, their 2208 frequent acceptions explored from their combinatory profile and sorted out in two semantic categories: healthcare subject-specific vocabulary and science specific trans-disciplinary vocabulary. The English lemmas were translated into the four Romance languages, and similarity measurements were carried out with the longest common substring method.

The interlexicon contains 47% of the frequent acceptions. Analogy is even higher by language pairs: English – French, 66%, English – Italian, 65%, English - Spanish, 63%, English – Portuguese, 58%. Consequently, this analogue vocabulary could form a transfer basis in learning activities of L3 French for health care providers, and L2 English seems to be a possible bridge language toward Romance languages. Plurilingual activities are built on concordances extracted from multilingual aligned corpora (EMEA, Europarl). Metalinguistic questions in English point out (morpho)syntactic features of French; the analogies between both languages are systematically enhanced, and in case of lexical opacity, those between English and the other Romance languages.

English title : Emphasising lexical analogies between English and Romance languages : prospective study towards a plurilingual learning device of French for healthcare

Keywords: lexical analogies (English-Romance languages), French for Specific Purposes / French for healthcare, multilingual aligned corpora, combinatory profile, longest common substring, bridge language

Laboratoire LIDILEM – EA 609

Université Grenoble Alpes

CS40700

38058 Grenoble cedex 9

# Table des matières

## Table des matières

Remerciements.....	3
Résumé.....	4
Abstract.....	5
Table des matières.....	7
Index des images.....	15
Index des graphiques.....	16
Index des tableaux.....	17
Introduction.....	18
Première partie - Analogies linguistiques : concepts voisins, origines et valorisation par des approches didactiques innovantes.....	25
Chapitre 1 - Analogies linguistiques : concepts proches et analogies de l'anglais et des langues romanes.....	27
1.1 Introduction du chapitre 1.....	27
1.2 Distinctions opératoires autour des concepts proches de similarité, d'identité et d'analogie.....	28
1.2.1 Analogie : fécondité d'un concept ambigu.....	28
1.2.2 Essai de clarification des concepts de similarité et d'identité.....	29
1.2.2.1 Part commune et part distincte.....	29
1.2.2.2 Un continuum similarité - identité dans la langue courante.....	29
1.2.2.3 Similarité / identité : ambigüité de certains mots.....	31
1.2.3 L'analogie, un rapport à quatre ou trois termes.....	31
1.2.3.1 Analogie vs similarité.....	31
1.2.3.2 Cas de la proportion discontinue.....	31
1.2.3.3 Cas de la proportion continue.....	32
1.3 De la confusion des concepts de similarité, d'identité et d'analogie appliqués au lexique.....	33
1.3.1 Des outils pour la comparaison lexicale intralinguistique.....	33
1.3.2 Mesurer la similitude / l'analogie interlinguistiques.....	35
1.3.2.1 Que comparer ? Forme et/ou sens ?.....	35
1.3.2.2 Examen de deux classifications utilisées en intercompréhension.....	37
1.3.2.3 Outils de mesure de similarité.....	40
1.3.2.4 Définitions : intersystème, interlexème, interlexique, interforme et intersémème.....	41
1.3.3 Un type de similarité lexicale interlinguistique : les <i>faux amis</i> .....	43
1.3.3.1 Définition habituelle de <i>faux amis</i> et failles de cette définition selon Kauffer.....	43
1.3.3.2 Réflexion et propositions autour des faux amis.....	44
1.4 Approche des principaux types d'analogies intralinguistiques et interlinguistiques.....	48
1.4.1 Analogies syntaxiques.....	48
1.4.1.1 Analogies syntaxiques intralinguistiques.....	48
1.4.1.2 Analogies syntaxiques interlinguistiques.....	48
1.4.2 Analogies morphosyntaxiques.....	49
1.4.2.1 Analogies morphosyntaxiques intralinguistiques.....	49



1.4.2.2 Analogies morphosyntaxiques interlinguistiques.....	49
1.4.3 Analogies morphologiques.....	50
1.4.3.1 Analogies morphologiques intralinguistiques.....	50
1.4.3.2 Analogies morphologiques interlinguistiques.....	51
1.5 Principales analogies et dissemblances entre l'anglais et les langues romanes.....	52
1.5.1 Un vaste champ d'étude où l'histoire de la langue est omniprésente.....	52
1.5.1.1 Ce que nous n'abordons qu'accessoirement.....	52
1.5.1.2 Incidence phonéticohistorique de l'accent à l'initiale des langues germaniques.....	53
1.5.2 Analogies et dissemblances syntaxiques de l'anglais et des langues romanes..	54
1.5.3 Analogies morphologiques de l'anglais et des langues romanes.....	55
1.5.3.1 Analogies lexicales de l'anglais et des langues romanes.....	55
1.5.3.2 Particularités de la dérivation de verbes anglais.....	56
1.5.3.3 Analogies du lexique de la santé et de la chimie.....	56
1.5.4 Analogies et dissemblances morphosyntaxiques.....	57
1.5.4.1 Analogies et dissemblances morphosyntaxiques des noms.....	57
1.5.4.2 Analogies et dissemblances morphosyntaxiques des adjectifs et des adverbess.....	58
1.5.4.3 Analogies et dissemblances morphosyntaxiques des verbes de l'anglais et des langues romanes.....	59
1.5.4.4 Analogies et dissemblances morphosyntaxiques des déterminants.....	63
1.5.4.5 Analogies et dissemblances morphosyntaxiques des pronoms.....	63
1.5.5 Récapitulatif partiel des analogies entre l'anglais et les langues romanes.....	64
1.5.6 Notre définition de l'analogie lexicale et de l'interlexique anglais-espagnol-français-italien-portugais.....	66
1.6 Conclusion du chapitre 1.....	68
Chapitre 2 - Raisons historiques des analogies entre l'anglais et les langues romanes.....	70
2.1 Introduction du chapitre 2.....	70
2.2 Origines génétiques communes des langues germaniques et romanes : des langues indoeuropéennes.....	72
2.2.1 Utilisation actuelle des langues indoeuropéennes.....	72
2.2.2 De la reconstruction phonéticohistorique des langues indoeuropéennes : à la recherche d'analogies.....	73
2.2.3 Hypothèses sur la diffusion des langues indoeuropéennes.....	74
2.2.3.1 D'une langue supposée à un peuple supposé.....	74
2.2.3.2 Théories sur le foyer d'origine d'hypothétiques Protoindoeuropéens.....	75
2.2.3.3 L'apport de la génétique à la recherche sur les migrations en Eurasie.....	76
2.2.4 Mise en question des modèles généalogiques.....	77
2.2.4.1 Implications des représentations arborescentes.....	77
2.2.4.2 Classification des langues romanes et germaniques.....	79
2.3 Des analogies interlinguistiques liées aux contacts des langues en Europe de l'Ouest.....	81
2.3.1 Concepts et outils autour de la notion de contacts de langues.....	81
2.3.1.1 Substrat, superstrat, adstrat.....	81
2.3.1.2 Le concept de continuum.....	83
2.3.1.3 Conflit entre la possibilité d'une hybridité et le sentiment linguistique...	84
2.3.1.4 L'emprunt, effet du contact de langue.....	85
2.3.2 Le latin comme dénominateur commun.....	86
2.3.2.1 Origines latines du lexique de l'anglais et des langues romanes.....	86

2.3.2.2	Les langues celtiques de l'Antiquité face au latin.....	87
2.3.2.3	Influences du latin dans les langues des Germains.....	89
2.3.2.4	Le latin, langue liturgique véhiculaire et langue de l'administration.....	91
2.3.2.5	Le latin, langue de la connaissance.....	92
2.3.2.6	Emprunts lexicaux au latin et au grec en biologie et en médecine.....	94
2.3.3	Influences des langues romanes.....	95
2.3.3.1	Influences normandes et françaises sur l'anglais.....	95
2.3.3.2	L'influence européenne du français du XVI <sup>e</sup> au XIX <sup>e</sup> siècle.....	96
2.3.4	Influences des langues germaniques.....	97
2.3.4.1	L'influence de la langue des Vikings.....	97
2.3.4.2	L'influence du bas-allemand sur le roman et l'anglais.....	99
2.3.4.3	Mondialisation des emprunts à l'anglais.....	100
2.4	Conclusion du chapitre 2.....	102
Chapitre 3 -	Valorisation des analogies interlinguistiques par l'intercompréhension et les approches sur corpus.....	104
3.1	Introduction du chapitre 3.....	104
3.2	L'analogie en didactique des langues : risques et promesses.....	106
3.2.1	Importance de l'analogie dans l'appropriation du savoir.....	106
3.2.2	Du risque d'interférences.....	106
3.2.3	Des influences des langues connues dans l'apprentissage d'une L3.....	108
3.2.3.1	Influences plus marquées de la L2.....	108
3.2.3.2	L'intérêt d'une L2 passerelle pour apprendre une L3.....	109
3.2.3.3	L'anglais, une langue passerelle vers les langues romanes.....	110
3.2.3.4	Psychotypologie et apprentissage d'une langue.....	111
3.2.3.5	Intelligibilité inhérente et intelligibilité acquise.....	111
3.3	Des approches novatrices qui développent l'intercompréhension.....	113
3.3.1	Ancrage et définition.....	113
3.3.2	But et principes de l'intercompréhension.....	113
3.3.3	Stratégies et techniques.....	114
3.3.4	Les langues romanes, des langues 'pluricentriques' ?.....	115
3.4	Apport des approches sur corpus pour la connaissance des analogies linguistiques.....	118
3.4.1	Définitions d'une discipline et d'un outil.....	118
3.4.2	Distinction entre corpus comparables et corpus parallèles.....	118
3.4.3	Vers une compréhension plus fine du discours de spécialité.....	120
3.4.3.1	Le lexique témoin d'une communauté de discours.....	120
3.4.3.2	Classer le lexique des écrits scientifiques.....	120
3.4.3.3	Une phraséologie des séquences collocatives analogues.....	122
3.5	Présentation du corpus ScienText utilisé pour l'étude lexicologique.....	124
3.5.1	Origine et principes de constitution de ScienText.....	124
3.5.2	Comment requêter à partir de ScienQuest.....	124
3.6	Conclusion du chapitre 3.....	126
Chapitre 4 -	Le lexique anglais dans le discours scientifique de la santé : délimitation des 2000 adjectifs, noms et verbes fréquents du corpus ScienText et de leurs 2208 acceptations fréquentes dans le corpus.....	130
4.1	Introduction du chapitre 4.....	130
4.2	Considération préalable : pourquoi ne pas avoir compilé des lexiques existants...132	
4.2.1	Défaut de critères de sélection des mots-vedettes dans les dictionnaires.....	132

4.2.2	Une évolution notoire : des dictionnaires bilingues qui définissent dans la langue source.....	132
4.2.3	Notre approche : le lexique et les acceptions fréquents.....	133
4.3	Constitution d'une liste des adjectifs, noms et verbes anglais fréquents du corpus ScienText.....	135
4.3.1	Plusieurs extractions nécessaires.....	135
4.3.1.1	Adaptation à un corpus évolutif.....	135
4.3.1.2	Première vérification de l'étiquetage des mots à 1 000 occurrences.....	135
4.3.2	Écrémage des mots.....	136
4.3.2.1	Écrémage des adjectifs.....	136
4.3.2.2	Écrémage des noms.....	137
4.3.2.3	Écrémage des verbes.....	137
4.3.2.4	Écrémage de mots ambigus ou redondants.....	138
4.3.3	Des erreurs d'étiquetage liées à des particularités linguistiques de l'anglais. .	138
4.3.3.1	Erreurs liées au faible nombre de flexions en anglais.....	138
4.3.3.2	Erreurs d'étiquetage dues aux homographes en anglais.....	139
4.3.3.3	Homogénéisation de la graphie.....	141
4.3.4	Sélection des 2000 mots les plus fréquents.....	141
4.4	Détermination du sens des 2000 mots anglais fréquents.....	143
4.4.1	Évolutions de l'approche suivie pour cerner le sens.....	143
4.4.1.1	Un premier examen manuel des 100 premières occurrences.....	143
4.4.1.2	Insuffisances de ce procédé.....	144
4.4.1.3	Seconde tentative : déterminer le sens par le profil combinatoire.....	144
4.4.1.4	Détermination de la monosémie / polysémie d'un mot du corpus ScienText .....	147
4.4.2	Comment définir les 2208 acceptions identifiées.....	148
4.4.2.1	Le choix du Wiktionary.....	148
4.4.2.2	Acceptions présentes / absentes du Wiktionary.....	149
4.4.2.3	Ajout de définitions dans le Wiktionary.....	150
4.5	Deux catégories sémantiques pour les 2208 acceptions fréquentes.....	152
4.5.1	Un tri manuel en deux catégories : lexique de spécialité et LST.....	152
4.5.2	Sous-catégories du lexique de spécialité de la santé.....	153
4.5.2.1	Anatomie générale.....	153
4.5.2.2	Anatomie comparée.....	154
4.5.2.3	Pathologie humaine : anatomopathologie générale.....	154
4.5.2.4	Pathologie humaine : biologie médicale.....	155
4.5.2.5	Thérapeutique.....	155
4.5.3	Sous-catégories du LST.....	156
4.5.3.1	La connaissance.....	156
4.5.3.2	Les dynamiques sociales.....	156
4.5.3.3	Les orientations de la recherche.....	156
4.5.3.4	La description des caractéristiques.....	157
4.5.3.5	La démarche scientifique.....	157
4.5.3.6	La mesure.....	157
4.5.3.7	L'interprétation.....	157
4.6	Conclusion du chapitre 4.....	158
Chapitre 5 - Traduction des 2000 mots anglais fréquents du corpus ScienText en espagnol, français, italien et portugais et mesures de similarité.....		160
5.1	Introduction du chapitre 5.....	160

5.2 Traduction des 2000 mots anglais fréquents en espagnol, français, italien et portugais.....	162
5.2.1 Une approche prioritairement cibliste.....	162
5.2.2 Trois cas rencontrés lors de la traduction.....	162
5.2.2.1 Équivalents de traduction connus.....	162
5.2.2.2 Latinité et grécité de mots anglais de spécialité : approche hypothético-déductive empruntée à l'intercompréhension.....	165
5.2.2.3 Équivalents de traduction manquants : recherches croisées.....	169
5.2.2.4 Termes non traduits dans le Wiktionary : requêtes dans Wikipédia et dans un moteur de recherche.....	169
5.2.3 Exemple du processus de traduction : le nom <i>assay</i> .....	170
5.2.3.1 Les équivalents de traduction du Wiktionary.....	170
5.2.3.2 Traduction d' <i>assay</i> en français.....	171
5.2.3.3 Traduction d' <i>assay</i> en italien.....	172
5.2.3.4 Traduction d' <i>assay</i> en portugais.....	173
5.2.4 Conséquences des différences lexicosémantiques interlinguistiques.....	173
5.2.4.1 De la difficulté de traduire les mots anglais courants.....	173
5.2.4.2 Mots anglais dont l'acception est plus étendue.....	174
5.2.4.3 Mots anglais dont l'acception est moins étendue que les langues romanes.....	176
5.2.5 Conséquences des divergences lexicosyntaxiques.....	176
5.2.5.1 Traduction des particularités lexicosyntaxiques des adjectifs anglais....	176
5.2.5.2 Multiplicité des constructions syntaxiques pour un même mot.....	177
5.2.5.3 Mots n'entrant que dans des figements.....	177
5.2.5.4 Autres cas de divergences syntaxiques.....	178
5.2.6 Convergence sémantique interlexicale.....	179
5.2.6.1 Le même rapport sémantique : l'exemple d' <i>early</i> .....	179
5.2.6.2 Convergence des séquences collocatives spécifiques.....	180
5.2.7 Conséquences de la variation intralinguistique.....	182
5.2.7.1 Variation diatopique des traductions des mots courants de la santé.....	182
5.2.7.2 Variation diastratique des traductions des mots courants de la santé.....	183
5.2.7.3 Variation diachronique et processus de néologisation en cours.....	184
5.2.7.4 Variation stylistique de l'anglais.....	185
5.2.8 Forme et ordre des mots traduits dans les tableaux de données.....	186
5.2.8.1 Forme des verbes.....	186
5.2.8.2 Ordre des traductions multiples dans les tableaux.....	186
5.3 Mesures de similarité : vers un interlexique et quatre lexiques bilingues.....	187
5.3.1 Un interlexique au niveau des acceptions et non des mots.....	187
5.3.2 Neutralisation des diacritiques des langues romanes.....	188
5.3.3 Mesure de similarité par le script CompareCols.csv.pl.....	189
5.3.3.1 Les trois critères appliqués pour mesurer la similarité.....	189
5.3.3.2 Réglages successifs des critères.....	189
5.3.3.3 Neutralisation des finales équivalentes.....	190
5.3.4 Étapes suivies pour mesurer la similarité formelle.....	191
5.3.5 Mesure de similarité des mots de moins de 4 lettres.....	193
5.3.5.1 Obtention des termes de moins de 4 lettres dans chaque langue.....	193
5.3.5.2 Approche formelle.....	193
5.3.5.3 Approche complémentaire plus intuitive.....	194
5.3.6 Similarité : trois catégories de mots traduits.....	195

5.3.7	Étapes pour obtenir quatre lexiques bilingues anglais-langue romane des mots analogues.....	196
5.4	Conclusion du chapitre 5.....	198
Chapitre 6 - Analogies interlinguistiques des mots anglais fréquents de la santé du corpus ScienText et de leurs traductions en espagnol, français, italien et portugais.....		
6.1	Introduction du chapitre 6.....	201
6.2	Étude des caractéristiques du lexique anglais fréquent de la santé du corpus scientifique de ScienText.....	202
6.2.1	Représentativité des occurrences des 2 000 mots anglais les plus fréquents dans le corpus ScienText.....	202
6.2.2	Catégories grammaticales des 2000 mots anglais fréquents.....	203
6.2.2.1	Résultats globaux : les catégories grammaticales des 2000 mots anglais fréquents.....	203
6.2.2.2	Part de chaque catégorie grammaticale parmi les mots anglais fréquents.....	204
6.2.2.3	Part de chaque catégorie grammaticale : comparaison de ScienText avec l'Oxford English Dictionary.....	205
6.2.2.4	Catégories grammaticales en fonction du rang des 2000 mots.....	205
6.2.2.5	Catégories grammaticales des occurrences en fonction du rang des 2000 lemmes.....	206
6.2.2.6	Les mots les plus fréquents : catégories grammaticales des 2000 mots en fonction des quartiles d'occurrences.....	207
6.2.3	Monosémie / polysémie des 2000 mots anglais les plus fréquents de ScienText.....	209
6.2.3.1	Données globales sur les monosémiques / polysémiques.....	209
6.2.3.2	Répartition globale des monosémiques / polysémiques.....	210
6.2.3.3	Répartition monosémiques / polysémiques par catégories grammaticales.....	210
6.2.3.4	Répartition des occurrences des mots monosémiques / polysémiques par catégories grammaticales.....	211
6.2.4	Caractérisation sémantique du lexique anglais de ScienText.....	212
6.2.4.1	Récapitulatif des données sur les catégories sémantiques.....	212
6.2.4.2	Représentation globale des catégories sémantiques des 2208 acceptions.....	213
6.2.4.3	Catégories grammaticales selon la catégorie sémantique des acceptions fréquentes.....	213
6.2.4.4	Catégorie sémantique et rang par acceptions.....	214
6.3	Convergence des monosémiques et des polysémiques à une traduction.....	220
6.3.1	Traduction des mots polysémiques.....	220
6.3.1.1	Des mots polysémiques à une traduction.....	220
6.3.1.2	Mots polysémiques selon le nombre de traductions.....	221
6.3.2	Analogie des mots monosémiques et polysémiques à une traduction.....	222
6.3.2.1	Un premier niveau : monosémiques / polysémiques.....	222
6.3.2.2	Analogie des monosémiques / polysémiques dans les cinq langues.....	222
6.3.2.3	Analogie des 1944 mots anglais à une traduction.....	223
6.3.2.4	Catégories grammaticales des 1944 mots à une traduction.....	224
6.3.2.5	Analogie des 1944 mots à une traduction.....	225
6.4	Un interlexique anglais-espagnol-français-italien-portugais des mots de la santé.....	227
6.4.1	Analogie selon l'acception.....	227

6.4.1.1	Résultats globaux : pour 1038 acceptions sur 2208, les équivalents sémantiques sont analogues dans les 5 langues.....	227
6.4.1.2	Catégories grammaticales des mots analogues dans les 5 langues exprimant les 1038 acceptions.....	227
6.4.2	Analogie dans les cinq langues (lexique spécifique de la santé).....	229
6.4.3	Analogie dans les cinq langues (LST).....	231
6.4.4	Comparaison de l’analogie dans les deux catégories sémantiques.....	232
6.4.5	Analogie, rang et catégorie sémantique.....	233
6.5	Analogie par couples de langues : quatre lexiques bilingues anglais-langue romane.....	238
6.5.1	Résultats globaux.....	238
6.5.2	Analogie anglais – langue romane selon la catégorie grammaticale.....	239
6.5.3	Analogie anglais-langue romane selon la catégorie sémantique.....	241
6.5.4	Améliorer les mesures de similarité par la prise en compte des allographèmes.....	242
6.6	Conclusion du chapitre 6.....	245
Chapitre 7 - Perspectives : vers un dispositif didactique plurilingue d'enseignement du FOS sur la base d'analogies interlinguistiques.....		247
7.1	Introduction du chapitre 7.....	247
7.2	Articulation de l’étude lexicologique et d’activités de FLE construites sur des analogies avec l’anglais.....	249
7.2.1	Des analogies lexicales anglais – langue romane pour l'apprentissage d'une L3 romane sur objectifs spécifiques.....	249
7.2.2	Configuration : L1 langue distante, L2 anglais L3 français.....	250
7.2.3	Quels mots de l’interlexique et des lexiques bilingues incorporer ?.....	250
7.2.3.1	LST vs. lexique de spécialité de la santé.....	250
7.2.3.2	Incorporer le lexique polysémique de la langue passerelle ?.....	251
7.3	Créer des conditions favorables à la perception d’analogies interlinguistiques.....	252
7.3.1	Lire une langue sans l’avoir apprise.....	252
7.3.2	Application de principes de la didactique des langues tierces.....	253
7.3.3	S’appuyer sur le lexique analogue.....	253
7.3.3.1	Rendre transparent le lexique analogue.....	253
7.3.3.2	Traitement de l’opacité lexicale : plusieurs langues en parallèle.....	254
7.3.4	Valorisation d’analogies et de dissemblances syntaxiques et morphosyntaxiques.....	255
7.3.4.1	Un questionnement métalinguistique contrastif.....	256
7.4	Requêtes dans des corpus multialignés et exemples d’activités.....	257
7.4.1	Des textes authentiques contenant des mots de l’interlexique.....	257
7.4.2	Présentation des corpus multialignés.....	257
7.4.2.1	Le corpus multiparallèle EMEA.....	257
7.4.2.2	Le corpus multiparallèle Europarl.....	258
7.4.3	Extractions de chaînes de caractères.....	259
7.4.4	Résultats partiels d’extractions dans les deux corpus multialignés.....	260
7.4.4.1	Écart entre la langue, le discours et la didactisation.....	260
7.4.4.2	Exemples de concordances didactisables du corpus EMEA.....	260
7.4.4.3	Exemples de concordances didactisables du corpus Europarl.....	263
7.4.4.4	Exemple de didactisation de la structure attributive.....	264
7.4.4.5	Autre piste prometteuse : des requêtes en deux langues.....	265
7.5	Conclusion du chapitre 7.....	267

Conclusion générale.....	268
Références bibliographiques.....	271
Annexes.....	291
Annexe 1 : les 2000 lemmes anglais fréquents du corpus ScienText.....	291
Annexe 2 : Les 2000 lemmes anglais du corpus ScienText traduits.....	298
Annexe 3 : Définitions à valider des experts.....	388
Annexe 4 : Tableau des allosuffixes.....	394
Annexe 5 : Liste du lexique spécifique de la santé.....	398
Annexe 6 : liste du lexique scientifique transdisciplinaire.....	403
Annexe 7 : l'interlexique : 1038 séries de mots analogues en anglais, espagnol, français, italien et portugais.....	408

# Index des images

Image 1 : Photographie d'un panneau d'affichage sur le quai de la gare Flughafen Graz-Feldkirchen.....	58
Image 2 : Répartition des langues indoeuropéennes dans le monde.....	72
Image 3 : Représentation des langues indoeuropéennes de Watkins.....	78
Image 4: Photographie d'une tablette de défixion, Bath (Angleterre).....	88
Image 5 : Photographie de la pierre runique d'Uppland n°194, Väsby (Suède).....	98
Image 6 : Capture d'écran d'un exemple de phrases alignées sur Reverso context.....	119
Image 7 : Extrait de la feuille contenant les 374 verbes.....	142
Image 8 : Copie d'écran du Lexicoscope.....	145
Image 9 : Copie d'écran d'un extrait de lexicogramme dans le Lexicoscope.....	146
Image 10 : Copie d'écran du cadre contenant les traductions de assay, Wiktionary.....	170
Image 11 : Trois exemples Sous-chaine commune maximale (SCM).....	189
Image 12 : Copie de la requête Nom être Adjectif.....	259



# Index des graphiques

Graphique 1 : Part des occurrences des 2000 mots anglais les plus fréquents dans le corpus ScienText.....	203
Graphique 2 : Part des 2000 mots fréquents par catégories grammaticales.....	205
Graphique 3 : Part des occurrences des 2000 mots fréquents par catégories grammaticales.....	205
Graphique 4 : Catégorie grammaticale des mots selon le rang des 2000 mots.....	207
Graphique 5 : Catégorie grammaticale des occurrences selon le rang des 2000 mots.....	208
Graphique 6 : Catégories grammaticales par quartiles d'occurrences.....	209
Graphique 7 : Répartition globale des monosémiques / polysémiques.....	211
Graphique 8 : Répartition globale des occurrences monosémiques / polysémiques.....	211
Graphique 9 : Répartition des mots monosémiques / polysémiques par catégories grammaticales.....	211
Graphique 10 : Occurrences des monosémiques / polysémiques par catégories grammaticales.....	212
Graphique 11 : Nombre d'occurrences de mots de chaque catégorie sémantique.....	214
Graphique 12 : Nombre d'acceptions de chaque catégorie sémantique.....	214
Graphique 13 : Catégorie grammaticale d'un mot selon la catégorie sémantique.....	214
Graphique 14 : Catégorie grammaticale de l'occurrence selon la catégorie sémantique.....	215
Graphique 15 : Distribution des acceptions du lexique de spécialité / du LST.....	217
Graphique 16 : Catégories grammaticales des acceptions du lexique de spécialité par quartiles.....	217
Graphique 17 : Catégorie grammaticale des acceptions du LST par quartiles.....	218
Graphique 18 : Distribution des occurrences de mots anglais du corpus ScienText en fonction de la catégorie grammaticale et de la catégorie sémantique.....	219
Graphique 19 : Mots anglais fréquents selon le nombre de traductions dans les cinq langues romanes.....	225
Graphique 20 : Occurrences des mots anglais fréquents selon le nombre de traductions dans les cinq langues romanes.....	225
Graphique 21 : Analogie dans les 5 langues des mots exprimant les 2208 acceptions.....	228
Graphique 22 : Analogie dans les 5 langues des occurrences des mots exprimant les 2208 acceptions.....	228
Graphique 23 : Catégories grammaticales des mots pour les 1038 acceptions où ils sont analogues dans les 5 langues.....	229
Graphique 24 : Catégories grammaticales des mots exprimant les 2208 acceptions.....	229
Graphique 25 : Catégorie grammaticale des occurrences des mots pour les 1038 acceptions où ils sont analogues dans les 5 langues.....	229
Graphique 26 : Catégories grammaticales des occurrences des mots exprimant les 2208 acceptions.....	229
Graphique 27 : Occurrences des adjectifs analogues / non analogues.....	230
Graphique 28 : Occurrences des verbes analogues / non analogues.....	230
Graphique 29 : Occurrences des noms analogues / non analogues.....	230
Graphique 30 : Nombre de langues avec lesquelles l'anglais est analogue.....	239

# Index des tableaux

Tableau 1 : Formation périphrastique du passif dans les langues romanes et germaniques.....	63
Tableau 2 : Principales analogies / dissemblances de l'anglais, du français et des langues romanes.....	66
Tableau 3 : Extrait de la liste des adjectifs en cours de nettoyage.....	136
Tableau 4 : Les 2000 mots fréquents de ScienText par catégories grammaticales.....	203
Tableau 5 : Les entrées de l'OED par catégories grammaticales.....	205
Tableau 6 : Les 53 mots du premier quartile d'occurrences.....	208
Tableau 7 : Monosémie / polysémie des 2000 mots fréquents de ScienText.....	209
Tableau 8 : Tableau récapitulatif des catégories sémantiques.....	212
Tableau 9 : Récapitulatif catégorie sémantique et rang.....	215
Tableau 10 : Répartition des deux catégories de polysémiques.....	221
Tableau 11 : Analogie des monosémiques / polysémiques.....	222
Tableau 12 : Analogie des polysémiques et des monosémiques selon la catégorie grammaticale.....	222
Tableau 13 : Mots à une ou plus d'une traduction.....	223
Tableau 14 : Répartition des monosémiques et des polysémiques à une traduction selon la catégorie grammaticale.....	225
Tableau 15 : Analogie des 2000 mots selon la catégorie grammaticale.....	226
Tableau 16 : Analogie en cinq langues du lexique anglais de spécialité de la santé.....	229
Tableau 17 : Analogie en cinq langues des occurrences du lexique anglais de spécialité de la santé.....	230
Tableau 18 : Analogie en cinq langues du lexique anglais transdisciplinaire.....	231
Tableau 19 : Analogie en cinq langues des occurrences du lexique anglais transdisciplinaire.....	231
Tableau 20 : Analogie par catégories grammaticales selon la catégorie sémantique.....	232
Tableau 21 : Séries analogues par quartiles.....	233
Tableau 22 : Analogie du lexique de spécialité et du LST selon le rang de l'acceptation.....	234
Tableau 23 : Récapitulatif catégorie sémantique et rang.....	235
Tableau 24 : Distribution des mots analogues et non analogues du lexique de spécialité de la santé et du LST en fonction des quartiles d'acceptations.....	236
Tableau 25 : Analogie du lexique des deux catégories selon le rang des acceptations.....	237
Tableau 26 : Nombre d'acceptations pour lesquelles le mot anglais est analogue avec 1, 2, 3 ou 4 langues romanes.....	238
Tableau 27 : Récapitulatif de l'analogie de l'anglais avec chacune des langues romanes.....	239
Tableau 28 : Analogie par catégories grammaticales et par langues.....	239
Tableau 29 : Répartition de chaque catégorie grammaticale dans les quatre lexiques bilingues.....	240
Tableau 30 : Acceptations pour lesquelles les paires de langues sont analogues, par catégories sémantiques.....	241
Tableau 31 : Allographèmes anglais - langues romanes.....	243

# Introduction

L'être humain perçoit des ressemblances entre certaines langues. Des écrits anciens de voyageurs témoignent déjà avant notre ère de cette faculté, ainsi que d'une variation linguistique qui n'empêche pas la compréhension mutuelle. Ainsi, dans sa *Géographie* (23, XV, chap. II, 8), Strabon rapporte les observations faites en Asie Centrale par Ératostène au III<sup>e</sup> siècle avant notre ère : selon ce dernier, les habitants de l'Ariané<sup>1</sup>, « parlent à peu de choses près la même langue »<sup>2</sup>. Lelièvre (2004 : 52), s'appuyant sur le chroniqueur Sima Qian (*Shiji*, 123), relate l'expérience de Zhang Qian, un Chinois qui a voyagé au II<sup>e</sup> siècle dans la même région : « quant aux caractéristiques de ces étrangers, on disait qu'ils possédaient, à l'ouest du Dayuan jusqu'à l'Anxi, des langages différents mais que leurs coutumes étaient voisines et qu'ils se comprenaient aisément »<sup>3</sup>.

La comparaison interlinguistique est l'une des stratégies déployées par la didactique des langues tierces (L3) en vue de développer le plurilinguisme. Selon Hufeisen et Neuner (2003 : 5), « on appelle langues tierces les langues étrangères qui sont apprises successivement après une première langue étrangère, autrement dit la deuxième langue étrangère, la troisième, la quatrième, etc. »<sup>4</sup>. Nous souscrivons à l'idée que l'enseignement d'une L3 devrait prendre en compte le répertoire langagier de l'apprenant, autrement dit les langues dans lesquelles il dispose déjà de compétences : sa langue maternelle (L1), sa première langue étrangère (L2), et ses éventuelles autres L3.

Nous nous inspirons des approches plurielles des langues et des cultures (Candelier, 2008), qui visent l'appropriation de compétences partielles. Parmi ces approches, celles qui se réclament de l'intercompréhension (Dabène, 1994 ; Klein & Rutke, 1997 ; Blanche-

---

1 C'est-à-dire d'une partie des antiques Perse, Médie, Bactriane septentrionale et Sogdiane.

2 « εἰσὶ γὰρ πῶς καὶ ὁμόγλωττοι παρὰ μικρὸν », traduction française d'Amédée Tardieu, <http://remacle.org/bloodwolf/erudits/strabon/livre152.htm>.

3 Lebedynsky (2006 : 10) mentionne ces expériences d'Ératostène et de Zhang Qian. *Dayuan* désignait en chinois la vallée de la Ferghana, dans les actuels Ouzbékistan, Kirghizistan et Tadjikistan. *L'Anxi* est l'ancien nom d'un district de la province du Gansu, dans l'ouest de la Chine.

4 « Als Tertiärsprachen bezeichnet man diejenigen Fremdsprachen, die in der zeitlichen Abfolge nach einer ersten Fremdsprache, d.h. als 2., 3., 4., etc. erlernt werden », notre traduction. On les appelle également en français les *langues tertiaires*.

Benveniste, 1997 ; Degache, 1997 ; Castagne, 2007A, etc.) s'appuient très largement sur les analogies interlinguistiques.

Les recherches dont nous rendons compte dans cette thèse portent sur des langues appartenant à deux branches différentes de la famille des langues indoeuropéennes : les langues germaniques et les langues romanes. L'orientation est didactique, mais nous recourons à différents champs de la recherche, dont la lexicologie et le Traitement Automatique des Langues (TAL).

Le public ciblé a un projet de mobilité dans un pays francophone et a besoin d'apprendre le français de la santé dans la perspective didactique que l'on appelle Français sur Objectifs Spécifiques (FOS). Il a comme L1 une langue éloignée des langues romanes et germaniques - par exemple l'une des langues de Chine, le coréen, le japonais, le vietnamien, ou encore l'arabe, et il a appris comme langue étrangère l'anglais. Ce public est composé d'étudiants et de professionnels du secteur de la santé au sens large, qui sont amenés à exercer ou exercent déjà en médecine, pharmacie, biologie cellulaire, biochimie, etc. Nous supposons qu'il a au moins le niveau B2 du CECRL<sup>5</sup> à l'écrit en anglais, et qu'il connaît les mots anglais usuels de son domaine de spécialité.

Un courant de la didactique des langues tierces étudie comment une L3 peut être enseignée en prenant en compte les analogies que présente une L2 avec cette L3, et les phénomènes de transfert qui en résultent. La L2 est alors envisagée comme « langue passerelle », ou « langue pont », vers la L3.

L'anglais paraît<sup>6</sup> particulièrement approprié comme passerelle vers le français de la santé. Tout d'abord, c'est la langue étrangère la plus apprise dans le monde, si bien que les apprenants d'une L3 l'ont déjà très souvent apprise en L2. Ensuite, l'anglais est classé parmi les langues germaniques, mais le latin et le français ont exercé pendant des siècles de profondes influences sur son lexique. Les langues romanes proviennent par définition du latin, leur lexique offre pour cette raison des similitudes avec l'anglais, par exemple dans les termes de la santé, où l'on recourt massivement aux racines grecques et latines. Enfin, l'anglais est *de facto* la langue de la recherche dans le domaine de la santé (Wulff, 2004 : 188 ; Faure, 2010 :

---

5 <https://frama.link/q6ga5XnD>

6 Nous utilisons l'orthographe du français rectifiée selon la loi du 6 décembre 1990, [https://frama.link/uvh\\_ZPxA](https://frama.link/uvh_ZPxA). Une liste des mots concernés est accessible à la page <http://www.renouvo.org/liste.php?t=3>.

73-86), aussi de nombreux termes de cette langue sont-ils empruntés par des locuteurs d'autres langues.

Cette thèse contribue à la connaissance des analogies lexicales entre l'anglais et les langues romanes dans le domaine de la santé, ainsi qu'à la réflexion sur l'élaboration d'activités écrites de FOS à partir d'analogies interlinguistiques.

Nous supposons qu'un enseignement plurilingue fondé sur des analogies lexicales et syntaxiques qui existent entre l'anglais et les langues romanes peut permettre à un apprenant de comprendre directement des écrits authentiques simples en français dans un domaine de spécialité : l'écrit, tantôt en français seulement, tantôt en plusieurs langues romanes, fait l'objet d'une comparaison interlinguistique guidée par un questionnement métalinguistique en anglais.

Nous avons étudié dans notre mémoire de master (Gilles, 2011) l'évolution de la compréhension des mots lexicaux et grammaticaux dans des textes de langues romanes par des élèves de 11 ans, et nous avons constaté que le sens des mots récurrents, qu'ils soient analogues ou non au français, a été mieux mémorisé. Nous présupposons que les mots fréquents sont compris des lecteurs. C'est pourquoi notre étude prospective s'appuie sur le lexique anglais fréquent dans le discours scientifique de la santé. Par *discours scientifique*, Boch et Rinck (2010) entendent le « discours produit dans le cadre de l'activité de recherche à des fins de construction et de diffusion du savoir ». Ce lexique a été identifié au moyen d'outils de Traitement Automatique des Langues (TAL). Les 2000 mots anglais les plus fréquents du sous-corpus scientifique de la santé de ScienText ont été obtenus, puis leurs 2208 acceptions fréquentes déterminées en fonction de leur profil combinatoire (Kraif & Diwersy, 2012).

Le français est la langue cible de notre travail. Néanmoins, nous pensons que les autres langues romanes peuvent aider à la compréhension du français dans certains cas d'opacité lexicale. C'est pourquoi le lexique anglais fréquent de la santé a été traduit en français, espagnol, italien et portugais, les quatre langues romanes les plus parlées dans le monde. La similarité formelle des mots anglais, espagnols, français, italiens et portugais a été mesurée pour chacune des 2208 acceptions fréquentes du corpus. Il en a résulté un interlexique anglais-français-espagnol-italien-portugais, autrement dit un lexique en cinq langues composé de mots anglais fréquents de la santé et de leurs équivalents de traduction qui pour une acception donnée sont similaires dans les quatre langues romanes.

Enfin, nous avons interrogé les corpus EMEA et Europarl où nous avons cherché des exemples de séquences collocatives composées d'éléments de cet interlexique destinées à un dispositif didactique plurilingue.

Des enjeux de plusieurs natures sont associés à cette étude de la terminologie multilingue de la santé. D'abord, des enjeux scientifiques. Nous attendons de cette recherche des retombées pour la lexicologie synchronique :

- l'étude des 2000 mots anglais fréquent du corpus ScienText permettra de mieux connaître le lexique dans le genre des écrits scientifiques de la santé, à commencer par les particularités du lexique de spécialité de la santé (abrégé parfois en lexique de la santé) et celles du lexique scientifique transdisciplinaire (LST) ;
- la comparaison du lexique fréquent de la santé avec les données des dictionnaires de la santé pourrait éclairer certains choix éditoriaux (représentativité des mots-vedettes, acceptions traitées, exemples) de ces dictionnaires de spécialité ;
- l'interlexique anglais-espagnol-français-italien-portugais permettra de mieux connaître les analogies interlinguistiques dans le lexique de spécialité de la santé / le lexique scientifique transdisciplinaire.

Ce troisième point intéresse la lexicologie, mais aussi la didactique :

- les analogies interlinguistiques pourraient orienter la réflexion sur l'enseignement des langues romanes sur la base d'analogies avec l'anglais dans des dispositifs en présentiel, à distance (adaptés par exemple à des usages nomades, sur tablette et téléphone portable), ou hybrides ;
- les données ont été élaborées à partir de 35,2 millions de mots extraits de textes authentiques de la spécialité ; il semble légitime qu'un enseignement de l'anglais de spécialité s'appuie sur les plus fréquents de ces 2000 mots, en fonction de la sous-spécialité de l'apprenant (biochimie, thérapeutique, etc.) ;
- les apprenants de français L3 disposent déjà de connaissances en anglais et d'une expérience d'apprenants de langue étrangère ; dispensée en amont, une préparation linguistique adaptée des professionnels étrangers conditionne en partie le succès de leur projet de mobilité.

À ces enjeux s'en ajoutent qui sont d'ordre socioculturel ou sociétal :

- les dictionnaires généralistes, qui définissent certains mots du lexique scientifique, contribuent à la vulgarisation de la recherche ; les termes de ce domaine essentiel qu'est la santé devraient être compréhensibles par tous les citoyens ;
- malgré l'omniprésence de l'anglais, la recherche médicale continue en français et dans les autres langues romanes (catalan, etc.) ; apprendre à comprendre ces langues, d'abord à l'écrit, puis à l'oral, prépare le séjour d'étude pendant lequel des compétences en anglais serviront peu avec les patients et les collègues.

Le lexique anglais a été fortement influencé par le latin et le français. Nous avons étudié les analogies entre l'anglais et les langues romanes dans le domaine de la santé, dans les deux catégories sémantiques que sont le lexique de spécialité de la santé et le LST. La question de la polysémie des mots fréquents s'est posée, ainsi que celle de la pertinence des outils de TAL dans le but d'obtenir des contextes afin de bâtir une intercompréhension de spécialité. Nos recherches ont été orientées par les hypothèses lexicologiques suivantes.

Hypothèse 1 :

Dans le discours scientifique de la santé écrit en anglais, le lexique de spécialité fréquent contient une majorité de mots analogues en anglais, espagnol, français, italien et portugais.

Cette hypothèse nous conduit à observer dans ce type de discours s'il existe une convergence entre le lexique de spécialité de l'anglais et celui des quatre langues romanes. Les mesures effectuées n'indiquent pas si l'analogie lexicale est liée à l'appartenance à une même famille de langue ou à des contacts entre les langues, mais elles permettent de savoir s'il semble prometteur de chercher dans des corpus alignés de spécialité des contextes qui contiennent des mots analogues.

Hypothèse 2 :

Les mots anglais polysémiques sont moins analogues à leurs équivalents de traduction dans les langues romanes que les mots monosémiques.

En comparant l'analogie des mots monosémiques à l'analogie des mots polysémiques, nous souhaitons prévenir le risque de difficulté qu'un apprenant éprouverait dans les cas de polysémie : faudrait-il éviter d'incorporer dans des activités d'apprentissage du FLE des mots français qui sont analogues à des mots anglais polysémiques ?

Hypothèse 3 :

Dans le discours scientifique de la santé écrit en anglais, le lexique de spécialité présente une plus grande analogie avec l'espagnol, le français, l'italien et le portugais que le LST.

L'examen de cette dernière hypothèse permettra d'abord de discerner la part du lexique de spécialité de celle du LST. Nous verrons comment nous procéderons pour distinguer les deux catégories sémantiques. La part de chacune montrera quelle est la place du LST dans ce type de discours et les résultats pourront éventuellement orienter les contenus d'enseignement de l'anglais. La comparaison de la similarité des mots des deux catégories sera significative pour l'enseignement des langues romanes : on saura quels sont les mots qu'un apprenant qui a l'anglais en L2 peut apprendre très vite en L3 français (espagnol, etc.) du fait qu'ils sont analogues.

Notre propos est structuré en trois parties. Dans la première partie, nous étudions le concept d'analogie, les origines des analogies lexicales entre l'anglais et les langues romanes et la valorisation des analogies interlinguistiques par des approches didactiques innovantes. La deuxième partie est consacrée à la méthodologie suivie pour extraire des données du corpus ScienText et élaborer un interlexique anglais-espagnol-français-italien-portugais sur la base de l'analogie lexicale entre ces langues. Les résultats sont analysés dans la troisième partie, où nous présentons également les bases d'un dispositif didactique plurilingue d'apprentissage du FLE appliqué au domaine de la santé.

La première partie est constituée de trois chapitres. Le chapitre 1 envisage les rapports du concept d'analogie lexicale avec les concepts proches de similitude et d'identité. Des outils de mesure de similarité sont abordés et notre positionnement discuté. Les principales analogies entre l'anglais et les langues romanes y sont récapitulées, et la définition de l'interlexique



présentée. Le chapitre 2 explique pourquoi il existe des analogies entre l'anglais et le français (et d'autres langues) en donnant des raisons historiques : les origines indoeuropéennes communes sont rappelées et les incessants contacts de langues mis en avant. Le chapitre 3 montre comment les approches plurielles des langues et des cultures, dont l'intercompréhension, s'appuient sur les analogies interlinguistiques pour développer le plurilinguisme et comment les approches sur corpus renouvellent la connaissance du fonctionnement du lexique.

La méthodologie est exposée dans la deuxième partie, qui se compose de deux chapitres. Le chapitre 4 expose la méthodologie d'extraction et de traitement du lexique anglais fréquent. La méthodologie d'élaboration de l'interlexique des mots fréquents de la santé fait l'objet du chapitre 5 : ce chapitre décrit comment ont été traduits en espagnol, français, italien et portugais les mots anglais exprimant les 2208 acceptions fréquentes identifiées dans le corpus, et comment a été mesurée la similarité des 8832 équivalents de traductions avec les 2000 mots anglais.

La troisième partie est consacrée aux résultats et aux perspectives. Elle donne lieu à deux chapitres. Dans le chapitre 6, nous étudions les résultats de l'étude lexicologique : nous analysons les caractéristiques syntaxiques et sémantiques du lexique anglais fréquent présent dans le corpus ScienText, l'interlexique et les lexiques bilingues anglais-langue romane. Le chapitre 7 ouvre des perspectives didactiques : nous y entamons une réflexion sur l'élaboration d'activités plurilingues sur la base de l'interlexique obtenu, d'analogies syntaxiques et de concordances extraites dans des corpus multilingues alignés.

Il nous fallait nous limiter à un nombre de pages raisonnable, aussi les annexes sont-elles abrégées. Leur version complète ainsi que certaines feuilles de calcul sur lesquelles se fondent les résultats seront consultables sur le portail hal<sup>7</sup>.

---

7 <https://hal.archives-ouvertes.fr/hal-01565589v1>

# **Première partie - Analogies linguistiques : concepts voisins, origines et valorisation par des approches didactiques innovantes**

Nous examinons dans cette première partie trois questions qui permettent de comprendre la place des analogies interlinguistiques dans la lexicologie et la didactique du plurilinguisme et de justifier certains de nos choix.

Premièrement, de quelle manière le concept d'analogie s'applique-t-il en linguistique, et particulièrement en lexicologie ? Deuxièmement, pourquoi existe-t-il des analogies entre l'anglais et les langues romanes ? Troisièmement, comment la prise en compte des analogies lexicales interlinguistiques apporte-elle un renouveau dans l'enseignement et la connaissance des langues ?

L'analogie est un principe fondamental dans la cognition. Les phénomènes d'analogies intra- et interlinguistiques sont à l'origine de l'appropriation du fonctionnement du système d'une langue : la compétence de compréhension et celle de production reposent sur l'utilisation d'analogies systémiques, dont les analogies lexicales. Nous verrons ce qu'on entend communément lorsqu'on dit que deux mots se ressemblent. Que ce soit dans leur graphie ou dans leur réalisation orale, à l'intérieur d'une même langue ou entre deux langues, on note des similitudes formelles et/ou sémantiques, mais la prise en compte de la forme et du sens est indispensable pour pouvoir parler d'analogies lexicales interlinguistiques. Ce sera le chapitre 1, Analogies linguistiques : concepts proches et analogies de l'anglais et des langues romanes, à partir de la page 27.

Les analogies entre l'anglais, le français et les autres langues romanes (et germaniques, slaves, etc.) ont été attribuées à l'origine des langues et des peuples qui les parlent. Cette question a suscité et continue à susciter des débats parmi la communauté scientifique. Nous avons jugé bon d'en traiter car l'existence supposée d'une protolangue qui serait à l'origine de centaines de langues ne va pas de soi. Dans le chapitre 2, Raisons historiques des analogies entre l'anglais et les langues romanes, à partir de la page 70, nous examinons deux raisons

avancées pour justifier ces analogies : leur appartenance à une même famille de langues, les langues indoeuropéennes, et les contacts de langues qui ont nourri de multiples influences mutuelles.

Longtemps restreinte à des approches monolingues, la didactique des langues accueille depuis les années 1980 des courants qui réfléchissent à l'enseignement de plusieurs langues. Dans le chapitre 3, Valorisation des analogies interlinguistiques par l'intercompréhension et les approches sur corpus, à partir de la page 104, nous inscrivons nos recherches dans les approches qui se réclament de l'intercompréhension et nous examinons les atouts que présentent les approches sur corpus pour renouveler la lexicologie et la didactique du plurilinguisme.

# Chapitre 1 - Analogies linguistiques : concepts proches et analogies de l'anglais et des langues romanes

## Sommaire du chapitre 1

1.1 Introduction du chapitre 1.....	27
1.2 Distinctions opératoires autour des concepts proches de similarité, d'identité et d'analogie.....	28
1.2.1 Analogie : fécondité d'un concept ambigu.....	28
1.2.2 Essai de clarification des concepts de similarité et d'identité.....	29
1.2.3 L'analogie, un rapport à quatre ou trois termes.....	31
1.3 De la confusion des concepts de similarité, d'identité et d'analogie appliqués au lexique.....	33
1.3.1 Des outils pour la comparaison lexicale intralinguistique.....	33
1.3.2 Mesurer la similitude / l'analogie interlinguistiques.....	35
1.3.3 Un type de similarité lexicale interlinguistique : les <i>faux amis</i> .....	43
1.4 Approche des principaux types d'analogies intralinguistiques et interlinguistiques..	48
1.4.1 Analogies syntaxiques.....	48
1.4.2 Analogies morphosyntaxiques.....	49
1.4.3 Analogies morphologiques.....	50
1.5 Principales analogies et dissemblances entre l'anglais et les langues romanes.....	52
1.5.1 Un vaste champ d'étude où l'histoire de la langue est omniprésente.....	52
1.5.2 Analogies et dissemblances syntaxiques de l'anglais et des langues romanes...	54
1.5.3 Analogies morphologiques de l'anglais et des langues romanes.....	55
1.5.4 Analogies et dissemblances morphosyntaxiques.....	57
1.5.5 Récapitulatif partiel des analogies entre l'anglais et les langues romanes.....	64
1.5.6 Notre définition de l'analogie lexicale et de l'interlexique anglais-espagnol-français-italien-portugais.....	66
1.6 Conclusion du chapitre 1.....	68

## 1.1 Introduction du chapitre 1

Le sens attribué au mot *analogie* est assez étendu pour qu'il nous semble nécessaire de commencer ce travail de linguistique et de didactique en distinguant plusieurs concepts proches. Ce sera la section 1.2, p.28.

Nous verrons dans la section 1.3, p.33 des outils qui décrivent les ressemblances et dissemblances du lexique d'une même langue et de plusieurs langues.

La section 1.4, p.48 dégagera sommairement plusieurs types d'analogies qui dépassent le niveau intra- ou interlinguistique. Nous les utilisons dans la section 1.5,p.52, afin de classer des traits qui unissent ou distinguent l'anglais, le français, et les langues germaniques et romanes en général.

## 1.2 Distinctions opératoires autour des concepts proches de similarité, d'identité et d'analogie

### 1.2.1 Analogie : fécondité d'un concept ambigu

D'après le Trésor de la Langue Française informatisé (TLFi)<sup>8</sup>, le mot français *analogie* est un emprunt au latin *analogia*, « rapport, conformité », provenant lui-même du grec *αναλογία*, *analogia*, qui a signifié d'abord « proportion mathématique », puis « correspondance ». Ce mot a été emprunté par des locuteurs de langues parlées en Europe et classées dans des branches ou familles différentes : *αναλογία* (grec moderne), *analogia* (basque, catalan, finnois, italien, polonais, portugais), *аналогия* (russe), *analogía* (espagnol), *аналогија* (serbe), *analogi* (norvégien, suédois), *Analogie* ou *analogie*, (allemand, néerlandais, tchèque), *analoji* (turc), etc. Les emprunts dans toutes ces langues, et dans d'autres, témoignent de la place essentielle de ce concept.

Voici la définition du mot-vedette *analogie* dans le TLFi :

« Rapport de ressemblance, d'identité partielle entre des réalités différentes préalablement soumises à comparaison ; trait(s) commun(s) aux réalités ainsi comparées, ressemblance bien établie, correspondance ».

Cette définition reflète les usages ambivalents qui sont faits de ce mot. Deux interprétations sont en effet possibles : l'analogie est entendue tantôt comme un rapport de ressemblance, tantôt, par métonymie, comme la ressemblance elle-même, ce qui implique que le terme *analogie* puisse être utilisé en concurrence avec le terme *similitude*.

Pour une analyse approfondie de ce concept, nous renvoyons à l'Habilitation à Diriger des Recherches de Lepage (2003). S'appuyant sur Salamanca (1984), Lepage (ibid. : 46) signale par exemple que le grec *analogia* est traduit par plusieurs mots en latin, parmi lesquels « *aequalitas* (égalité), *comparatio* (comparaison), *similitudo* (ressemblance, similitude) ». L'évolution sémantique de ce concept et sa restitution hésitante a vraisemblablement eu sa part dans la confusion qui demeure aujourd'hui.

Initialement, les Grecs ont fait usage du mot dans les domaines de la mathématique et de la logique. Le préfixe *ana-* signifie par en haut, encore une fois. Comme ce préfixe est associé à

---

8 <http://atilf.atilf.fr/dendien/scripts/tlfiv4/showps.exe?p=combi.htm;java=no>;

*logos*, c'est-à-dire la parole, la raison, on peut comprendre qu'*analogia* réfère à une répétition, à une superposition d'éléments comparés.

Cette idée d'éléments soumis à comparaison est commune à la similarité et l'identité. Il est utile de distinguer ces deux concepts car ils tendent à être confondu dans le langage courant.

## **1.2.2 Essai de clarification des concepts de similarité et d'identité**

### ***1.2.2.1 Part commune et part distincte***

La similarité se distingue de l'identité en ce qu'il existe entre des entités similaires une part distincte, alors que deux entités identiques ne se distinguent pas l'une de l'autre. Le trait commun entre des choses similaires peut être un de leurs éléments constitutifs, leur fonctionnement, leur structure, leur nature, leur cause, une conséquence qu'elles produisent, etc. Nous notons ce point *a priori* évident parce que ce qui est commun est fondamental dans notre recherche, et que la prise en compte de ce qui est commun et de ce qui est distinct inspire les classifications utilisées en linguistique contrastive.

Un trait commun rapproche deux choses similaires. Mais cette ressemblance peut être absolument fortuite, et quel que soit leur degré de ressemblance, il est vain de chercher à tout prix un lien de causalité entre elles. Nous pouvons garder à l'esprit que l'identité entre deux entités est un type particulier, extrême, de similarité, et que la similarité n'a pas forcément de cause.

La définition de la part commune et de la part distincte ne va pas de soi, peut-être en raison de ce qu'il existe un continuum entre la similarité et l'identité. Nous nous proposons d'examiner ci-après les marques de ce continuum dans la langue courante.

### ***1.2.2.2 Un continuum similarité - identité dans la langue courante***

Nous avons tenté de trier dans cette section les mots auxquels renvoie le *Thésaurus* (Péchoin, 1999) pour *similarité* et *identité*. L'examen des mots de sens voisins nous permet de distinguer, artificiellement, cinq degrés dans un continuum qui va de l'absence de similarité à l'identité.

1) Toute similarité est niée :

*X et Y n'ont aucun point en commun ; X et Y n'ont rien à voir.*

2) L'expression du contraste révèle que des points non mentionnés sont similaires :

*Sur ce point, X et Y tranchent (divergent, contrastent) ; par rapport à X, certains aspects de Y ressortent ; en cela, X et Y diffèrent.*

3) L'accent est mis sur la part commune :

*X et Y ont en commun etc. ; X et Y présentent des points communs ; tel point de X se retrouve dans Y ; X et Y ont une caractéristique semblable ; X évoque Y ; X rappelle Y ; X et Y présentent des similitudes ; X et Y se ressemblent en ce que etc. ; X et Y ont une ressemblance.*

4) Le trait commun n'est pas précisé, la frontière avec l'identité semble proche

*X ressemble à Y ; X est semblable à Y ; X est pareil à Y ; X et Y sont similaires ; X est comme Y.*

5) Deux entités en tous points similaires ont un rapport d'identité

*X et Y sont identiques ; X et Y ne font qu'un ; X et Y ne se distinguent pas l'un de l'autre.*

D'autres formulations ne figurent pas dans ce tri : *à l'identique ; indistinctement ; avoir des affinités*, etc. Cette ébauche sur l'expression de la similarité et l'identité en français pourrait être menée en contexte et dans d'autres langues afin de mettre en lumière des fonctionnements sémantiques et syntaxiques intra- et interlinguistiques.

La zone d'incertitude qui existe parfois, comme dans le 4), traduit peut-être la complexité du monde, qui échappe à nos sens et à nos capacités langagières. Toujours est-il que certains mots ou formulations entraînent une ambiguïté.

### **1.2.2.3 Similarité / identité : ambiguïté de certains mots**

Examinons l'ambiguïté du mot *comparable*. On pourrait comparer deux choses et arriver à la conclusion qu'elles n'ont pas de point commun. Or, lorsqu'on affirme que deux choses sont *comparables*, on ne veut pas seulement dire qu'elles peuvent être comparées. Par l'emploi de ce mot, on signifie qu'on les a comparées, et qu'on a observé entre elles une similitude. La comparabilité implique justement souvent qu'on sait qu'elles sont similaires. Par conséquent, *comparable* ne signifie pas seulement « *Qui peut faire l'objet d'une comparaison* »<sup>9</sup>, mais qui présente un degré de similarité.

*Pareil* fait aussi clairement partie de ces mots ambigus. En effet, lorsqu'on dit dans le langage courant que *X et Y sont pareils*, on veut plutôt dire qu'ils sont similaires, mais ils peuvent aussi être identiques.

Les concepts de similarité et d'identité impliquent une comparaison, à l'image du concept d'analogie ; c'est sur ce dernier que nous allons désormais concentrer notre attention.

## **1.2.3 L'analogie, un rapport à quatre ou trois termes**

### **1.2.3.1 Analogie vs similarité**

La différence fondamentale entre la similarité et l'analogie réside dans le nombre minimal d'éléments en présence, qui sont soumis à comparaison : deux éléments minimum dans la similarité et trois à quatre dans l'analogie. En outre, et c'est une différence essentielle, l'analogie dévoile un fonctionnement parallèle : A est à B ce que C est à D. Si A subit une altération, B en subit aussi une pour que le rapport d'analogie demeure. Le nombre d'éléments impliqués, quatre ou trois, dépend de ce que la proportion est discontinue ou continue.

### **1.2.3.2 Cas de la proportion discontinue**

Dans sa *Poétique*, Aristote (1996 : 120) écrit au IV<sup>e</sup> siècle avant notre ère :

« Il y a le même rapport entre la vieillesse et la vie qu'entre le soir et le jour ; le poète dira donc du soir, avec Empédocle, que c'est « la vieillesse du jour », de la vieillesse que c'est « le soir de la vie » ou « le couchant de la vie »<sup>10</sup>.

---

9 Article « Analogie » dans le *TLFi*.

10 Cité par Lepage (2003 : 40-41), cette traduction d'Aristote est de J. Hardy.



Dans la proportion discontinue, A est donc à B ce que C est à D. Moins poète qu'Empédocle, nous traduirons cette image par l'équation suivante à quatre termes :

vieillesse : vie :: soir : jour

La proportion discontinue existe dans les analogies linguistiques telles que les analogies syntaxiques, ou les analogies lexicales, qu'elles soient intralinguistiques ou interlinguistiques.

### 1.2.3.3 *Cas de la proportion continue*

La proportion continue intéresse au premier plan cette recherche. Selon Aristote (1997 : 228-229),

« Que la proportion discontinue implique quatre termes, cela est évident, mais il en est de même aussi pour la proportion continue, puisqu'elle emploie un seul terme comme s'il y en avait deux et qu'elle le mentionne deux fois ; par exemple, ce que la ligne *A* est à la ligne *B*, la ligne *B* l'est à la ligne *C*, la ligne *B* est donc mentionnée deux fois, de sorte que si l'on pose *B* deux fois, il y aura quatre termes proportionnels »<sup>11</sup>.

Lorsqu'on compare des équivalents de traduction, on est dans le cas d'une proportion continue car trois termes entrent en compte :

- 1) le mot de la langue source
- 2) une acception
- 3) le mot de la langue cible

Voici un exemple d'analogie entre équivalents de traduction<sup>12</sup> :

*act* : « *To have an effect* » :: *agir* : « Pouvoir propre à l'homme de transformer (...) »

Les dictionnaires bilingues reposent sur ce rapport de correspondance sémantique plus ou moins exact entre deux mots pour une acception. La prise en compte de la forme des deux équivalents sémantiques complexifie la comparaison lexicale.

---

11 Cité par Lepage (ibidem : 41).

12 Nous tirons ces définitions du *Wiktionary*, <https://en.wiktionary.org/wiki/act#Verb> et du *TLFi*.

## 1.3 De la confusion des concepts de similarité, d'identité et d'analogie appliqués au lexique

### 1.3.1 Des outils pour la comparaison lexicale intralinguistique

En morphologie, l'analogie est omniprésente : les verbes d'une classe sont conjugués selon un paradigme flexionnel, les noms et adjectifs déclinés selon des paradigmes de déclinaison, et bien des mots sont construits en adjoignant des affixes à un radical. Nous revenons plus loin (Approche des principaux types d'analogies intralinguistiques et interlinguistiques, p.48), sur les différents types d'analogies inter- et intralinguistiques car les analogies qui existent au sein d'une langue correspondent à des analogies qui existent au sein d'une autre langue, et c'est sur ce système d'analogies du deuxième ordre que nous proposons une approche du FOS. Dans un premier temps, cependant, nous nous en tenons aux analogies lexicales, et dans cette section, aux analogies lexicales intralinguistiques.

Dans une même langue, deux mots se distinguent par leur forme (orale, écrite) et/ou leur signification<sup>13</sup>. La plupart du temps, deux mots d'une langue se distinguent en ce qu'ils n'ont ni la même forme, ni la même signification, par exemple (fra)<sup>14</sup> *série* et *ruine*. Cependant, les langues connaissent des mots qui peuvent être associés en raison de sèmes qu'ils ont en commun, ou de leur forme semblable. Les dictionnaires analogiques, par exemple le *Nouveau dictionnaire analogique* de Larousse (Niobey *et al.*, 1980) présentent un rapport de similarité sémantique entre les mots qu'ils associent, bien plus qu'un rapport d'analogie. Il en est de même d'autres outils tel l'excellent *Thésaurus* (Péchoin, 1999).

Geysen (1990 : 8, note 4) explique ainsi le phénomène de l'analogie lexicale :

« Le mécanisme d'analogie est la tendance naturelle et inconsciente à donner à des mots nouveaux, par imitation d'un modèle existant, une même forme se reproduisant invariablement, toujours la même pour tous les cas semblables, dans une série de mots, les façonnant tous à son image ».

Deux cas extrêmes de l'identité, sémantique ou formelle, se détachent : la synonymie et l'homonymie. Dans ces deux cas, on peut parler d'analogie. Deux mots synonymes ont

---

13 Mais aussi leur rôle syntaxique, etc.

14 Pour le codage des noms de langues, nous utilisons la norme ISO 639-3 que nous mettons entre parenthèses : (eng) désigne l'anglais, (fra) le français, etc.

(quasiment) la même signification pour au moins l'une de leurs acceptions. La synonymie est un rapport de sens identique qu'on pourrait formaliser ainsi :

mot A : sens X :: mot B

L'homonymie regroupe les sous-notions d'homographie et d'homophonie<sup>15</sup>. Deux mots homophones présentent la même réalisation orale mais pas la même signification : l'analogie entre *air* et *ère*, en français, est formalisable ainsi :

mot A : réalisation orale X :: mot B

Deux mots homographes présentent la même forme écrite mais pas la même signification : le nom français pluriel *notions* et la forme écrite de l'imparfait de l'indicatif du verbe *noter* à la première personne du pluriel (nous) *notions*. Dans cet exemple, l'homographie s'explique par le paradigme de conjugaison en français : la flexion de l'imparfait à la première personne du pluriel est *-ions*. L'équation suivante exprime ce type d'analogie intralinguistique :

Mot A : forme écrite X :: mot B

Les grammaires décrivent ces particularités, et les élèves qui apprennent la langue dans laquelle ils sont scolarisés y sont sensibilisés. Le phénomène de l'analogie lexicale interlinguistique déclenche plus de passions.

---

15 Pour des statistiques portant sur 640 d'homophones en français, voir Ferrand (1999 : 687-708).

## 1.3.2 Mesurer la similitude / l'analogie interlinguistiques

### 1.3.2.1 *Que comparer ? Forme et/ou sens ?*

Geysen (ibidem : 17) dresse dans son *Dictionnaire des formes analogues en 7 langues* des séries de mots latins, italiens, espagnols, français, anglais, néerlandais et allemands. Il précise ainsi sa méthode pour ce qui est du sens des mots :

« La signification d'un mot peut avoir évolué dans le temps depuis ses origines [...]. S'il y en a plusieurs, elles peuvent ne pas se recouvrir toutes ('match', 'performance' en E et F) ; il suffit cependant qu'une au moins corresponde (D eitel, N eidel, E idle, F vain), les mots étant choisis avant tout en fonction de leur forme. »<sup>16</sup>

Dans sa thèse, Caure (2009 : 26-27) mentionne LeBlanc et Séguin (1987), lesquels dressent à partir de deux dictionnaires une liste de 6 447 mots homographes entre l'anglais et le français, et de 16 713 paragraphes, autrement dit de mots qui présentent une forte similarité formelle. Mais Séguin (1994 : 196) reconnaît que les paragraphes ont été répertoriés « au jugé », par conséquent sans appliquer de critères définis.

Cette question de la délimitation des similarités formelles est soulevée par Kauffer, qui note (2016 : 339) :

« On peut certes encore comprendre ce qu'est une identité de forme, à condition bien sûr de préciser si cette identité inclut ou non l'orthographe, la prononciation, l'accentuation, le genre grammatical etc. Mais que signifie une « ressemblance » de forme, c'est-à-dire une identité partielle de forme ? Comment en dégager des critères rigoureux ? À partir de quand deux mots se « ressemblent-ils » ou ne se « ressemblent-ils plus ? ».

Plusieurs termes sont employés en concurrence en lexicologie multilingue pour désigner ces mots qui ont une forme identique ou proche. Parmi ces appellations, on compte *cognats*, *cognates*, et *congénères*. Mais ce sont souvent à la fois la forme et le sens qui sont au cœur de la comparaison interlinguistique.

Les cognats, selon Victorri (2008 : 119), sont « des mots qui se correspondent de langue à langue, tant sur le plan phonétique que sémantique ». *Cognat* n'est pas défini dans une acception linguistique dans le *Trésor de la Langue Française*, ni dans le *Dictionnaire Larousse*. Pourtant, des ressources comme *TermSciences* donnent (fra) *cognat* comme

---

16 Les lettres D, N, E et F, que Geysen met en gras, désignent l'allemand, le néerlandais, l'anglais et le français.

équivalent de (eng) *cognate*<sup>17</sup>. *Cognate* est traduit en français par TERMIUM Plus comme « mots apparenté »<sup>18</sup>. Il est aussi emprunté en français par certains chercheurs.

Cognat / cognate peut suggérer par son étymologie que les deux mots proviendraient d'une seule et même langue. Or la ressemblance lexicale procède de toutes sortes de cas, dont une éventuelle évolution parallèle de langues à partir d'un même étymon, certes, ou bien un emprunt tardif à une langue ancienne (cf. 2.3.2.6, p.94), ou bien encore un emprunt à une autre langue par laquelle le mot a transité (cf.2.3.3.2, p.97), etc.

D'autres chercheurs n'incluent pas l'origine étymologique. Bogaards (1994 : 153) définit des *cognates* comme des « mots de langues différentes ayant (à peu près) les mêmes formes et (à peu près) les mêmes sens ». Cette définition, qui englobe des cas sémantiquement et formellement divers, ne s'accompagne pas des critères rigoureux dont Kauffer déplore le manque. De la même manière, pour Robert (2004B, note 29), les cognates présentent à la fois « similitude entre deux (ou plusieurs) signifiants de langues différentes et similitude partielle au niveau du signifié ». Enfin, *mots congénères* est défini par *Le Grand Robert de la langue française* (Rey, 1988 : 819) comme « de la même famille », autrement dit *congénère* fait aussi référence à une origine commune des mots. D'autres chercheurs considèrent cependant les congénères comme des mots qui, sans être étroitement liés par leur étymologie, ont un degré d'analogie fort (Degache, 2006 : 43). Nous avons beaucoup hésité sur l'appellatif à utiliser, mais face à ces définitions qui englobent des cas différents, nous parlons de *mots analogues*, donnant ainsi la priorité à un rapport entre une forme et un sens pour des mots de deux ou plus de deux langues.

Nous verrons les travers qu'implique la recherche de liens génétiques entre des familles de langues (cf. D'une langue supposée à un peuple supposé, p.74). Dans cette recherche, nous étudions la ressemblance formelle indépendamment de toute origine avérée ou supposée, quand bien même nous sommes sensible à l'héritage du grec et du latin. Pour anticiper notre propos, notons une évidence : deux mots qui entretiennent une ressemblance formelle, ou même une identité formelle, n'ont pas pour autant la même signification (cf. Définition habituelle de faux amis et failles de cette définition selon Kauffer, p.43).

Mais à partir de quand considérer que deux mots de langues différentes ont une forme similaire ? Partant d'une similitude morphosémantique partielle ou totale, des classifications ont été proposées par des chercheurs de l'intercompréhension. Nous en examinons deux.

17 <https://frama.link/4qwAT57X>

18 <https://frama.link/otkH0Pd1>

### 1.3.2.2 Examen de deux classifications utilisées en intercompréhension

Degache (ibidem : 63-65) classe les analogies linguistiques en quatre types :

1) l'analogie totale réfère aux cas de correspondance morphosémantique totale.

Deux cas sont distingués : la forme graphique (quasi-) originale, par exemple<sup>19</sup> les emprunts intégraux (spa) *tisú* et (fra) *tissu*, et les congénères homographes tels que *televisión* et *télévision*, ou *hamaca* et *hamac* ;

2) l'analogie partielle englobe les cas où la correspondance morphosémantique est partielle.

Quatre sous-types en font partie : emprunts intégrés, dont l'acception initiale a été perdue, comme *bombón*, qui désigne un bonbon au chocolat ; congénères parographes, comme *diente* et *dent* ; congénères homonymes, comme *farsa* et *farce* ; para-congénères pour lesquels un ajustement sémantique est nécessaire, comme *receta* qui signifie *ordonnance*, le français *recette* référant lui aussi à un type d'écrit.

3) l'analogie marginale est « la correspondance morphologique totale ou partielle entre items sémantiquement éloignés et difficilement assimilables ».

Degache distingue là quatre cas : congénères improbables, que l'on rencontre dans des contextes différenciés, comme *gafas* (lunettes) et *gaffes* ; congénères divergents, qui ont une acception très éloignée du sens de base, comme *escribano* (notaire) et *écrivain* ; congénères polysémiques en langue cible, comme *esperar*, qui signifie à la fois *espérer* et *attendre* en français ; et (quasi-)homographes fortuits, comme *col*, un légume en espagnol (le chou) alors qu'en français, le *col* désigne initialement le cou, et par extension une partie d'un vêtement.

4) l'analogie trompeuse est la « correspondance morphologique totale ou partielle entre items sémantiquement divergents mais facilement assimilables ».

Ce type d'analogies est subdivisé en cinq : les emprunts divergents en langue source ou langue cible, comme *chalé* (pavillon, villa) et *chalet* (maison en bois, souvent en montagne) ; les congénères périphériques, qui risquent d'être trompeurs dans un contexte ambigu, comme

---

19 Les exemples sont donnés pour l'espagnol langue cible pour des francophones.

*pasaje* (billet) et *passage* ; les congénères discursifs, comme *después* (après) et *desde* ; les congénères sémiques, avec des traits sémiques communs, comme *constipado* (enrhumé) et *constipé* ; et enfin les congénères sémantiques dont l'un est polysémique en langue source ou langue cible, comme *contestar*, qui signifie à la fois *contester* et *répondre*.

Degache (ibid. : 66) note que ce classement propose « un changement de perspective qui permet d'apprécier plus finement, d'une part les possibilités de procéder à un glissement philologique, morphologique ou sémantique, et, d'autre part, à identifier différentes catégories d'analogies problématiques et à les hiérarchiser ».

Pour certains des sous-types identifiés par Degache, cependant, il s'agit, semble-t-il, moins d'un rapport d'analogie que d'un rapport de similarité entre deux termes. En effet, nous avons vu que l'analogie implique que ce que A est à B, B l'est aussi à C. On le retrouve dans plusieurs cas.

Quand la correspondance morphosémantique est totale, les trois termes de l'équation sont présents, avec une acception commune :

*hamaca* : « *Red alargada (...)* » / « *Rectangle de toile ou filet suspendu (...)*. »<sup>20</sup> :: *hamac*

Il en est de même dans certains cas de correspondances morphosémantiques partielles, à l'instar de<sup>21</sup> :

*diente* : « *(...) órgano de masticación o de defensa* » / « *Organe dur, blanchâtre* » :: *dent*

Les congénères polysémiques montrent également une analogie si les deux mots partagent une acception. Mais dans les autres cas, quand les mots ont uniquement la forme en commun, nous parlerions plutôt de similarité formelle interlinguistique que d'analogie.

---

20 Les définitions proviennent du *Diccionario de la Lengua Española* et du *TLFi*.

21 <http://dle.rae.es/?id=DiDC7iY>; et *TLFi*;

Castagne (2007A : 157-158) a proposé une classification dans laquelle il distingue les *mots transparents* des *mots non transparents*. Il définit ainsi les premiers :

« (...) nous appellerons « mots transparent » tout terme lexical, présent dans un texte en une LE donnée qui d'une part possède une forme qui suggère un terme lexical en LM de forme (écrite ou orale) quasi identique, et qui d'autre part se trouve posséder une signification quasi-identique dans les deux langues. »<sup>22</sup>

Par *mot non transparent*, Castagne entend :

« tout terme présent dans un texte en LE qui peut être jugé illisible par un lecteur « débutant » dans cette même LE à partir des connaissances extraites dans sa LM »<sup>23</sup>.

Ces deux définitions réfèrent à une capacité subjective à reconnaître dans la forme d'un mot d'une langue un mot d'une autre langue. Cette capacité, difficile à mesurer quantitativement, dépend entre autres facteurs de la L1 du lecteur, des autres langues de son répertoire et de leur typologie. Murillo Puyal *et al.* (2008 : 343) critiquent l'utilisation du terme de transparence :

« la notion de « transparence », souvent invoquée pour rendre compte des phénomènes d'intercompréhension spontanée entre locuteurs de langues romanes distinctes et, d'une manière générale, entre locuteurs de langues distinctes mais typologiquement et/ou génétiquement apparentées, est par trop impressionniste et imprécise, et la référence à des étymons communs, dont le « locuteur moyen » actuel n'a en général ni conscience ni connaissance, n'est pas susceptible d'applications didactiques ».

Nous préférons ne pas utiliser l'adjectif *transparent* pour désigner ce lexique. En effet, notre étude lexicologique multilingue a été menée dans la perspective d'un public cible de locuteurs de langues distantes diverses, typologiquement éloignées de l'anglais et des langues romanes, un public avec différentes traditions d'apprentissage (Robert, 2009A). En outre, le dispositif envisagé n'a pas fait l'objet d'un test. Nous ne savons donc pas si le lexique que nous avons obtenu peut être considéré comme transparent. Pour ces raisons, nous parlerons de lexique similaire et de lexique analogue. Le lexique analogue est déterminé à partir de calculs de similarité d'équivalents de traduction en langues romanes de mots anglais.

---

22 LE signifie langue étrangère, LM langue maternelle.

23 LM signifie langue maternelle.



### 1.3.2.3 Outils de mesure de similarité

Les outils de TAL qui mesurent la similarité formelle prennent en compte uniquement la forme des mots. Ils occultent le processus subjectif qui a lieu dans la compréhension. Ce faisant, ils sont réducteurs aux yeux du didacticien : la compréhension d'un mot ne dépend pas seulement de sa similarité avec un mot connu du lecteur : le domaine du texte, les connaissances du lecteur, son habitude à inférer le sens d'une langue étrangère, le contexte phrastique, etc., comptent aussi<sup>24</sup>. Aussi, pour une même acception, le lexique similaire que nous avons obtenu grâce au TAL n'est que le lexique analogue. Il serait impropre de le considérer comme lexique transparent.

Examinons quelques uns de ces outils. La distance de Levenshtein (1965 : 845-8) mesure le coût du passage d'une chaîne de sons à une autre en fonction des opérations de suppression, d'insertion et de substitution qui sont nécessaires pour passer d'un mot à un autre. Caure (ibid. : 29) précise : « dans d'autres travaux (Kessler, 1995 ; Nerbonne *et al.*, 1999), toutes les opérations n'ont pas le même coût (la substitution d'une voyelle à une autre voyelle coûte par exemple moins que la substitution d'une voyelle à une consonne) ». Cet aspect qualitatif se retrouve dans certains outils qui comparent la part commune entre deux chaînes de caractères.

Le coefficient de Dice multiplie par deux le nombre de caractères communs et ramène le résultat ainsi obtenu sur le nombre de caractères total. Caure (ibid. : 29) note la limite de ce coefficient appliqué au lexique : les anagrammes semblent présenter une similarité parfaite. Ce n'est pas le cas avec la distance de Jaro, qui tient compte de l'éloignement des caractères identiques. La distance de Jaro-Winkler (Winkler, 1990) améliore cette dernière en valorisant la présence d'un préfixe commun. Appliquée à la recherche de mots similaires, la méthode des 4-grams (Simard *et al.*, 1992) considère deux mots comme cognats s'ils ont au moins leurs quatre premiers caractères en commun. Une autre méthode d'identification donne de meilleurs résultats, celle qui recherche les sous-chaînes communes maximales (SCM), Kraif (2014 : 22) : la comparaison des chaînes de caractères permet d'identifier la plus longue sous-chaîne commune. C'est la méthode que nous avons utilisée (cf. Mesure de similarité par le script CompareCols.csv.pl, p.189).

---

24 Nous signalons un outil pour l'intercompréhension de l'oral, l'indice de transparence sonore proposé par Jamet (2007 : 333-355), qui a étudié le degré de reconnaissance à l'oral de mots français par des locuteurs italiens qui n'avaient pas appris le français. L'indice prend en compte plusieurs facteurs pondérés : le degré de ressemblance vocalique entre les deux congénères, le degré de ressemblance consonantique, le nombre de syllabes du mot français, le degré de ressemblance de la première syllabe, le nombre de phonèmes identiques, la place respective de l'accent tonique, et la longueur respective des deux mots.

Aucun de ces outils n'intègre les systèmes graphémiques des langues, si bien que les allographèmes que sont par exemple <ph> et <f> en anglais (en allemand, en français) ne sont pas pris en compte dans les mesures de similarité. Pourtant, un apprenant d'espagnol L3 qui connaît (eng) *philosophy* devrait reconnaître ce mot dans le mot (spa) *filosofia*, dans lequel le phonème /f/ est transcrit par l'allographème <f> qui existe aussi en anglais et est plus répandu que <ph>. Nous proposons plus loin d'en tenir compte (cf. Améliorer les mesures de similarité par la prise en compte des allographèmes, p.242).

#### **1.3.2.4 Définitions : intersystème, interlexème, interlexique, interforme et intersémème**

Comme Caure (ibid. : 83), nous nous inscrivons non pas dans le système de chacune des langues étudiées, mais dans l'intersystème, « le système dans lequel s'organisent les différents éléments sinon identiques, du moins très proches, entre langues ». L'intersystème contient des interlexèmes, définis ainsi par l'équipe d'EuroCom (Klein & Rutke : 1997) :

« L'interlexème européen est une unité de base du lexique plurilingue qui représente les (allo-)lexèmes de différentes langues qui, du point de vue du concept, sont équivalents et, du point de vue de la forme, se présentent comme congruents (très proches). »

Caure (ibid. : 83) précise qu'un interlexème n'est pas un mot, mais

« l'unité théorique supérieure qui « représente » les mots effectifs des langues, qui possède les traits communs aux différentes réalisations [...] L'ensemble des interlexèmes pour les langues données constitue leur « interlexique » (Meißner *et al.*, 2004 : 39) ».

Elle le reformule (ibid. : 85) comme suit :

« Un interlexème est une unité théorique supérieure à des lexèmes donnés de plusieurs langues. Il est la somme des points communs à ces différents lexèmes. »

Caure précise également les notions d'*interforme* et d'*intersémème*, qu'elle définit ainsi (ibid. : 85) :

« Une interforme est une unité théorique supérieure à des formes données attestées dans des langues différentes. Elle rend compte du versant formel de la transparence directe entre plusieurs mots, sans prise en compte aucune du versant sémantique ».

Elle donne l'exemple de la forme graphique *sein* qui n'a pas la même signification en français, allemand et néerlandais, et la rapproche de la notion d'homonymie et de paronymie. Nous comprenons tout autrement l'idée d'une interforme : si l'interforme correspond à la définition qu'en donne Caure, elle réfère plutôt à une forme théorique qui représente les formes congruentes des langues comparées, une sorte de dénominateur commun de la forme, une représentation de la sous-chaine commune. Dans l'exemple qu'elle donne, l'interforme de (por) *comunicação*, (spa), *comunicación*, (ita) *comunicazione*, (eng) / (fra) *communication*, (deu) *Kommunikation* et (ndl) *communicatie*, pourrait être quelque chose comme :

\*com(C)unicaCVV(C)(V) (interforme à l'écrit)

\*kVmV(V)nikVC(C)(C)V(C)(V) (interforme à l'oral)

Dans cette représentation de l'interforme de ces sept langues, C figure une consonne et V une voyelle qui varient selon la langue, et (C) une consonne et (V) une voyelle qui s'ajoutent à la sous-chaine commune dans certaines de ces langues.

Caure énonce plus loin (ibid. : 101) la règle suivante :

« pour que deux mots soient proches formellement directement, ils doivent avoir la même « trame consonantique », c'est-à-dire les mêmes consonnes dans le même ordre, avec éventuellement dans un des mots une consonne de plus ou de moins au contact de ses consonnes initiales ou médianes, et la dernière consonne, si elle est au moins la troisième, éventuellement différente ».

La trame consonantique, ou racine bilittère / trilitère / quadrilitère, est un outil de premier ordre en arabe et en hébreu pour classer les mots de la même famille, mais aussi pour former les mots de ces langues par analogie, et nous pensons qu'il est sous-exploité pour décrire les similitudes interlinguistiques. Caure définit par ailleurs (ibid. : 85) l'intersémème :

« Un intersémème est une unité théorique supérieure aux significations de mots donnés de langues différentes. Il rend compte du versant sémantique de la transparence directe entre plusieurs mots, sans prise en compte aucune du versant formel ».

Elle donne comme exemple de représentants d'un même intersémème (fra) *lunettes*, (eng) *glasses* et (deu) *Brille*. Si un interlexème est la somme des points communs à différents

lexèmes, on peut de manière analogue se figurer un intersémème comme la somme des sèmes communs à ces lexèmes. Caure signale que Braun (2002 : 187 sq.) a forgé le mot *intersémème*. Braun lui attribue le sens d' « acception commune à des mots de langues différentes ». Le terme répété dans les équations analogiques que nous avons formulées correspond à la définition de Braun. Si l'on voulait représenter plus précisément un intersémème, on pourrait établir le sémème de chacun des mots envisagés et mettre en évidence les sèmes qu'ils partagent.

Nous nous en sommes tenus, dans ce travail, aux mots dont nous avons mesuré la similarité en fonction de critères formels pour une même acception. L'interlexique anglais-espagnol-français-italien-portugais est un lexique en cinq langues composé de mots anglais fréquents dans un corpus de spécialité et de leurs équivalents de traduction similaires dans les quatre langues romanes. Il repose sur la similarité des équivalents de traduction pour les 2008 acceptions fréquentes des 2000 adjectifs, noms et adverbes anglais fréquents d'un corpus de textes scientifiques de spécialité.

Lorsque deux représentants d'une interforme n'expriment pas les mêmes sèmes, ces mots sont appelés des *faux amis*.

### **1.3.3 Un type de similarité lexicale interlinguistique : les *faux amis***

#### ***1.3.3.1 Définition habituelle de faux amis et failles de cette définition selon Kauffer***

L'appellation *faux amis* est due à Koessler et Derocquigny (1928), qui listent 1200 mots anglais similaires dans leur forme à des mots français mais de sens différent. Le titre entier de leur ouvrage, *Les faux Amis ou les Trahisons du vocabulaire Anglais (conseils aux traducteurs)*, donne la mesure de la défiance qui accompagne cette question. Nous abordons également l'image négative d'une autre langue dans la section 2.3.1.1 (Substrat, superstrat, adstrat), p.81.

La ressemblance formelle fortuite contredit la tendance de l'esprit humain à décoder ce qui présente des similitudes avec le connu. Les couples de *faux amis* ont une forme écrite et/ou orale si proche qu'elle crée chez le locuteur / lecteur l'impression qu'ils devraient avoir le même sens. Mais (eng) *rate* et (fra) *rate*, par exemple, sont seulement homographes, ils n'ont sémantiquement rien à voir, ils ne partagent pas de sèmes. Loin d'être un phénomène annexe,

les faux amis ont fait l'objet de dictionnaires entiers (par exemple Thody et al. (1985) pour le français et l'anglais, ou Cantera Ortiz de Urbina *et al.* (1988) pour le français et l'espagnol. Kauffer (2016 : 337) en rappelle la définition communément admise :

« On entend ordinairement par faux amis des *mots* de *deux langues* ayant une grande *ressemblance de forme* mais des *sens différents*, ce qui est source de confusion et donc d'*erreur* pour le locuteur »<sup>25</sup>.

Puis, Kauffer critique plusieurs points de cette définition : les faux amis peuvent être non seulement des mots, mais des expressions ; ils peuvent être liés à une tierce langue à laquelle deux langues ont emprunté, ou même à une seule langue ; la forme peut être proche par pure coïncidence ; le phénomène des faux amis est avant tout individuel, l'erreur étant liée à la connaissance linguistique du locuteur ; la différence de sens définitoire peut être très diverse ; la ressemblance de forme peut être due à une origine commune.

Les critiques de Kauffer s'appuient sur la nature des faux amis, les causes de la similitude formelle, l'extension de la différence sémantique et la perception individuelle du phénomène. Si nous nous en tenons à la forme et au sens, nous pouvons tenter une représentation du phénomène des faux amis.

### ***1.3.3.2 Réflexion et propositions autour des faux amis***

On pourrait formaliser les faux amis comme suit, où X et X<sup>(0)</sup> représentent deux formes (quasi-)identiques :

Forme (écrite ou orale) X dans une langue A : sens Y

Forme (écrite ou orale) X<sup>(0)</sup> dans une langue B : sens Z

On n'observe pas d'analogie entre ces quatre termes, il y a seulement une similarité (éventuellement une identité) entre la forme X et la forme X<sup>(0)</sup> des mots comparés. On pourrait donc en tirer une interforme.

Il peut être utile de faire observer à des apprenants de quoi procèdent les faux amis, que l'on pourrait appeler plutôt des *sosies* de manière à créer un rapport moins affectif au phénomène.

<sup>25</sup> Les italiques sont de Kauffer.

Les sosies pourraient être considérés avec moins de passion qu'ils ne le sont, comme un phénomène interlinguistique comparable à ce qui existe dans une même langue.

Des sosies qui ont exactement la même graphie, comme (eng) *singer* (en français : chanteur.euse) et (fra) *singer* (imiter comme peut le faire un singe), sont semblables à des formes homographes au sein d'une langue, par exemple en anglais la forme écrite *bear*, qui est un verbe (en français : porter) et un nom (ours).

Des sosies qui ont la même prononciation<sup>26</sup>, comme (spa) *rato* (en français : moment, instant) et (fra) *râteau* (instrument pour ratisser), sont semblables à des homophones au sein d'une langue, par exemple en anglais l'adverbe *no* (non) et le verbe *know* (savoir).

Des sosies qui ont une forme écrite très proche sans aucun rapport de sens, comme (spa) *sopa* (en français : soupe) et (eng) *soap* (savon), sont semblables à des paronymes homographes (ou parographe), c'est-à-dire des mots d'une langue dont seules une ou quelques lettres diffèrent, par exemple en français *complément* / *compliment*.

Des sosies qui ont une forme orale proche sans aucun rapport de sens, comme (eng) *tray* (en français : plateau) et (fra) *très* (intensément) sont semblables à des paronymes homophones (qu'on pourrait appeler des *parophones*), c'est-à-dire des mots d'une langue dont peu de sons diffèrent, par exemple en français *panser* (soigner) et *pincée* (petite quantité).

Nous pouvons résumer ces quatre cas en deux phrases qui, si on y réfléchit, expriment des évidences. Une forme écrite ou orale peut référer à deux réalités différentes, que ce soit au sein d'une langue ou entre deux langues. Deux formes semblables à l'écrit ou à l'oral peuvent référer à deux réalités différentes, que ce soit au sein d'une langue ou entre deux langues.

Mais ces homonymes ou ces faux amis méritent-ils vraiment toute la place qui leur est donnée ? Rencontrer des mots isolés de tout contexte n'est pas si fréquent dans le discours : ordinairement, un mot est éclairé par le contexte, c'est-à-dire par ses cooccurrents. La redondance des sèmes, activés en parallèle grâce à la présence de cooccurrents qui entretiennent des relations sémantiques, guide généralement vers une interprétation correcte. La résolution de l'ambiguïté tient probablement à la faculté du lecteur à inférer le sens, la connaissance linguistique du locuteur que mentionne Kauffer. Castagne (2007A : 159) estime du reste que l'opacité des faux amis n'est pas un problème fondamental parce que, notamment, « le contexte nous remet très rapidement sur le droit chemin ».

---

26 Indépendamment de l'accent tonique et de la réalisation différente du /r/ et du /o/ dans notre exemple.

Ce qui nous semble intéressant dans les sosies, ce sont les cas où ils partagent certains sèmes. Les bilingues connaissent bien ce problème, qui doivent constamment établir des rapports sémantiques entre l'une et l'autre langue parce que les sèmes de deux équivalents sémantiques ne sont pas entièrement superposables (et nous retrouvons là l'un des sens de *ana-* en grec).

Nous préférons considérer qu'il est intéressant pour un didacticien que deux mots qui ont une forme similaire dans deux langues puissent diverger dans leur sens. Ces mots peuvent en effet être examinés en diachronie, c'est-à-dire resitués dans leur évolution sémantique et/ou morphologique. Ce faisant, l'enseignant peut en faire aborder la portée métaphorique et symbolique. Un tel travail fait aussi appel à l'évolution de la L1 de l'apprenant, qui connaît également des emprunts avec un sens restreint. Prenons l'exemple du français *fémur*, emprunté au latin *femur*. En latin, ce mot désigne la cuisse. Le français l'a emprunté avec le sens d'os de la cuisse afin de pouvoir désigner précisément cette partie du système ostéo-articulaire. Du besoin d'enrichir le lexique procède la polysémie, que Charaudeau (1992 : 63) définit en ces mots :

« Au sens étroit, la polysémie désigne, en synchronie, le fait qu'un même signifiant recouvre des signifiés différents entre lesquels (contrairement à l'homonymie) existe une intersection sémantique ».

La polysémie s'explique aussi en diachronie, et en retraçant l'évolution du sens, on s'affranchit d'approches didactiques des faux amis caractérisées par la défiance vis-à-vis des autres langues. On rattache ainsi les glissements sémantiques au phénomène général d'adaptation du code aux besoins langagiers d'une société.

Pour ce type de mots, un enseignant peut bien entendu recourir aux dictionnaires étymologiques. Il peut également par une approche sur corpus faire comparer aux apprenants le contenu sémémique de deux sosies. Partant de dizaines de concordances, le sémème de chacun des sosies est établi en faisant analyser quels sèmes sont activés dans différents contextes. Cette observation contrastive génère une mise à distance dans laquelle peuvent être dégagés les sèmes qui coïncident, l'intersémème, et selon le cas ceux qui font que le sens de chacun des mots diffère<sup>27</sup>.

Notre objectif terminologique a été de délimiter un interlexique, mais les analogies linguistiques ne se situent pas uniquement au niveau du lexique. Cette réflexion sur les types

---

27 Au demeurant, on peut aussi faire établir le sémème d'un mot d'une langue à partir de concordances tirés d'un corpus.

d'analogies, qu'elles soient intralinguistiques ou interlinguistiques, fait partie plus largement d'un enseignement de la langue par une méthode contrastive, comme nous le décrivons dans la section Perspectives : vers un dispositif didactique plurilingue d'enseignement du FOS sur la base d'analogies interlinguistiques, p.247.



## 1.4 Approche des principaux types d'analogies intralinguistiques et interlinguistiques

Nous avons annoncé (cf. 1.3.1, p.333) que nous reviendrions sur les différents types d'analogies inter- et intralinguistiques du fait que les analogies qui existent au sein d'une langue correspondent à des analogies qui existent au sein d'une autre langue, et que c'est sur ce système d'analogies du deuxième ordre que nous souhaitons proposer une approche du FOS. Aussi récapitulons-nous ci-après les principales analogies concernées - analogies syntaxiques, morphosyntaxiques et morphologiques.

### 1.4.1 Analogies syntaxiques

#### 1.4.1.1 Analogies syntaxiques intralinguistiques

Au sein d'une même langue, l'analogie syntaxique a lieu par exemple dans l'ordre des constituants d'une phrase :

*Il apprend la langue des signes en ce moment*

est analogue à :

*La rivière a recouvert le chemin empierré près du bois*

Dans les deux phrases, on a en effet : sujet, verbe, objet direct, circonstant.

Cette analogie n'empêche pas qu'un autre ordre puisse être possible, ni ne signifie que l'usage de cet ordre soit majoritaire. Nous nous situons là uniquement dans la comparaison de l'existant. En outre, l'analogie ne s'observe ici qu'au niveau de l'ordre et de la fonction de tous ces constituants, elle n'existe pas au niveau de certains de ces syntagmes. Ainsi, *il* est un pronom quand *la rivière* est un groupe nominal, etc.

#### 1.4.1.2 Analogies syntaxiques interlinguistiques

La typologie syntaxique classe les langues en fonction de l'ordre des actants : les langues SOV (Sujet Objet Verbe), les langues SVO (Sujet Verbe Objet), etc. D'autres classifications dépendent de la place des circonstants, celle des modificateurs, etc.

Nous pensons qu'il devrait être possible de s'appuyer sur des analogies lexicales et sur des analogies syntaxiques dans le but de faire acquérir des traits syntaxiques et morphosyntaxiques convergents, ainsi que les traits qui divergent d'une langue à l'autre.

## 1.4.2 Analogies morphosyntaxiques

### 1.4.2.1 Analogies morphosyntaxiques intralinguistiques

La flexion des verbes est l'une des analogies morphosyntaxiques à l'intérieur d'une langue. Gosme (2012 : 20) l'illustre par le futur en français :

« Un autre exemple en conjugaison est *vous deviendrez : ils deviendront :: vous permettrez : ils permettront* où la marque de la troisième personne du pluriel au futur simple « -ront » commute avec la marque de la deuxième personne du pluriel « -rez ». »

Les flexions liées au genre (masculin / féminin / neutre), les flexions liées au nombre (singulier / pluriel) et les déclinaisons relèvent également des analogies morphosyntaxiques.

### 1.4.2.2 Analogies morphosyntaxiques interlinguistiques

Des analogies d'ordre morphosyntaxique justifient qu'on puisse comparer le paradigme des verbes français du premier groupe, *chanter*, avec le paradigme des verbes espagnols du groupe *cantar*. C'est parce qu'on a :

*je chante : tu chantes :: canto : cantas*

qu'on a aussi pour *aimer / amar* :

*j'aime : tu aimes :: amo : amas*

Comme nous l'avons vu pour les analogies intralinguistiques, les flexions du singulier / du pluriel relèvent aussi des analogies morphosyntaxiques. Les analogies suivantes entre l'italien et le français l'illustrent :

*la personne : les personnes :: la persona : le persone*

*le chat : les chats :: il gatto : i gatti*

Mais étant donné que le pluriel d'un nom en italien dépend de son genre grammatical et que des équivalents sémantiques n'ont pas nécessairement le même genre, il faudra combiner les deux séries précédentes pour déduire le quatrième terme de l'analogie :

*la période : les périodes :: il periodo : ?<sup>28</sup>*

Ce type d'équation montre que l'existence d'une similitude formelle de certaines formes de mots, par exemple le nom singulier *canal* en anglais, espagnol, français et portugais, n'oblitére pas les règles de morphosyntaxe propres à chacune des langues : le pluriel de *canal* est (eng) *canals*, (fra) *canaux*, (spa) *canales* et (port) *canais*.

Des équations permettent de rendre compte du fonctionnement d'autres morphèmes grammaticaux selon les langues, par exemple le déterminant possessif de la première personne du singulier en arabe (Chavanes, 2014 : 10) :

*stylo : mon stylo :: kalam : kalamy*

Entrent aussi dans la catégorie des analogies interlinguistiques morphosyntaxiques le fonctionnement des pronoms clitiques, les classificateurs des langues qui en utilisent, etc. Nous ne traitons évidemment pas tous les cas ni ne les mentionnons tous.

### **1.4.3 Analogies morphologiques**

#### ***1.4.3.1 Analogies morphologiques intralinguistiques***

La dérivation au moyen d'affixes (suffixes, préfixes, et infixes dans des langues agglutinantes comme l'estonien) fournit de nombreux exemples d'analogies morphologiques qui induisent un rapport entre la forme et le sens : si l'*adaptation* est le fait d'*adapter*, la *conservation* est (probablement) le fait de *conserver*.

Le lexique des langues synthétiques, telles que les langues romanes et les langues germaniques, se structure en grande partie sur la base de ce type d'analogies de surface.

---

28 *I periodi*.

### 1.4.3.2 Analogies morphologiques interlinguistiques

Le système d'affixation, développé dans les langues romanes et les langues germaniques, permet de formaliser certaines analogies interlinguistiques.

Ainsi, pour ne prendre qu'un exemple évident, on a en anglais<sup>29</sup> et en français des noms en *-ization* / *-isation*, du type (eng) / (fra) construits sur des verbes en (eng) *-ize* / (fra) *-iser* :

*hospitalize : hospitalization :: hospitaliser : hospitalisation*

Des séries entières existent sur cette analogie : *immunize, immunization ; immuniser, immunisation ; localize, localization ; localiser, localisation, etc.*

Là encore, l'existence d'une forme dans l'une des langues ne présage pas de l'existence de son équivalent dans l'autre langue : l'un des mots similaires peut avoir évolué par métaphore, ou bien encore, devenu désuet, il peut avoir été supplanté dans l'usage par un autre.

---

29 En anglais britannique, la variante de ces deux suffixes est homographe avec le français : *-ise* et *-isation*, si bien que ces six mots sont homographes dans les deux langues.

## 1.5 Principales analogies et dissemblances entre l'anglais et les langues romanes

### 1.5.1 Un vaste champ d'étude où l'histoire de la langue est omniprésente

#### 1.5.1.1 *Ce que nous n'abordons qu'accessoirement*

Certains traits de l'anglais et des langues romanes convergent et devraient faciliter un enseignement du FOS qui s'appuierait sur l'anglais comme langue passerelle. Nous voyons plus bas comment des analogies entre l'anglais, le français et d'autres langues sont liées à des origines communes (Origines génétiques communes des langues germaniques et romanes : des langues indoeuropéennes, p.72) et des contacts entre ces langues (et d'autres) (Des analogies interlinguistiques liées aux contacts des langues en Europe de l'Ouest, p.81). Nous synthétisons ici les analogies actuelles dans les domaines du lexique, de la morphologie et de la syntaxe tout en mettant entre parenthèses certains points (syntaxe de la proposition subordonnée, de la négation, adverbes, système des temps, etc.). En effet, nous ne visons pas une comparaison exhaustive, mais une mise en relief de traits qui pourraient se prêter à un apprentissage plurilingue contrastif.

De la même façon, nous négligeons la sémantique comparative, puisque nous souhaitons avant tout étudier ce qu'un apprenant pourrait reconnaître : la forme des unités du lexique et l'ordre dans lequel elles se rencontrent dans la phrase. Cependant, une comparaison sémantique ferait apparaître d'autres analogies : Teyssier (2012 : 106) donne l'exemple des verbes (eng) *play*, (fra) *jouer* et (deu) *spielen*, dont il compare les quatre sens principaux :

« *To play* et *spielen* ont des origines très différentes de celles du français *jouer*, mais les champs sémantiques que ces verbes délimitent sont découpés de façon identique ».

Les faits relevant spécifiquement de la prosodie sont également écartés, étant donné que notre approche repose sur les analogies écrites<sup>30</sup>. Nous signalons cependant en premier lieu un trait phonologique des langues germaniques qui a eu des effets considérables sur le français et l'anglais : l'accentuation tonique de la syllabe à l'initiale.

---

30 Les troisième et quatrième tamis d'EuroCom semblent plus difficiles à valoriser dans une approche de l'écrit.

### ***1.5.1.2 Incidence phonéticohistorique de l'accent à l'initiale des langues germaniques***

Nous éclairons plus loin le contexte historique (cf. De la reconstruction phonéticohistorique des langues indoeuropéennes : à la recherche d'analogies p.73) dans lequel les linguistes comparatistes en sont venus au XIX<sup>e</sup> siècle à reconstituer un système de *correspondances* entre des langues qu'ils ont appelées les langues indoeuropéennes. Ces correspondances procèdent étroitement du principe d'analogie<sup>31</sup>.

Ainsi, les comparatistes ont déduit qu'à l'accent de hauteur mélodique qu'on suppose en indoeuropéen s'est substitué en germanique commun un fort accent d'intensité sur la syllabe initiale des mots simples et sur la syllabe radicale des mots dérivés.

Or la fin des mots, en germanique comme dans les autres langues indoeuropéennes, portait une désinence. L'accent d'intensité sur la première syllabe a entraîné une articulation moindre des syllabes qui ne portaient pas l'accent, et l'amuissement de certains phonèmes. En conséquence, les désinences se sont nivelées et le nombre de formes s'est réduit : en anglais moderne, il n'en subsiste plus qu'une au prétérit, par exemple *had*, là où en gotique on distinguait quatre formes (ibid : 8).

La syntaxe en a également été modifiée : à la fin de la période du vieil-anglais, les désinences casuelles nominales se sont si affaiblies que l'emploi des prépositions s'est étendu afin de pallier la fonction syntaxique qu'avaient exercée jusqu'alors les cas. Le phénomène d'accentuation primaire, qui a touché tous les mots des langues germaniques, explique la présence de nombreux monosyllabes en anglais moderne (Crépin, 1994 : 35). Elle éclaire aussi la contraction syllabique du français, particulièrement dans les mots issus du bas-latin, par rapport à leurs équivalents formels dans les autres langues romanes. L'influence de la prononciation par les Francs des mots du latin régionalisé (Banniard, 1997 : 34) au nord de la Loire a eu un impact considérable : « les Francs auraient accentué les voyelles toniques en syllabe ouverte plus que ne le faisaient les autochtones, tout en les allongeant » (Gadet & Ludwig, 2015 : 27). La tendance à une chute des syllabes finales, qui va à l'encontre d'une reconnaissance de la parenté du français avec les autres langues romanes, est contrebalancée à l'écrit de deux manières. D'une part par une orthographe du français conservatrice, d'autre part

---

31 Il a été déduit par exemple que l'occlusive sourde /p/ des mots d'une protolangue supposée, nommée par commodité l'indoeuropéen, a progressivement muté en spirante sourde /f/ en germanique commun. La loi de Grimm énonce ce phénomène. Dans les langues non germaniques, cette mutation n'a pas eu lieu, ce qui explique l'opposition à l'initiale de (fra) *poisson*, (spa) *pez*, (bre) *pesk*, d'une part, et de (eng) *fish*, (sve) *fisk*, (deu) *Fisch*, d'autre part.

par les très nombreux emprunts savants au latin (cf. section Le latin, langue de la connaissance) que l'on retrouve dans d'autres langues, y compris l'anglais.

### **1.5.2 Analogies et dissemblances syntaxiques de l'anglais et des langues romanes**

Harris (1976), Delmonte et Fernández Lagunilla (1987), Reinheimer Rîpeanu (1993), Muller (2002), et Godard (2003), traitent d'un point de vue contrastif la syntaxe des langues romanes. Nous ne passerons pas en revue toutes les structures syntaxiques. Nous noterons seulement que les langues romanes et les langues germaniques connaissent toutes la syntaxe SVO à la forme déclarative<sup>32</sup>. Ce point commun nous paraît essentiel dans l'optique de faire comprendre immédiatement des phrases dans une langue romane dont la syntaxe est analogue à celle de son équivalent sémantique en anglais.

Certaines des structures syntaxiques panromanes du cinquième tome d'EuroCom (Klein, 2000 ; cf. aussi 7.3.4 p.255) existent également en anglais :

N + V + N

N + V + Adj

N + V + O

N + V

N + V + syntagme prépositionnel

etc.

Le déterminant se place avant le nom dans les langues romanes et les langues germaniques. L'article défini est à part en roumain et dans les langues scandinaves, où il est enclitique : (ron) *un om* (un homme), mais *omul* (l'homme) ; (swe) *en person* (une personne) mais *personen* (la personne).

Dans les langues romanes, l'adjectif se place après le nom dans le cas général. Quelques adjectifs peuvent être antéposés au nom, avec parfois une différence sémantique. En anglais,

---

<sup>32</sup> Sans exclure d'autres possibilités, en cas d'emphase, par exemple.

comme dans les autres langues germaniques, l'adjectif est antéposé au nom. En outre, les langues romanes et les langues germaniques connaissent des tournures attributives.

La simplification progressive du système de déclinaisons dans les langues romanes et les langues germaniques<sup>33</sup> a eu pour effet une syntaxe moins libre et une généralisation du recours à des prépositions quand le système casuel pouvait se passer de prépositions. Lors de cette reconfiguration de la syntaxe, l'usage des prépositions a été élargi, l'anglais a du reste emprunté des prépositions au scandinave.

Il faudrait envisager l'interrogation et la négation ; nous ne développerons pas ce point ici. Nous retiendrons que la syntaxe interrogative SVO se répand dans les usages anglophones à l'oral et que les adverbes de négation *no(n)* des langues romanes présentent une analogie avec l'adverbe anglais *no(t)*.

Les analogies syntaxiques entre des langues sont utilisées dans les systèmes de traduction par analogie (Gosme, *ibidem* : 20-21) qui recourent à la résolution d'équations analogiques.

### **1.5.3 Analogies morphologiques de l'anglais et des langues romanes**

#### ***1.5.3.1 Analogies lexicales de l'anglais et des langues romanes***

Des jugements sont parfois portés sur « la pauvreté de la morphologie de l'anglais » (Viel, 1981 : 11). Des points de comparaisons existent cependant entre cette langue et les langues romanes.

Pour la forme des unités du lexique, la méthode des 7 tamis d'EuroCom (Klein, 2008) distingue le premier tamis, le lexique international, et le deuxième tamis, le lexique panroman. Bien des termes de l'une et l'autre de ces catégories existent sous une forme similaire en anglais.

Au point de vue de la morphologie du lexique, les langues romanes et les langues germaniques sont analogues dans leur recours à :

- la dérivation, principalement au moyen de préfixes et suffixes, dont certains sont analogues d'une langue à l'autre : (eng) *prevalence*, (spa) *prevalencia*, (fra) *prévalence*, (ita) *prevalenza*, (por) *prevalência* ;

---

33 Le roumain et l'allemand sont à part, qui ont conservé davantage de déclinaisons.



- la composition à partir de composants plus ou moins autonomisables : (eng) *side effect, amino acid* ;
- la conversion : (fra) (eng) *orange* peut être nom et adjectif.

C'est le septième tamis d'EuroCom (préfixes, suffixes, « eurofixes »).

Au sujet de la dérivation, un large usage de diminutifs et d'augmentatifs est encore vivant dans les langues romanes. Le français contemporain est à part, qui n'en utilise plus guère (Teyssier, 2012 : 117), et ressemble en cela à l'anglais.

La perspective d'utiliser l'anglais comme langue pont vers les langues romanes nous amène à faire les deux remarques suivantes. Premièrement, l'anglais recourt à la composition de mots dans lesquels le déterminant est placé devant le déterminé (*teenage, stem cell*<sup>34</sup>, etc.). Bacquet (ibid. : 115) note que ce procédé rapproche l'anglais considérablement des (autres) langues germaniques. En revanche, l'usage de prépositions est majoritaire dans les langues romanes. Deuxièmement, la conversion lexicale est très fréquente en anglais<sup>35</sup>.

### **1.5.3.2 Particularités de la dérivation de verbes anglais**

La similarité de certains verbes anglais avec une ou plusieurs langues romanes dépend de la manière dont ils ont été dérivés du latin. Mossé (1947 : 99-100) signale :

« Pour les verbes de la première conjugaison latine en -āre, on a le choix entre deux procédés : on les tire tantôt de l'infinitif comme en français, *consultāre* > fr. *consulter* ; angl. *to consult*, tantôt, suivant un procédé propre à l'anglais, du participe passé, *consolidatus* > angl. *to consolidate* ».

### **1.5.3.3 Analogies du lexique de la santé et de la chimie**

Le lexique de la santé (cf. Deux catégories sémantiques pour les 2208 acceptions fréquentes, p.152), a des origine latines et grecques, comme nous le voyons dans la section Emprunts lexicaux au latin et au grec en biologie et en médecine, p.94. Les procédés de néologie qui ont cours dans ces domaines s'appuient souvent sur des préfixes / suffixes / suffixoïdes grecs et latins (Corvez, 2011).

34 En français : cellule souche.

35 Le phénomène de conversion en anglais aurait été favorisée par les emprunts au français (Lee ; 1948).

Dans certaines spécialités scientifiques, et notamment en médecine et en chimie, on systématisé la dérivation, voire la composition : pour des termes tels que *paradichlorobenzène* (para-di-chloro-benz-ène) ou cholecystite (chole-cyst-ite), les procédés lexicaux employés sont systématiques et répondent au besoin pragmatique d'intelligibilité immédiate pour les spécialistes qui connaissent les racines lexicales du domaine. Un grand avantage pour l'apprenant de langue réside dans la construction analogue d'une grande partie du lexique de spécialité dans les langues de notre étude (et dans d'autres).

## 1.5.4 Analogies et dissemblances morphosyntaxiques

### 1.5.4.1 Analogies et dissemblances morphosyntaxiques des noms

Les éléments morphosyntaxiques constituent le sixième tamis d'EuroCom.

Les langues romanes et l'anglais marquent de manière inégale le cas, le genre, et le nombre. Le nom roumain peut avoir trois formes distinctes selon le cas<sup>36</sup>, mais dans les autres langues romanes et en anglais, le système casuel qui existait en latin a disparu pour les noms.

En vieil-anglais, il existait trois genres et cinq cas. Les cas ont disparu de la morphologie des noms en anglais moderne, par simplification, sauf le génitif, sous la forme du cas possessif, ou génitif saxon. En anglais, il est en concurrence avec la tournure périphrastique *the N of N* et avec les noms composés.

Deux genres grammaticaux existent dans les noms des langues romanes : le masculin et le féminin. Le genre est la plupart du temps arbitraire. En catalan, espagnol, italien et portugais, le genre de la plupart des noms se devine du fait d'un suffixe. En français également, certains suffixes sont typiques de l'un ou l'autre genre, mais il n'est pas rare que la finale n'en permette pas l'identification : les mots en -e, très fréquents, peuvent être masculins ou féminins.

En outre, certains noms des langues romanes sont épïcènes. En français, ce sont souvent des mots en -e : *collègue, adulte, arabe*. Aux noms s'ajoutent des suffixes épïcènes : *-gène, -logue*, etc. On retrouve ce trait dans les autres langues romanes, mais il ne concerne pas nécessairement les mêmes mots. Dans nos données, (fr) *collègue*, (ita) *collega* et (por) (spa) *colega* sont épïcènes. En revanche, alors que (por) *estudiante* et (spa) *estudiante* sont épïcènes,

---

36 Une forme est commune au nominatif et à l'accusatif, une autre au datif et au génitif, une dernière peut exister pour le vocatif.

leurs équivalents sémantiques se différencient au masculin et au féminin : (ita) *studente / studentessa* et (fra) *étudiant / étudiante*.

Ni leur déterminant, ni leur morphologie ne distinguent les noms anglais en fonction d'un genre.

Le pluriel des noms en français (et en catalan, espagnol, occitan et portugais) provient de l'accusatif latin ; dans le cas général, il se forme par l'adjonction d'un -s<sup>37</sup>. Il en est de même en anglais : dans le nord de l'Angleterre, un pluriel en -s a été général après 1250<sup>38</sup>.

Lorsqu'un mot anglais est identique dans sa forme à un mot français, le pluriel est bien souvent identique<sup>39</sup>. Dans l'image 1, on ne peut dire si l'inscription en allemand est traduite en français ou en anglais puisque la formulation est valable dans les deux langues.



Image 1 : Photographie d'un panneau d'affichage sur le quai de la gare Flughafen Graz-Feldkirchen, Autriche. Crédit : Gilles CC-BY-NC

#### **1.5.4.2 Analogies et dissemblances morphosyntaxiques des adjectifs et des adverbes**

Les adjectifs des langues germaniques connaissent une double déclinaison, forte ou faible. La déclinaison faible en -e(n) se rencontre après un déterminant, par exemple (dan) *den røde liste* (la liste rouge), (deu) *die rote Liste* (idem), (nld) *de lange weg* (le long chemin). La déclinaison forte se rencontre à la forme indéfinie, par exemple sous la forme d'un -t en norvégien : *et grønt eple* (une pomme verte).

37 La forme plurielle de l'italien et du roumain rappelle le nominatif latin.

38 On trouve bien entendu des exceptions, mais nous n'insisterons pas sur ce qui est dissemblable : (eng) *woman ~ women, child ~ children, etc.*

39 Mis à part le pluriel des mots français en -al, etc.

L'adjectif en vieil-anglais avait deux flexions (Mossé, 1947 : 21). Aujourd'hui, il n'en a plus qu'une.

Dans les langues romanes, en revanche, la forme de l'adjectif dépend du genre et du nombre du nom auquel il se rapporte, y compris lorsqu'il est attribut ; il a donc quatre formes possibles. Cependant, certains adjectifs sont épïcènes, surtout des adjectifs en *-e*.

Le comparatif et le superlatif de supériorité des adjectifs des langues germaniques ont une formation synthétique, par la suffixation de *-er / -st*<sup>40</sup>. Mais l'anglais connaît un double système. Ayant subi l'influence du français (Crépin : 40), il recourt aussi à une formation périphrastique, avec les adverbes *more than / the most* antéposés aux adjectifs de plus de deux syllabes, la plupart du temps. Ce dernier point le rapproche des langues romanes en ce qu'elles utilisent aussi une corrélation adverbiale pour le comparatif : (cat) *més que*, (spa) *más que*, (fra) *plus que*, (ita) *più che*, (por) *mais que*, (ron) *mai decât*. Pour le superlatif, on retrouve les formes du comparatif : (cat) *el més* ; (spa) *el más* ; (fra) *le plus* ; (ita) *il più* ; (por) *o mais* ; (ron) *cel mai*.

Dans les langues germaniques et les langues romanes, il existe peu de formes très irrégulières, et certaines se retrouvent dans des mots sémantiquement équivalents, comme (fra) *bon, meilleur, le meilleur* ; (ita) *buono, meglio, il buonissimo / il ottimo* ; (eng) *good, better, the best* ; (deu) *gut, besser, das beste*, etc.

Les adjectifs des langues germaniques se distinguent des adjectifs des langues romanes en ce qu'ils sont invariables à la forme attributive.

Enfin, l'anglais et les langues romanes dérivent de leurs adjectifs des adverbes de manière au moyen de suffixes.

### **1.5.4.3 Analogies et dissemblances morphosyntaxiques des verbes de l'anglais et des langues romanes**

Les verbes des langues romanes et des langues germaniques connaissent une conjugaison par des flexions adjointes à la fin du verbe, ainsi que des tournures périphrastiques.

L'indoeuropéen n'avait pas de temps, mais des aspects. L'alternance vocalique ou *apophonie*, autrement dit l'utilisation de plusieurs voyelles à partir d'un même radical, permettait de

---

40 L'allemand et le néerlandais connaissent un comparatif relatif : (deu) *eine ältere Dame* ; (nld) *een ouder man* (une dame d'un certain âge, un homme d'un certain âge).

distinguer trois aspects : le présent, le parfait et l'aoriste. Ce système a évolué dans les langues romanes et germaniques vers un système de temps et de modes.

Dans les langues germaniques, on compte deux types de verbes, les verbes forts et les verbes faibles. La différenciation morphologique est encore en partie perceptible aujourd'hui dans les verbes forts, qui peuvent être répartis en six classes, selon la classification de Sweet (1898), en fonction de l'alternance de voyelles, à l'instar des verbes (eng) *ring, rang, rung* ; (deu) *finden, fand, gefunden* ; (dan) *binde, bandt, bundet*, qui appartiennent à la classe 3, la classe *bind*.

Dans les langues romanes, certains verbes vérifient également une alternance vocalique dans leur radical, avec deux ou trois voyelles différentes dans les formes construites sur le même radical : il en est ainsi par exemple de (fra) (je) *fais, (je) fis, (je) ferai* ; (ita) *facio, feci* ; (spa) *hago, hice, hecho* ; (por) *faço, fiz, feito*.

Certains apprenants de langues romanes, qui connaissent une forme d'alternance vocalique dans leur langue, peuvent reconnaître ce phénomène. Il en est ainsi :

- des locuteurs de langues finno-ougriennes et turciques (estonien, finnois, hongrois, turc, pour ne citer que des langues parlées en Europe), qui appliquent l'harmonie vocalique<sup>41</sup> ;
- des locuteurs d'arabe, dont la déclinaison des noms consiste en une alternance de voyelles : *-u* au nominatif, *-a* à l'accusatif et *-i* au génitif.

Des langues de l'Asie de l'Est – chinois, coréen, japonais et vietnamien – ne connaissent pas l'alternance vocalique, leurs locuteurs pourront néanmoins reconnaître ce phénomène du fait qu'ils ont appris les verbes irréguliers anglais.

Les verbes faibles (Grimm, 1822 : 558) des langues germaniques forment leur prétérit et leur participe passé<sup>42</sup> par la suffixation d'une dentale *-t / -d*. Le paradigme du verbe peut donc se limiter à quelques formes : des 14 formes du vieil-anglais (Mossé : 21), il en subsiste quatre en anglais aujourd'hui : *walk, walks, walking, walked*. Ce sont les verbes faibles qui sont productifs aujourd'hui.

---

41 Ceci dit, elle procède d'une autre logique : les affixes verbaux varient en fonction du point d'articulation de la voyelle du verbe.

42 Hormis en allemand et en néerlandais, où le participe passé de la plupart des verbes se forme aussi avec le préfixe *-ge* (ce trait a disparu du vieil-anglais au XV<sup>e</sup>).

En latin, quatre classes regroupaient les verbes. Dans les langues romanes modernes, il en reste trois<sup>43</sup>, dont deux sont productives<sup>44</sup> (Tasmowski & Reinheimer, 2003 : 22).

La notion de temps et de mode est délicate à définir car la morphologie emprunte à la fois à la dérivation et à la combinaison, et la sémantique implique des valeurs temporelles et aspectuelles. Très schématiquement, les langues romanes présentent de multiples formes verbales, selon la personne, le temps et le mode, à la différence de l'anglais. La maîtrise de la conjugaison des verbes des langues romanes est pour cette raison un processus couteux. Là encore, certains locuteurs connaissent des formes multiples dans leur L1 : les slavophones, les arabophones, les locuteurs de langues turciques, etc. Au contraire, les apprenants qui ont pour L1 une des langues de l'Asie de l'Est ignorent l'accord sujet-verbe et les multiples temps.

En revanche, les langues germaniques et les langues romanes convergent quant au recours à des formes périphrastiques. Alors que le germanique commun se contentait de trois temps, l'anglais a développé dès le vieil-anglais des temps périphrastiques avec des auxiliaires suivis d'un participe passé (Mossé, *ibid.* : 22). Les langues romanes ont développé ce même trait dès le bas-latin<sup>45</sup>.

Aujourd'hui, ce phénomène est étendu en anglais et dans les langues romanes, qui construisent de nombreuses formes périphrastiques verbales<sup>46</sup> avec des valeurs temporelles et/ou aspectuelles, par exemple :

- l'aspect terminatif : (fra) *venir de*, (spa) et (por) *acabar de* ;
- le futur proche ou immédiat : (fra) *aller*, (ita) *stare per*, (spa) *ir a*, (eng) *be going to* ;
- le présent progressif : (fra) *être en train de*, (ita) *stare* + gérondif, (spa) *estar* + gérondif, (spa) *andar* + gérondif, (por) *estar a*, (ron) *a fi pe cale să*, (eng) *be doing* ;
- l'obligation : (spa) *tener que*, (spa) *haber de*, (por) *ter que*, (por) *ter de*, (fra) *devoir*, *il faut*, *avoir à*, (ita) *avere da*, (ron) *trebuie să* ; (eng) *must*, *have to* ;
- la possibilité : (fra) *pouvoir*, (ita) *potere*, (spa) (por) *poder*, (ron) *a putea*, (eng) *can* ;
- la probabilité : (fra) *devoir*, (ita) *dovere*, (spa) *deber de*, (ron) *trebuie să*, (eng) *could* ;

---

43 Les quatre subsistent en roumain.

44 La classe en *-a / -er / -as* et la classe en *-is*.

45 Par exemple la conjugaison du futur.

46 La terminologie de ce trait varie selon des critères sémantiques, syntaxiques, etc., nous ne distinguons pas ici ce que d'aucuns appellent *auxiliaire*, *semi-auxiliaire*, (spa) *forma verbal atemporal*, *verboide*, etc.

- l'habitude ou la répétition : (spa) *soler* + infinitif, (ita) *venire* + gérondif, (por) *ter* + participe passé, (eng) *would*, (eng) *used to* ;
- l'aspect inchoatif ou ingressif : (spa) *echarse a*, (spa) *ponerse a*, (fra) *se mettre à*, (ita) *mettersi a*, (por) *meter-se a*, (ron) *începe să*, (eng) *start to*, (eng) *begin to*.

Il n'est pas indispensable d'entrer ici dans les nuances sémantiques entre (eng) *be going to* et *be about to*, qui exprime mieux l'imminence, etc. Ce n'est en effet pas notre propos, et nous illustrons seulement le fait qu'un recours au procédé périphrastique est possible et répandu dans ces langues. De même, d'autres tournures périphrastiques existent ; nous ne cherchons pas ici l'exhaustivité.

C'est également une forme périphrastique qui sert à construire le passif, dans les langues germaniques et dans les langues romanes, avec un ou deux auxiliaire(s) suivi d'un participe passé. Le tableau 1 l'illustre où l'on voit que les langues romanes et l'anglais utilisent deux auxiliaires quand l'allemand et le néerlandais utilisent le verbe équivalent à *devenir*.

Langue	Phrase au passif
français	Il a été vacciné.
italien	È stato vaccinato.
espagnol	Ha sido vacunado.
portugais	Tenho sido vacinado
roumain	A fost vaccinat.
anglais	He's been vaccinated.
allemand	Er wurde geimpft.
néerlandais	Hij wordt gevaccineerd.

Tableau 1 : Formation périphrastique du passif dans les langues romanes et germaniques

Par ailleurs, l'anglais présente la particularité par rapport aux langues romanes d'avoir développé, particulièrement depuis le XIX<sup>e</sup> siècle, des verbes suivis d'une postposition (*put in*, *set out*, etc.)<sup>47</sup>.

<sup>47</sup> Ils existent également en italien, par exemple *buttare via* (jeter), *andare su* (augmenter), etc.

#### **1.5.4.4 Analogies et dissemblances morphosyntaxiques des déterminants**

Les déterminants et les pronoms sont utilisés extrêmement fréquemment, ils ont subi pour cette raison une forte érosion. Des analogies sont encore cependant visibles.

Au singulier, les déterminants des langues romanes expriment le genre grammatical du nom : (spa) *la, el, una, un, esta, este*, etc. ; (fra) *la, le, une, un, cette, ce/cet*, etc. ; (ita) *la, il, una, un, questa, questo*, etc. ; (por) *a, o, uma, um, esta, este*<sup>48</sup>, etc.

Au pluriel, les déterminants de l'espagnol, de l'italien et du portugais marquent le genre : (spa) *las, los, algunas, algunos*, etc. ; (ita) *le, li/gli, alcune, alcuni*, etc. ; (por) *as, os, algumas, alguns*, etc.

Le français est à part, dans la mesure où plusieurs déterminants ont une forme épïcène : *les, ces, des, plusieurs, quelques*, etc. (mais *certaines, certains*, etc.). Par ses formes plurielles épïcènes, le français ressemble à l'anglais qui ne distingue jamais le féminin du masculin, ni au singulier, ni au pluriel<sup>49</sup>.

Le français et l'italien ont un article indéfini pluriel à la différence de l'espagnol et du portugais. Ces derniers ont un fonctionnement identique à l'anglais, ils n'utilisent pas d'article :

(fra) *Il y a des médecins qui...* ; (ita) *Ci sono dei medici che...* ; (spa) *Hay médicos que...* ; (por) *Há médicos que...* ; (eng) *There are doctors who...*

#### **1.5.4.5 Analogies et dissemblances morphosyntaxiques des pronoms**

Il existe des pronoms masculins, féminins et neutres dans les langues romanes et en anglais. Les pronoms personnels des langues romanes ont tendance à avoir une forme commune au neutre et au masculin :

(spa) *lo sé ; lo conozco* ; (fra) *je le sais ; je le connais* ; (ita) *lo so ; lo conosco* ; (por) *o sei ; o conheço*.

---

48 On notera la forme épïcène des déterminants élidés en français et en italien : *l'abdomen* (masculin), *l'admission* (féminin), *l'acido* (masculin) et *l'attività* (féminin). L'élision est plus répandue en italien : *un'altra* (une autre), etc.

49 Il en est de même en néerlandais pour l'article, identique au singulier et au pluriel.



Certains pronoms personnels sont épïcènes : en français, les pronoms sujets de la première et de la deuxième personne *je, tu, nous, vous*, tous les pronoms compléments *me, te, se, lui, nous, vous, se, leur*. Il en est de même du pronom complément élidé *l'*.

Les langues romanes connaissent le tutoiement et le vouvoiement. La distinction T-V n'existe plus en anglais, qui a éliminé la deuxième personne du singulier<sup>50</sup> et n'utilise plus que *you*, la deuxième personne du pluriel.

Nous ne traiterons pas en détail de la place différente qu'occupent les pronoms compléments : en anglais, ces pronoms sont en enclise, et en français, « les pronoms clitics sont en proclise devant le verbe conjugué et après la négation » (Marchello-Nizia, 2011 : 581).

### **1.5.5 Récapitulatif partiel des analogies entre l'anglais et les langues romanes**

Si nous récapitulons les analogies que nous avons vues précédemment, nous obtenons le Tableau 2 : Principales analogies / dissemblances de l'anglais, du français et des langues romanes, p.65. Encore une fois, les analogies entre ces langues n'y apparaissent pas de manière exhaustive<sup>51</sup>. Ce sont les plus évidentes. On pourrait les prendre en compte pour enseigner une langue tierce romane en adoptant une approche contrastive à partir de l'anglais. Nous y revenons dans la section Perspectives : vers un dispositif didactique plurilingue d'enseignement du FOS sur la base d'analogies interlinguistiques, p.247.

Dans ce tableau comparatif, une croix dans une case indique que la langue / l'ensemble de langues possède(nt) le trait, une double croix indique que le trait est une caractéristique remarquable de la langue / l'ensemble de langues, et une croix entre parenthèses indique que le trait existe sans être le cas général. Nous distinguons le français des autres langues romanes parce qu'il s'agit de notre L3 cible.

Nous regroupons au début du tableau les traits sur lesquels les cinq langues de notre étude sont analogues. Les lignes de la fin du tableau mettent en relief les différences actuelles entre l'anglais et les langues romanes. Sur ces dissemblances, il doit être possible de faire porter un enseignement des langues romanes destiné à des apprenants qui ont l'anglais dans leur répertoire.

---

50 C'est ce qu'exprimait le pronom *thou.*, devenu désuet à la fin du XVI<sup>e</sup> (Walter, 2001 : 84).

51 La subordonnée, relative ou circonstancielle, existe dans toutes ces langues, mais nous ne l'avons pas abordée.

	anglais	français	autres langues romanes
Syntaxe SVO	x	x	x
Syntaxe attributive	x	x	x
Article indéfini singulier	x	x	x
Recours à des prépositions	x	x	x
Dérivation lexicale par construction	x	x	x
Dérivation lexicale par conversion	xx	x	x
Dérivation lexicale par composition	x	x	x
Comparatif / superlatif périphrastique	x	x	x
Adverbes dérivables d'un adjectif	x	x	x
Conjugaison des verbes	x	xx	xx
Alternance vocalique dans la conjugaison	x	x	x
Temps périphrastiques	x	x	x
Autres formes périphrastiques	x	x	x
Genres des pronoms à la 3 <sup>e</sup> personne du singulier	x	x	x
Pronoms épïcènes 1 <sup>ère</sup> et 2 <sup>e</sup> personne	x	x	x
Déterminant antéposé au nom	x	x	(x)
Pluriel des noms en -s	x	x	(x)
Existence de déterminants pluriels épïcènes	x	x	
Genres des pronoms de la 3 <sup>e</sup> personne du pluriel			x
Adjectif antéposé au nom	x	(x)	(x)
Genres des noms		x	x
Accord des adjectifs		x	x
Genres des déterminants		x	x
Adjectif postposé au nom		x	x
Verbes à postposition(s)	x		

Tableau 2 : Principales analogies / dissemblances de l'anglais, du français et des langues romanes

On peut dire que l'anglais, le français et les langues romanes comptent de nombreuses analogies pour des traits essentiels de la syntaxe, de la morphologie lexicale et de la morphosyntaxe. En outre, le tableau fait apparaître la place intermédiaire du français, qui partage certains traits avec l'anglais et certains autres avec ses consœurs romanes.

### **1.5.6 Notre définition de l'analogie lexicale et de l'interlexique anglais-espagnol-français-italien-portugais**

Dans la présente recherche, nous comprenons l'analogie lexicale comme un double rapport : un rapport d'identité de sens et une similarité formelle. La proportion est continue en ce que l'équation implique trois termes : un mot anglais, quatre équivalents de traduction en espagnol, français, italien et portugais, et une acception :

- 1) le mot anglais est l'un des 2000 mots fréquents du corpus scientifique anglais de ScienText ;
- 2) les équivalents dans les quatre langues romanes traduisent l'un des 2000 mots anglais fréquents dont ils sont similaires formellement ;
- 3) l'acception est l'une des 2208 acceptions fréquentes des 2000 mots anglais fréquents.

L'interlexique est l'ensemble des séries de cinq mots qui sont des représentants d'un intersémème et d'une interforme, autrement dit qui ont à la fois une acception commune et une forme similaire. La similarité des mots anglais fréquents de quatre lettres et plus du corpus ScienText avec leurs équivalents de traduction en espagnol, français, italien et portugais est mesurée automatiquement (cf. 5.3.3 Mesure de similarité par le script CompareCols.csv.pl, p.189). Elle est calculée manuellement pour les mots de moins de 4 lettres (cf. Obtention des termes de moins de 4 lettres dans chaque langue, p.193).

La ressemblance formelle est au cœur de notre recherche lexicologique car nous misons sur la faculté de l'apprenant à reconnaître dans un mot du français, la L3 cible romane, un mot qu'il connaît en anglais. L'anglais est la langue passerelle vers les langues romanes grâce à des origines latines communes (cf. Stefenelli,1992).

Comme la forme d'un mot d'une langue romane est comparée à celle de l'anglais, on n'observe pas de symétrie de la conformité (Lepage 2003 : 112) : ce n'est pas parce qu'un mot espagnol est analogue à un mot anglais qu'il est également analogue à un mot français, italien ou

portugais. Les mots des quatre langues romanes de l'interlexique ont la même acception mais leur forme n'est pas nécessairement similaire, même si chacun est similaire à l'anglais<sup>52</sup>. Nous avons pour cette raison établi quatre lexiques bilingues anglais-langue romane, qui rassemblent les paires d'équivalents de traduction dans les quatre langues romanes de cette étude (cf. Analogie par couples de langues : quatre lexiques bilingues anglais-langue romane, p.238).

---

52 En voici un exemple : (spa) *mayor* et (fra) *majeur* ne sont pas analogues parce que le rapport de la sous-chaine commune (m a r, trois caractères) vis-à-vis de celle de la plus courte des deux sous-chaines (*mayor*, cinq caractères) est inférieur à 0,7.

## 1.6 Conclusion du chapitre 1

Ce chapitre a montré comment le concept d'analogie s'applique en linguistique, et plus particulièrement en lexicologie. Tout d'abord, nous avons distingué certains concepts proches qui contiennent l'idée de similitude : similarité, identité et analogie. Alors qu'on pourrait penser que les mots et expressions qui réfèrent à la similarité et à l'identité se répartissent dans la langue courante dans un continuum allant de l'absolument distinct à l'identique, la signification de certains d'entre eux est très large et traduit une similarité qui peut aller jusqu'à l'identité, nourrissant ainsi une ambiguïté certaine. Les mots mêmes qui désignent les concepts de similarité et d'analogie sont parfois utilisés l'un pour l'autre.

Nous avons mentionné les rapports de synonymie et d'homonymie, qui correspondent à des analogies intralinguistiques. Nous avons décidé par ailleurs de parler d'analogies interlinguistiques lorsque des mots de deux langues ont une forme similaire pour une même acception.

Nous avons examiné des appellations qui désignent des mots avec une certaine ressemblance de forme et/ou de sens : cognats, cognates, congénères et mots transparents. Étant donné leur définition imprécise ou très englobante, nous évitons ces appellatifs et utilisons plutôt la désignation de *mots analogues*, qui ne contient pas l'idée que ces mots seraient apparentés. Nous estimons que l'enseignement des langues pourrait, plutôt que d'inquiéter les apprenants sur des interférences linguistiques, mettre l'accent sur les similitudes et les analogies, qu'elles se justifient en diachronie ou qu'elles soient fortuites.

Nous avons également passé rapidement en revue des outils de mesure de similarité. Les données de notre interlexique sont calculées en faisant le rapport de la sous-chaine commune maximale (SCM) vis-à-vis de la plus courte des deux sous-chaines.

Puis, dans la continuité d'EuroCom et de Caure, nous avons expliqué ce que nous entendons par intersystème, interlexique, interlexème, interforme et intersémème. L'interlexique en cinq langues que nous avons élaboré repose sur la similarité des 2208 équivalents de traduction identifiées exprimant les acceptions fréquentes des 2000 mots anglais fréquents d'un corpus de textes scientifiques de spécialité. Nous avons rattaché la notion d'interforme à celle de faux amis et nous avons essayé de dédramatiser cette question pour proposer d'exploiter le potentiel que représentent ce que nous appelons les sosies dans l'étude de la polysémie, et surtout de la néologie.

Enfin, nous avons adopté un point de vue moins restrictif en élargissant la réflexion à d'autres catégories d'analogies intra- et interlinguistiques, nous avons étudié celles qui s'appliquent dans la comparaison de l'anglais, du français et des langues romanes et les avons récapitulées avant de préciser en quoi l'analogie s'applique dans les séries de mots de l'interlexique. L'initiation à une langue pourrait porter en priorité sur le contraste (morpho)syntaxique des systèmes linguistiques que connaît l'apprenant et du système de la langue cible.

Les nombreuses analogies que nous avons notées entre l'anglais, le français, et plus largement les langues germaniques et les langues romanes, sont le fruit de l'évolution parallèle de ces langues. C'est l'objet du chapitre 2.

# Chapitre 2 - Raisons historiques des analogies entre l'anglais et les langues romanes

## Sommaire du chapitre 2

2.1 Introduction du chapitre 2.....	70
2.2 Origines génétiques communes des langues germaniques et romanes : des langues indoeuropéennes.....	72
2.2.1 Utilisation actuelle des langues indoeuropéennes.....	72
2.2.2 De la reconstruction phonéticohistorique des langues indoeuropéennes : à la recherche d'analogies.....	73
2.2.3 Hypothèses sur la diffusion des langues indoeuropéennes.....	74
2.2.4 Mise en question des modèles généalogiques.....	77
2.3 Des analogies interlinguistiques liées aux contacts des langues en Europe de l'Ouest.....	81
2.3.1 Concepts et outils autour de la notion de contacts de langues.....	81
2.3.2 Le latin comme dénominateur commun.....	86
2.3.3 Influences des langues romanes.....	95
2.3.4 Influences des langues germaniques.....	97
2.4 Conclusion du chapitre 2.....	102

## 2.1 Introduction du chapitre 2

Afin de comprendre comment l'anglais, le français (et d'autres langues) en sont venus à partager des analogies interlinguistiques, nous adoptons dans ce chapitre un point de vue diachronique. Les langues germaniques, auxquelles appartient l'anglais, et les langues romanes, auxquelles appartient le français, sont classées dans la famille des langues indoeuropéennes<sup>53</sup>. De plus, comme l'anglais et le français sont parlés dans des zones géographiques contigües, la situation de contacts qui en résulte est en soi un motif premier d'influences interlinguistiques.

La section 2.2, p.72, expose des raisons d'ordre génétique qui expliquent en partie la situation géolinguistique actuelle. Nous examinerons quatre points : l'utilisation actuelle des langues indoeuropéennes, en 2.2.1, p.72, la méthode de reconstitution phonéticohistorique qui a

53 Nous soudons *indo* et *européen* conformément à la loi du 6 décembre 1990 sur l'orthographe rectifiée : [http://www.academie-francaise.fr/sites/academie-francaise.fr/files/rectifications\\_1990.pdf](http://www.academie-francaise.fr/sites/academie-francaise.fr/files/rectifications_1990.pdf) (cf.aussi le site <http://www.renouvo.org>).

permis de classer ces langues, en 2.2.2, p.73, les hypothèses sur leur diffusion, en 2.2.3, p.74, et enfin la mise en question des modèles généalogiques, en 2.2.4, p.77.

Nous récapitulons dans la section 2.3, p.81 les principaux motifs d'influences mutuelles de l'anglais et du français, et plus largement entre d'autres langues de l'Europe de l'Ouest, au gré des politiques linguistiques et des échanges culturels.

Nous pouvons considérer que le latin est un dénominateur commun des langues parlées dans l'Ouest de l'Europe, qu'elles soient romanes ou germaniques, de l'Antiquité jusqu'à nos jours. Le lexique des langues germaniques se compose de nombreux mots que reconnaissent des apprenants de FLE (et de langues romanes, en général) parce que la langue latine a exercé une influence considérable dans l'ouest de l'Europe, mais aussi parce que le français a modifié l'anglais en profondeur, tout en s'enrichissant lui-même au contact de cette langue.

La section 2.3.1, p.81 présente la notion de contacts de langues. L'influence omniprésente du latin est traitée dans la section 2.3.2, p.86, et les influences mutuelles des langues romanes et des langues germaniques dans les sections 2.3.3, p.95, et 2.3.4, p.97.



## 2.2 Origines génétiques communes des langues germaniques et romanes : des langues indoeuropéennes

### 2.2.1 Utilisation actuelle des langues indoeuropéennes

Les usagers de langues indoeuropéennes représenteraient près de la moitié de la population mondiale (Walter, 2014 : 18). Deux milliards de personnes en utiliseraient une comme L1 sur les cinq continents. La fiabilité des estimations se pose toutefois car les enjeux symboliques sont très forts dans la bataille des langues (Cassen, 2008). Aussi est-il délicat de représenter où les langues indoeuropéennes sont effectivement utilisées.

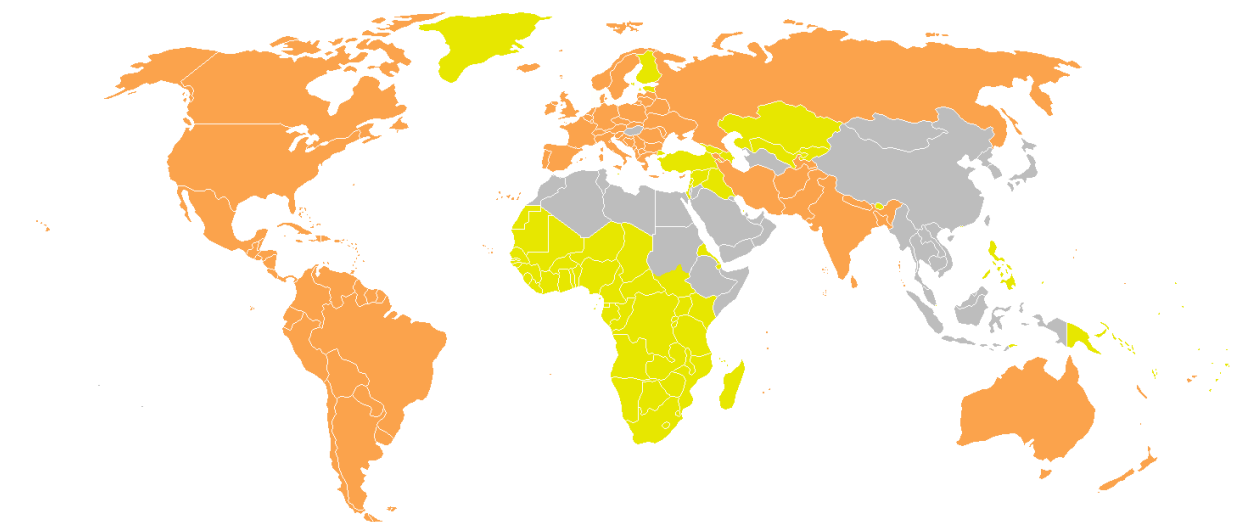


Image 2 : Répartition des langues indoeuropéennes dans le monde, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=189307>

La carte de l'image 2 prend en compte à la fois l'utilisation spontanée des langues indoeuropéennes et leur statut de langue (co-)officielle. Selon ce document de *Wikipédia*, on trouve une « majorité de locuteurs d'une langue indo-européenne » en Amérique, dans la partie septentrionale de l'Eurasie, en Iran et Afghanistan, dans le sous-continent indien et le sud de l'Océanie (en orange), une « minorité de locuteurs d'une langue indo-européenne ayant un statut officiel » dans des poches en Eurasie et presque toute l'Afrique (en jaune), hormis le

nord, et une « minorité significative de locuteurs d'une langue indo-européenne sans statut officiel » dans le reste du monde, autrement dit le nord de l'Afrique, l'Asie de l'Est et du Sud-Est (en gris).

Quelles que soient les faiblesses de cette carte<sup>54</sup>, on observe que les langues indoeuropéennes sont utilisées dans une majorité des États du monde, et que la plupart des États européens en ont une pour langue officielle. Les langues dont il est surtout question ci-après, les langues romanes et les langues germaniques, le sont dans l'ouest de l'Europe, et au nord-est du Danube.

### **2.2.2 De la reconstruction phonéticohistorique des langues indoeuropéennes : à la recherche d'analogies**

Les témoignages écrits les plus anciens sur les langues indoeuropéennes datent de la première moitié du II<sup>e</sup> millénaire avant notre ère (Lamberterie, 2011 : 399). La connaissance des états antérieurs est donc indirecte.

À partir du XVI<sup>e</sup> siècle, des érudits ont remarqué des analogies écrites entre les langues anciennes. Jones (1799) a affirmé que le sanscrit, le latin, le grec et le persan, du fait de leurs fortes affinités, « ont dû jaillir d'une source commune ». Cependant, jamais il n'en a fourni aucune démonstration philologique (Demoule, 2014 : 23).

Bopp (1816) le premier a démontré la similitude de certaines langues en comparant leur système de flexions respectif : en établissant des *correspondances*, il a fondé la linguistique historique et la linguistique comparée. Bergounioux (2010 : 14) définit la correspondance comme une « corrélation établie entre des mots appartenant à des langues de la même famille linguistique au moyen de lois phonétiques qui justifient leur différence de forme et leur rapprochement ».

Par des comparaisons lexicales et morphologiques (Jucquois, 2009), la linguistique génétique a énoncé des règles de reconstruction phonéticohistorique.

Concrètement, les mots de sens proche dans plusieurs langues ont été comparés. Des régularités ont été constatées, dont ont été déduits des changements successifs du point d'articulation de certains phonèmes. Ainsi, dans certains rameaux des langues

---

54 Il faudrait afficher clairement les critères suivis dans l'élaboration de cette carte et expliciter « minorité de locuteurs » et « minorité significative de locuteurs ». Par ailleurs, l'Angola ne devrait pas être en jaune car le portugais est la L1 majoritaire.

indoeuropéennes, des mots ont à l'initiale la dorsopalatale /k/, dans d'autres rameaux, leur équivalent sémantique a la glottale /h/, et dans d'autres encore la prédorsoalvéolaire /s/<sup>55</sup>.

L'ensemble de ces mots forme une série, et de multiples séries permettent de déduire une concordance. D'un faisceau de concordances entre plusieurs langues, il a été déduit qu'elles étaient apparentées génétiquement, autrement dit, qu'elles provenaient d'étymons appartenant à une même protolangue. Cette différenciation a eu lieu pendant plusieurs millénaires. Lepage (2004 : 54) note :

« On peut dire, avec le recul, que la linguistique de la fin du XIXe siècle, c'est-à-dire l'ensemble des travaux des linguistes comparatistes et des Néogrammairiens (*Junggrammatiker* en allemand), apparaît comme traversée par la recherche des invariants et des variantes : comment ramener ce qui change entre différentes langues à des prototypes communs historiques ? ».

Cependant, constater des correspondances dans deux langues ne suffit pas pour en conclure que l'une provient de l'autre. Aussi les recherches sur la génétique des langues font-elles appel à d'autres disciplines scientifiques.

## 2.2.3 Hypothèses sur la diffusion des langues indoeuropéennes

### 2.2.3.1 *D'une langue supposée à un peuple supposé*

Plusieurs homonymes sont employés en français et en anglais, nous nous proposons de préciser leur sens respectif.

Les *langues indoeuropéennes*, en anglais *Indo-European languages*<sup>56</sup> (les deux mots sont ici des adjectifs), sont des langues qui partagent certains traits - les correspondances dont il a été question dans la section précédente. L'hyperonyme *langues indoeuropéennes*<sup>57</sup> apparaissait *a priori* comme une appellation logique puisque ces langues étaient parlées<sup>58</sup> en Eurasie, de l'ouest de l'Europe jusqu'à l'Inde<sup>59</sup>.

---

55 Nous pouvons illustrer ce phénomène au moyen de l'exemple classique du français *cœur* et de ses équivalents : en latin, grec et dans les langues celtiques, l'initiale est en /k/ (grec *kardia* ; breton *kalon* ; latin *cor*) ; dans les langues germaniques et en sanscrit, l'initiale est en /h/ (anglais *heart* ; suédois *hjärta* ; sanscrit *hrd-*, *hrdaya-*) ; dans les langues balto-slaves, l'initiale est en /s/ (letton *sirds* ; russe *serdce* ; serbe *srce*).

56 Il existe aussi la variante graphique *Indoeuropean*.

57 Le nom *Indo-European* a été forgé par Thomas Young en 1813.

58 Avant la colonisation européenne des XV<sup>e</sup> – XX<sup>e</sup> siècles.

59 On sait depuis les travaux sur le tokharien que cette zone va jusqu'au bassin du Tarim, à l'ouest de la Chine.

Les linguistes allemands du XIX<sup>e</sup> siècle ont fait l'hypothèse que ces langues qui présentent des analogies provenaient d'une protolangue originelle, l'*indo-européen*, (eng) *Indo-European* (ces deux mots sont des noms). On parle de protolangue puisqu'il n'en existe pas de traces écrites. Elle est également désignée sous le nom de *protoindo-européen*, (eng) *Proto-Indo-European*.

Certains ont imaginé que cette langue originelle supposée, (deu) *Ursprache*, avait été parlée par un seul peuple originel, *Urvolk*, dans un foyer originel, *Urheimat*. La quête de ce peuple supposé, les *Indo-européens* (*the Indo-Europeans*), a commencé. Bergounioux (2010 : 15-16) résume comme suit comment on est passé de la langue au peuple :

« Et de cette langue de synthèse est inféré un peuple, les « Indo-européens », dont les archéologues rêvent de découvrir le berceau. En quarante ans, l'attention s'est déplacée de propriétés morpho-phonologiques particulières, observables par des correspondances, à la restitution de la langue qui en serait l'ancêtre puis à la quête du peuple qui l'aurait parlée. »

Le foyer d'origine de ce peuple hypothétique demeure controversé.

### **2.2.3.2 Théories sur le foyer d'origine d'hypothétiques Protoindo-européens**

Plusieurs théories se fondent sur le présupposé qu'auraient existé les (Proto)Indo-européens, un groupe humain culturellement homogène qui aurait parlé une même langue, le protoindo-européen, et aurait provenu d'une même zone géographique.

Les hypothèses s'appuient sur les reconstructions linguistiques dont nous avons parlé, ainsi que sur la comparaison de témoignages culturels (pratiques funéraires, techniques, mythes, etc.) mettant en œuvre des approches interdisciplinaires : l'archéologie met au jour des restes matériels, les datent et les associent à une culture. Mais un mobilier funéraire tel qu'une poterie ne peut pas avoir de signification ethnique (Legendre, 2016 : 73). Tout au plus peut-on parler d'une aire culturelle où ce type de poterie s'est échangé.

Une théorie a longtemps dominé, celle que l'archéologue Gimbutas (1956) a formulée : l'hypothèse des Kourganes. Mallory (1997) a repris cette théorie, selon laquelle les Indo-européens auraient été des peuples guerriers nomades originaires des steppes de l'Ukraine

méridionale ou du sud de la Russie, et qu'ils auraient introduit la pratique de l'agriculture au Néolithique, il y a 5000 à 6000 ans.

Renfrew (1987), a contredit l'hypothèse des Kourganes et a formulé l'hypothèse anatolienne : les Indoeuropéens, dont la migration aurait commencé il y a 8000 ans, seraient originaires d'Anatolie, il auraient apporté l'agriculture et l'élevage, entraînant ainsi la néolithisation<sup>60</sup> de l'ouest de l'Europe. La thèse anatolienne attire actuellement de nombreux suffrages.

Outre ces deux théories, nous ne ferons qu'en mentionner deux autres, marquées par l'idéologie. La théorie nordique, adoptée par Kossinna (1902), a dominé en Allemagne après la première guerre mondiale. Elle plaçait le foyer des *Indogermains* dans le nord de l'Europe, et c'est « l'une des composantes idéologiques essentielles du nationalisme allemand et de son incarnation politique du moment, le nazisme » (Demoule, *ibid.* : 219). Enfin, selon la théorie de la continuité paléolithique (Alinei, 1996), des humains auraient déjà parlé des langues indoeuropéennes au paléolithique, et les Celtes auraient migré d'ouest en est. Les dernières recherches génétiques<sup>61</sup> infirment cette théorie.

### **2.2.3.3 L'apport de la génétique à la recherche sur les migrations en Eurasie**

À partir des années 1970, l'anthropologie biologique a incorporé des techniques et approches empruntées à la génétique. Cavalli-Sforza (2000 : 39), par exemple, a étudié le produit d'expression des gènes tel que le groupe sanguin des individus.

Bramanti (2013) retrace les migrations par l'étude de l'ADN mitochondrial<sup>62</sup> des chasseurs-cueilleurs du Mésolithique, qu'elle compare à celui des agriculteurs du Néolithique. L'énorme disparité génétique signifierait que ces derniers auraient migré, et par conséquent qu'ils auraient apporté avec eux l'agriculture. Cependant, certains variants d'ADN des chasseurs-cueilleurs présentent encore une fréquence non négligeable aujourd'hui dans certaines populations européennes. Par ailleurs, «chez les agriculteurs de la culture rubanée, la répartition des gènes diffère de celle des Européens actuels » (Berger, 2014 : 22). Enfin, il semblerait que les représentants d'une autre culture, la culture de la céramique cardiale, soient

---

60 Dès lors, l'agriculture et l'élevage complètent la cueillette et la chasse, et des poteries sont fabriquées pour cuire et conserver des aliments.

61 Analysant des gènes polymorphes, Haak (2015) montre que des mouvements de populations massifs ont eu lieu d'est en ouest jusqu'aux Iles Britanniques.

62 Il permet de reconstituer les lignées maternelles.

eux aussi issus d'une vague de peuplement provenant du Moyen Orient, ce qui complexifie les données sur l'origine des Européens actuels.

D'autres progrès sont attendus de la génétique car depuis 2008, des marqueurs spécifiques permettent de reconstituer d'éventuels liens de parenté dans un ensemble funéraire (Crubézy et al., 2010 : 66). En outre, le séquençage de l'ADN autosomal est désormais pratiqué sur des restants humains, ce qui augure une compréhension plus fine de la manière dont les descendants d'une même population ont migré en Eurasie au Néolithique.

Mais ces approches anthropologiques linéaires des migrations ne peuvent permettre de déduire quelle langue a parlé une population car cela supposerait que les représentants d'une même lignée ont parlé une même langue. Or des mariages mixtes ont nécessairement eu lieu, et dans ce processus, « des ressources sociales, économiques et culturelles s'entremêlent » (Dumănescu et al., 2014 : 4). Dans le cas où deux groupes parlaient une (variété de) langue distincte, quelle(s) (variétés(s) de) langue(s) les descendants de telles unions ont-ils continué à parler ? D'autre part, les situations de plurilinguisme ne semblent pas envisagées, alors qu'elles ont dû être fréquentes, comme elles sont extrêmement fréquentes aujourd'hui à l'échelle planétaire.

Parce qu'il s'appuie sur des régularités - les *correspondances* -, on pourrait s'attendre à ce que l'énorme travail de reconstruction phonéticohistorique des comparatistes occasionne moins de débats. Pourtant, les modèles de restitution des langues indoeuropéennes que la linguistique génétique a proposés sous la forme d'arborescences soulèvent d'autres critiques.

## **2.2.4 Mise en question des modèles généalogiques**

### ***2.2.4.1 Implications des représentations arborescentes***

La classification des langues indoeuropéennes est souvent représentée dans des sortes d'arbres qui se ramifient à mesure que les langues ont divergé. Cette représentation implique une conception généalogique - en témoigne une expression telle que *famille de langues*, que nous utilisons ici par la force de l'habitude ancrée dans notre discipline.

Le botaniste allemand darwinien Schleicher le premier (1861) a établi un *arbre généalogique* des langues à une époque de questionnement intense sur l'identité allemande, où était recherché un modèle alternatif à celui de la Bible et d'une langue originelle qui serait l'hébreu.

Ces représentations arborescentes ont subi l'influence de la recherche menée au XIX<sup>e</sup> siècle dans d'autres disciplines scientifiques, particulièrement la biologie. Elles risquent de suggérer qu'à partir d'un moment, les langues des différentes branches n'auraient plus entretenu de rapports les unes avec les autres, et en cela, elles peuvent induire en erreur.

Watkins (2000) distingue quinze branches : huit encore vivantes (baltoslave, germanique, celtique, italique, albanais, hellénique, arménien, indo-iranien) et sept éteintes (vénète, messapien, illyrien, thrace, phrygien, anatolien, tokharien). Il propose la représentation suivante en éventail, image 3<sup>63</sup>, où figurent de gauche à droite les branches d'ouest en est.

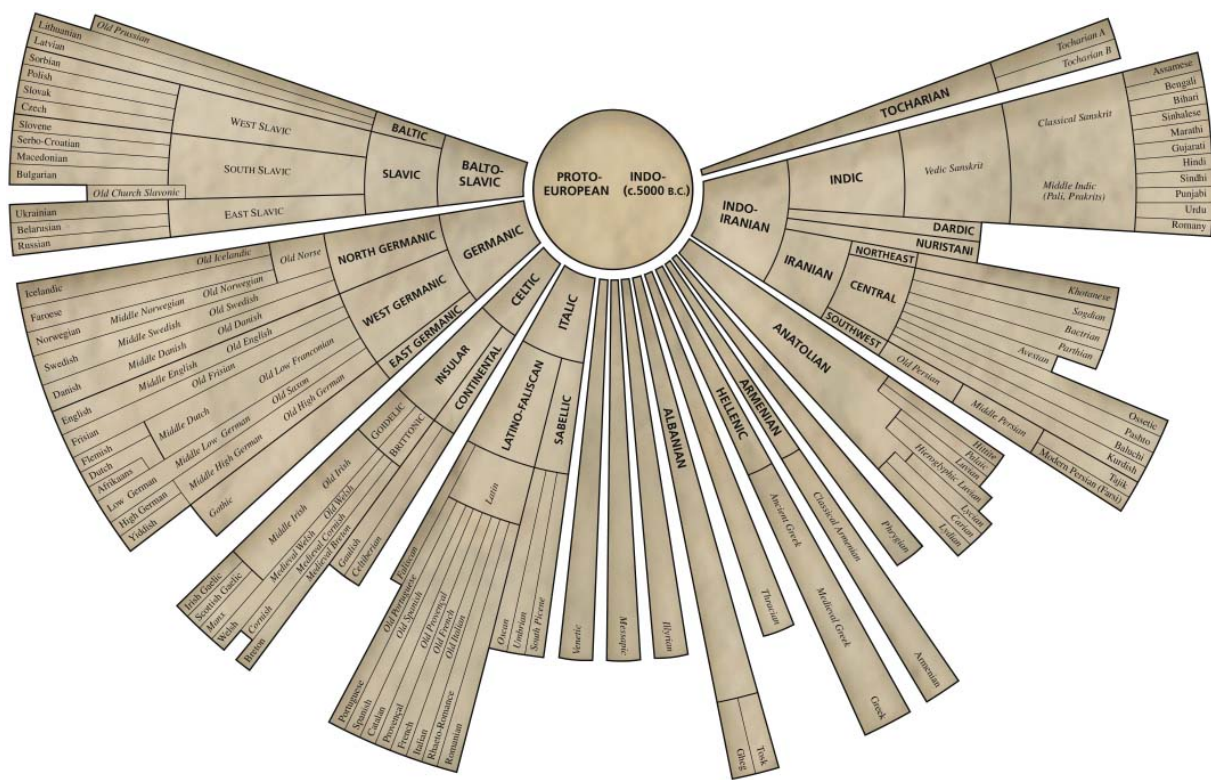


Image 3 : Représentation des langues indoeuropéennes de Watkins. Crédit : Houghton Mifflin Company

En outre, des approches statistiques ont été développées depuis le début du XXI<sup>e</sup> siècle à l'aide de méthodes comparables à celles utilisées pour l'analyse de séquences d'ADN. Ainsi, Gray et Atkinson (2003 : 435-438) recourent à la modélisation d'arbres de parenté des langues. Selon eux, les racines de la famille indoeuropéenne remonteraient à 7800 à 9800 ans.

63 Nous tirons ce document de l'*American Heritage Dictionary of the English Language* (Morris, 1981) dans sa version de 2000, <http://homes.chass.utoronto.ca/~cpercyl/hell/families/indoeuropean>.

Elle serait née en Anatolie et se serait propagée avec l'expansion des premiers agriculteurs entre 6000 et 7000 avant notre ère. Une bifurcation aurait eu lieu il y a 6500 ans, entre une branche baltoslave et une autre branche, laquelle se serait divisée il y a 6100 ans entre une branche celtique et une autre branche, qui se serait à son tour ramifiée pour donner il y a 5500 ans les deux branches qui nous intéressent dans cette thèse : l'italique et le germanique.

Cette nouvelle approche a été critiquée en ce qu'elle ne prend pas en compte les données archéologiques, ni la complexité du langage, mais pour Pakendorf (2014 : 47), « la méthode comparative n'offre pas d'évaluation objective de la qualité d'une reconstruction », à l'inverse des méthodes phylogéniques, qui « fournissent des mesures objectives de la qualité d'un regroupement, car elles calculent non pas un seul arbre généalogique, mais des milliers », et qui de plus permettraient de « dater les racines des regroupements de langues » (ibid. : 48).

Nous quittons ici les controverses sur les origines des langues, leur classification et leur représentation pour nous focaliser sur les langues romanes et des langues germaniques.

#### **2.2.4.2 Classification des langues romanes et germaniques**

La classification des langues romanes et germaniques fait davantage consensus. La dérivation<sup>64</sup> des langues romanes à partir du latin est contrôlable puisque nous disposons de textes depuis l'époque romaine ; celle des langues germaniques à partir du germanique commun<sup>65</sup> a été montrée par des études portant sur le *Codex Argenteus*<sup>66</sup>, écrit au VI<sup>e</sup> siècle en gotique, et sur des inscriptions runiques.

Pour Sergent (1995), les langues romanes<sup>67</sup> sont un groupe qui fait partie de la sous-branche italique de la branche italoceltique des langues indoeuropéennes. Le latin, dont proviennent les langues romanes, est une langue italique au même titre que l'osque et l'ombrien, toutes deux éteintes.

Partant de critères morphologiques, syntaxiques et sociolinguistiques, Tagliavini (1969) subdivise les langues romanes en quatre groupes : galloroman, ibéroroman, italo-roman,

---

64 Abandonnons le terme *filiation* qui réfère à la génétique.

65 Le germanique commun (ou protogermanique) est la désignation de la langue hypothétique de laquelle proviendraient toutes les langues germaniques.

66 Une version numérisée est consultable sur la plateforme Alvin : <https://frama.link/tkKZdqP8> ; le texte transcrit est sur Wikisource : [https://wikisource.org/wiki/Bible,\\_Gothic,\\_Ulfila#II](https://wikisource.org/wiki/Bible,_Gothic,_Ulfila#II)

67 On a dit aussi *langues latines*.



balcanoroman. Dans la présente étude, nous avons établi des données portant sur le français, une langue du groupe gallo-roman, l'espagnol et le portugais, deux langues du groupe ibéroroman, et l'italien, une langue du groupe italo-roman.

Pour ce qui est des langues germaniques, on distingue trois rameaux issus du germanique commun: le rameau ostique (éteint), le rameau nordique et le rameau westique. Le rameau ostique comprenait les langues parlées par les Goths, les Burgondes et les Vandales. Au rameau nordique appartiennent les langues germaniques parlées actuellement en Scandinavie telles que le danois, le suédois et le norvégien. Le rameau westique regroupe les langues germaniques parlées actuellement dans l'ouest de l'Europe telles que le néerlandais, le frison, l'allemand et l'anglais. Pour certains chercheurs (Nielsen, 1981 ; Lass, 1995), il fait sens de distinguer un état de langue intermédiaire, le germanique du nord-ouest, qui précéderait la différenciation entre le rameau westique et le rameau nordique. L'anglais est classé comme langue germanique occidentale de la branche anglofrisonne du rameau westique.

Les origines génétiques ne sauraient suffire à expliquer les analogies qui existent entre les langues romanes et les langues germaniques. Des raisons sociohistoriques liées aux contacts de langues motivent la présence de similitudes formelles et d'analogies entre l'anglais, le français et d'autres langues.

## 2.3 Des analogies interlinguistiques liées aux contacts des langues en Europe de l'Ouest

### 2.3.1 Concepts et outils autour de la notion de contacts de langues

#### 2.3.1.1 *Substrat, superstrat, adstrat*

Une représentation arborescente ne rend pas compte des traces qu'ont laissées d'autres langues, à la différence du concept d'influence interlinguistique. On parle aussi d'*interférence linguistique*, mais l'élément *interférence*, souvent connoté négativement en français, avec l'idée de perturbation, de parasitage, d'ingérence, de contrariété, semble induire qu'une langue subirait des influences néfastes de la part d'autres langues. Parler d'interférences, ce n'est pas alors simplement constater une « superposition de [...] mouvements vibratoires de même nature, de fréquence identique ou voisine [...] »<sup>68</sup>, tel que le définit le TLFi dans le domaine de la physique<sup>69</sup>. La dénomination en français exprime une certaine méfiance vis-à-vis d'autres langues, méfiance qui trouve des prolongements dans l'enseignement des langues. Nous utiliserons préférentiellement *influence interlinguistique*, soit le calque de l'anglais *cross-linguistic influence*.

Les influences interlinguistiques peuvent être analysées sous la forme de trois types de rapports : le substrat, le superstrat et l'adstrat. Précisons l'évolution du sens de ces mots que nous utilisons dans la section suivante.

Wartburg (1950 : 155) a emprunté à la géologie les termes de substrat et de superstrat et leur a conféré un sens en linguistique diachronique :

- « 1. Quand un peuple occupe un pays habité par un peuple qui parle une autre langue, ce pays devient bilingue pendant un certain nombre de générations.
2. Cette situation peut devenir durable en cas de fort écart culturel. Toutefois, dans bien des cas, dans la plupart de ceux que nous connaissons du moins, l'une des deux langues supplante l'autre au fil du temps.

---

68 Les langues pourraient-elles également être considérées comme des phénomènes mouvants dont des éléments (lexiques, etc.) se superposent ?

69 Cf. la définition de ce mot dans le TLFi.

3. Si la langue de ceux qui ont vaincu et immigré l'emporte, la langue de ceux qui vivaient là auparavant entre, vis-à-vis de celle-ci, dans un rapport de langue substrat. Dans le cas inverse, la langue des nouveaux venus devient dans son rapport à la langue qui l'emporte une langue superstrat. »<sup>70</sup>

Ce sens a cependant évolué par la suite. Bourcier (1978 : 15, note 16) met en avant la domination physique : *substrat* « s'emploie à propos d'une langue supplantée par une autre à la suite d'une colonisation, d'une conquête militaire, etc. ». Par ailleurs, Gadet et Ludwig (2015 : 26) le définissent comme « une langue qui a disparu sous la pression sociale d'une langue dominante importée, tout en laissant des traces dans la nouvelle langue dominante ». Dans cette dernière définition, la pression sociale implique dans la reconfiguration de la situation sociolinguistique des raisons autres qu'un écart culturel. Comme nous le voyons plus loin, section 2.3.2.2 (Les langues celtiques de l'Antiquité face au latin), p.87, le gaulois est un substrat par rapport au français, lui-même issu du latin vernaculaire importé par les conquérants romains.

Le francique, la langue des Francs, est devenu un superstrat de l'ancien français (cf. section 2.3.4.2 p.99). Il en est de même de l'anglo-normand face à l'anglais. Tous deux ont laissé une influence dans les habitudes articulatoires et le lexique des locuteurs. Nous abordons cet aspect dans les sections 2.3.3.1 (Influences normandes et françaises sur l'anglais), p.95 et 1.5.1.2 (Incidence phonéticohistorique de l'accent à l'initiale des langues germaniques), p.53.

En complément de substrat et de superstrat, Valkhoff (1932 : 17 et 22) a proposé le concept d'*adstrat* pour désigner une langue qui exerce par des contacts directs une influence sur une autre langue. Au cours du temps, l'*adstrat* devient ou bien un superstrat, ou bien un substrat (Sinner, 2014 : 243-244). L'*adstrat* latin de la période romaine en Grande-Bretagne est ainsi devenu substrat à l'époque anglo-saxonne, cf. 2.3.2.2 (Les langues celtiques de l'Antiquité face au latin), p.87.

Par la force des contacts directs qui existent entre les locuteurs, les influences interlinguistiques sont vives aux marges de la zone de répartition de deux (variétés de) langues. Le concept de continuum permet d'envisager la communication à la périphérie.

---

70 Notre traduction. Le texte de Wartburg est cité par Geckeler et Kattenbusch (1992 : 124).

### 2.3.1.2 *Le concept de continuum*

Une représentation des langues, qu'elle soit géographique ou arborescente, occulte les influences linguistiques des locuteurs des confins avec leurs voisins. Le concept de *continuum*, combiné à celui de *frontière linguistique*, peut aider à saisir le phénomène de l'intercompréhension entre deux langues ou dialectes.

L'expression frontière linguistique s'utilise pour distinguer des zones où sont parlées deux langues. Le concept de continuum, parce qu'il contient l'idée de glissement progressif, nous semble plus approprié pour exprimer les influences interlinguistiques mutuelles et fluctuantes que connaissent les locuteurs d'une (variété de) langue au voisinage de locuteurs d'une autre (variété de) langue.

À l'intérieur d'un diasystème (Coşeriu, 1970 : 32), il y a continuum dialectal lorsque les locuteurs d'un point A du continuum peuvent comprendre les locuteurs d'un point B - et vice versa -, que les locuteurs du point B peuvent comprendre des locuteurs d'un point C – et vice versa -, mais que les locuteurs du point A ne peuvent comprendre les locuteurs du point C – et vice versa -, lesquels peuvent cependant comprendre les locuteurs d'un point D, etc. (Genetti, 2014 : 12). Si on le formule autrement, on peut dire qu'il existe une intercompréhension mutuelle entre locuteurs de zones contigües. Il faudrait nuancer cette définition car elle exclut que le locuteur du point A puisse comprendre pour partie au moins le locuteur du point C.

Un outil, l'isoglosse (Meillet, 1908: 3) permet de représenter les divergences dialectales : Nicolai (2010 : 69) explique que l'isoglosse « sur la base de la différenciation d'un trait linguistique particulier – souvent phonétique – consiste à tracer des lignes imaginaires ayant valeur de frontières entre des zones linguistiques géographiquement définies ». Cette distinction gomme cependant l'aptitude humaine à voir des analogies : que des locuteurs ne partagent pas les mêmes pratiques n'exclut pas qu'ils se comprennent. C'est pour cela que nous préférons rapprocher l'isoglosse de la notion phonologique d'*allophonème*<sup>71</sup>.

Si nous prenons l'exemple du français, deux locuteurs se comprennent si l'un prononce [e] et l'autre [ɛ]. Cette différence de prononciation, dont ils ne sont pas nécessairement conscients, n'est pas réhibitoire car le contexte phrastique et syntaxique permet de distinguer les formes

---

71 L'adjectif *allophone*, employé en sociolinguistique, qualifie un locuteur qui a pour L1 une langue différente de la langue du territoire où elle vit. On parle ainsi en France de la scolarisation des EANA, élèves *allophones* nouvellement arrivés (<https://frama.link/KfgZPrcw>, et absolument, des *allophones*. Afin d'éviter une confusion avec la signification du terme en phonologie, nous préférons employer le nom *allophonème* et l'adjectif *allophonémique*.

éventuellement concurrentes (*tu distinguais, tu as distingué, c'est distingué, etc.*). En outre, c'est cette souplesse dont font preuve les interlocuteurs au quotidien qui explique que des réalisations altérées par un déficit de la représentation perceptive (Krifi-Papoz, 2010 : 115) peuvent demeurer compréhensibles – que l'on pense à des énoncés relevant de la dyslexie phonologique (Adron & Hourcq, 2011 : 28).

Pareillement, on pourrait considérer qu'une isoglosse phonologique matérialise sur une carte les zones où une même unité phonologique se réalise de deux façons, au moyen de ce qui serait un allophonème dans une zone et un allophonème dans l'autre zone, quand bien même il n'est pas coutumier de relier ces deux réalisations dans l'une et l'autre langue<sup>72</sup>.

### 2.3.1.3 Conflit entre la possibilité d'une hybridité et le sentiment linguistique

Tôt dans l'histoire de la linguistique, certains linguistes ont pris leurs distances avec une représentation arborescente de l'évolution des langues indoeuropéennes empruntée à la biologie. Schmidt (1872) a formulé la théorie des ondes<sup>73</sup>, selon laquelle les innovations linguistiques se répandraient à la manière de vagues, d'une langue à une autre, dans un continuum linguistique.

D'autre part, Schuchardt a étudié les phénomènes de mélanges linguistiques dans des zones de contacts. L'analyse de pidgins utilisés à la fin du XIX<sup>e</sup> siècle en Europe, par exemple l'italoslave, l'ont conduit à noter : « Il n'existe pas de langue qui soit tout à fait sans mélange »<sup>74</sup> (1884 : 6). Demoule (2014 : 159-161) montre que deux visions se sont opposées. À la mixité des langues de Schuchardt - où les langues sont le produit de métissages – s'opposaient des linguistes comme Meillet, qui croyaient en une *permanence indoeuropéenne*, pour lesquels il ne saurait être question de langues qui se mélangeraient à d'autres. Pour ce linguiste (1921 : 102), les locuteurs ne peuvent qu'ajouter à un système des procédés d'un autre. En outre, selon Meillet, « Ce qui compte, ce n'est pas de déterminer la proportion de tel ou tel élément, mais de savoir quelle langue ont cru et voulu parler ceux qui ont fait la transmission entre les deux dates considérées »<sup>75</sup> (ibid. 104). Cette conception refuse donc

---

72 Que l'on pense à l'isoglosse dat-das entre les dialectes germaniques : /t/ et /s/ ne sont des allophonèmes ni en anglais, ni en allemand.

73 *Wellentheorie*, est aussi traduit par *théorie des vagues* en français.

74 « Es gibt keine völlig ungemischte Sprache ».

75 Cité dans Demoule (2014 : 160-161).

l'hybridité, qui remettrait en question l'arbre des langues de Schleicher et donne la priorité au sentiment linguistique sur les faits phonéticohistoriques.

Troubetzkoï (cf. Sériot, 1996 : 115-126), pour sa part, a formulé en 1923 l'idée d'union de langues<sup>76</sup> à l'intérieur d'une même aire linguistique. Des groupes de langues seraient liés entre eux par une série de ressemblances dans la syntaxe, la morphologie et le vocabulaire, sans qu'il y ait entre ces langues de correspondances. La linguistique aréale étudie ces « ressemblances entre les langues parlées dans une région données, ressemblances dues au contact de langues qui peuvent être ou ne pas être apparentées » (Sörös, 2008 : 23). En sont issues les premières représentations des langues indoeuropéennes sur une base statistique (Kroeber & Chrétien, 1960).

Nous avons vu que l'évolution des langues peut s'étudier à partir des altérations graduelles du code : c'est à quoi s'attèle la linguistique comparée. D'autres courants étudient plutôt les contacts de langues, car « l'étude du contact conduit à s'intéresser non seulement aux langues dans leur abstraction et leur matériau mais aussi à la façon dont elles sont pratiquées » (Nicolai, 2010 : 70). Dans le même esprit, pour Gadet & Ludwig (2015 : 52), « C'est avant tout par des contacts ordinaires répétés entre locuteurs (y compris entre scripteurs et lecteurs) que des effets langagiers se produisent ».

#### **2.3.1.4 L'emprunt, effet du contact de langue**

Pour décrire les effets de contacts répétés entre locuteurs sur l'une ou l'autre langue, nous utilisons la terminologie utilisée par les lexicologues. Johanson (2002 : 251) a critiqué cette terminologie :

« Je soutiens que l'utilisation même de métaphores telles que 'emprunt', 'importation' et 'transfert' conduisent souvent des linguistes à de fausses conclusions quant à la nature de changements induits par contacts, car cela suggère une espèce d'implantation d'éléments étrangers dans le corps d'un Code de Base »<sup>77</sup>.

---

76 En russe, *jazykovoï sojuz*, ou son calque allemand : *Sprachbund*. (Drinka, 2017 : 410)

77 « I argue that the very use of metaphors such as 'borrowing', 'import' and 'transfer' often lead linguists to false conclusions concerning the nature of contact-induced change, since it suggests some kind of implantation of foreign elements into the body of a Basic Code ». Pour le terme *code*, Johanson (ibid. : 262) précise : « the term 'code' is not simply used instead of 'language', but as a neutral label for various kinds of linguistic systems, specific creative techniques of linguistic production ».

Johanson propose donc de parler de *copie*. Gadet et Ludwig précisent ce terme : « **la copie lexicale ouverte** est le phénomène de contact le plus fréquent, qui correspond à l'**emprunt** dans la terminologie traditionnelle. Il désigne la reproduction, plus ou moins conventionnalisée, d'un terme du lexique »<sup>78</sup> (ibid. : 56). Ils désignent par ailleurs le calque par *copie ouverte* : le signifié est copié mais est aussi une structure sémantogrammaticale (ibid. : 57).

Le terme copie (lexicale etc.) paraît relativement négatif en ce qu'il semble nier la faculté des locuteurs à adapter le mot à leurs besoins sémantiques. En effet, copier implique reproduire, alors que le processus d'intégration d'un mot d'une autre langue équivaut davantage à une adoption de sa signification ou d'une partie de celle-ci, et à une adaptation de sa forme (orale et écrite) aux sons et aux graphies qui existent dans le système qui accueille ce mot. Nous nous inscrivons dans les usages de la lexicologie en utilisant les termes *emprunt* et *calque*.

Des raisons sociohistoriques expliquent l'évolution des langues parlées dans les territoires où l'anglais et le français sont nés et où les emprunts lexicaux au latin ont joué un rôle prééminent.

## **2.3.2 Le latin comme dénominateur commun**

### **2.3.2.1 Origines latines du lexique de l'anglais et des langues romanes**

Selon l'estimation déjà ancienne de Cressman (1926 : 671), l'anglais serait composé de plus de 50 % de mots d'origine latine. Mossé (1947 : 157) pour sa part indique que « l'élément classique et roman occupe les trois cinquièmes du dictionnaire ». Il ressort par ailleurs du vocabulaire représentatif<sup>79</sup> des langues romanes (Sala, 1988 : 419 ; 257 ; 179 ; 477) que la part de mots d'origine latine s'élève à 72 % en italien, 70 % en portugais, 66 % en espagnol et 63 % en français<sup>80</sup>. Toutefois, il est difficile de comparer ces chiffres qui reposent sur des choix méthodologiques différents. Il n'en reste pas moins étonnant qu'une langue classée parmi les langues germaniques soit autant marquée par une langue d'un autre rameau. Des contacts anciens et continus avec le latin le justifient.

---

78 La source comporte ces mots en gras.

79 La délimitation du vocabulaire représentatif (entre 2312 et 1611 mots pour les quatre langues) repose sur une combinaison de trois critères : un nombre minimum de significations, un nombre minimum de dérivés et l'indice minimal d'usage des dictionnaires de fréquence de Juilland (1964 ; 1970 ; 1973)

80 32 % en roumain.

### 2.3.2.2 *Les langues celtiques de l'Antiquité face au latin*

Cette profonde influence date du premier millénaire. Avant que l'Empire romain ne conquière les territoires de l'ouest de l'Europe, les langues dont il y était fait usage appartenaient en grande partie à la branche celtique<sup>81</sup>. Nous ne mentionnerons que les contacts qu'ont eu les locuteurs de brittonique et de gaulois avec le latin. Des influences interlinguistiques ont eu lieu dans les deux sens, avec cependant un avantage massif pour le latin.

On connaît en latin des emprunts celtes, avec des mots tels que *carrus* (chariot à deux roues). Mais les dialectes celtiques du gaulois, sur le continent, et du brittonique, sur l'île, jouissaient d'un prestige inférieur au latin, si bien qu'elles n'existent plus aujourd'hui en français et en anglais que comme substrat : le brittonique est un substrat pour l'anglais et le gaulois pour le français<sup>82</sup>. Au point de vue lexical, plus d'une centaine des mots du français proviennent du gaulois, et du brittonique sont issus seulement une douzaine de noms communs de l'anglais (Bacquet, 1974 : 25-27)<sup>83</sup>, dont la plupart sont tombés en désuétude. L'hypothèse d'un substrat phonologique gaulois /y/ en français, a perdu de la vigueur (Perret, 1999 : 30). L'évolution du brittonique et celle du gaulois au contact du latin doivent cependant être contrastées.

Sur l'île, le brittonique a été pénétré par des mots latins pendant l'époque de la *Provincia Britannia*<sup>84</sup>, c'est-à-dire de 43 à 407 (Le Bohec, 2012 ; Salway, 1997 : 319). Pendant cette période, des colonies (Galsterer, 1997) ont été fondées où prévalaient le droit et les coutumes de Rome. Ces colonies ont établi une constitution en latin à laquelle les peuples devaient se soumettre, et élisaient un sénat qui débattait en latin. En outre était concédée à certains candidats - légionnaires démobilisés, etc., autant de locuteurs de latin - la propriété d'un tiers des terres conquises. Les échanges économiques plurilingues qui résultaient du contact entre populations locales et colons ont contribué à une acculturation progressive des Celtes, surtout au sud et à l'est. Dans les autres zones, Sawyer (1998 : 69) suppose que le latin a surtout été parlé dans les villes, lors d'interactions entre les élites locales, bilingues, et l'administration romaine.

En effet, comme partout ailleurs dans l'Empire Romain, le latin était la langue de l'administration de la Britannia et de la Gaule, et les dirigeants locaux inféodés étaient formés en latin. Au sujet de la Britannia, Virgile (98 : XXI) affirme que dès les années 70 de notre

---

81 À l'exception notable du territoire des Vascons dans le sud-ouest de la Gaule.

82 L'hypothèse d'un substrat paléoeuropéen (Krahe, 1964) ou préindo-européen (Vennemann, 2003), perceptible par exemple dans des oronymes et des hydronymes, est controversée car elle pose des questions de méthode.

83 Cf. également : [https://en.wikipedia.org/wiki/List\\_of\\_English\\_words\\_of\\_Brittonic\\_origin](https://en.wikipedia.org/wiki/List_of_English_words_of_Brittonic_origin)

84 Nous employons par la suite le mot latin *Britannia* pour éviter la confusion avec le mot français *Bretagne*.



ère, le gouverneur Agricola, « faisait instruire les enfants des notables dans les arts libéraux [...], et ceux qui jusqu'alors rejetaient la langue de Rome, désiraient absolument la parler avec éloquence»<sup>85</sup>.

La prudence s'impose néanmoins face à ce discours de propagande : la romanisation a été inégale en raison de troubles qui ont secoué l'ouest et le nord (Lambert, 1997 : 22), et les chercheurs s'opposent quant aux suites sociolinguistiques de la domination romaine : les habitants de la Britannia ont-ils à partir d'un moment adopté le latin dans leurs usages quotidiens ? Liberman (2005) nie que la romanisation ait abouti à une vernacularisation du latin, tandis qu'au contraire, Greenough *et al.* (1962 : 107) le soutiennent. Jackson *et al.* (1953 : 365) ont également affirmé qu'une variété britannique de latin<sup>86</sup> se serait développée. Certaines inscriptions le confirmeraient. Certains mots sont considérés comme des régionalismes, et la confusion des sons [b] et [w]<sup>87</sup> (Mullen, 2016 : 585-586) font penser à un latin vernaculaire en voie de dialectalisation. Des tablettes de défexion<sup>88</sup> trouvées dans des contextes ruraux et pour de menus larcins (vol d'un manteau sur la tablette de l'image 4) plaident pour une utilisation générale du latin, même par le peuple (Mullen, 2013 : 271).

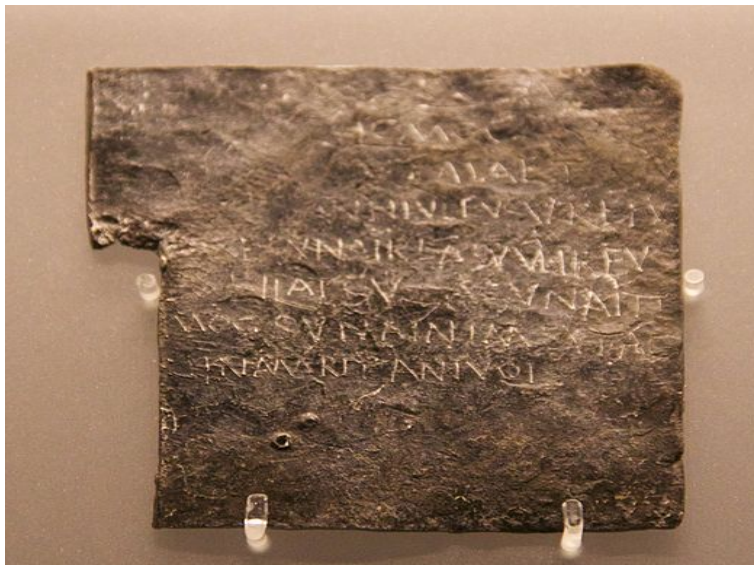


Image 4: Photographie d'une tablette de défexion, Bath (Angleterre), II-IVe siècles. Crédit : Mike Peel, CC BY-SA 4.0 <https://frama.link/v2m-64tT>

85 Notre traduction de « Iam vero principum filios liberalibus artibus erudire [...], ut qui modo linguam Romanam abnuebant, eloquentiam concupiscerent ».

86 Cette variété est appelée British Latin ou British Vulgar Latin en anglais.

87 Visible par exemple dans l'erreur *bissacio* pour *vissacio*, (fra) *besace*, dans l'inscription qui figure sur une tablette de bronze de Ratcliffe-on-Soar.

88 Les tablettes de défexion sont la plupart du temps des feuilles de plomb retrouvées dans des sources ou des puits ; souvent, on demande ainsi à une divinité aquatique de nuire à quelqu'un (voleur, partie adverse au tribunal, rival en amour, etc.).

Quant aux contacts entre Britannia et Gaule, Schrijver (2002 : 92-95) a montré que les inscriptions latines de l'île partageaient des traits avec le brittonique et le gaulois du nord-ouest (pour ce dernier point, voir aussi Adams, 2007 : 596-603).

Puis, l'effondrement des institutions romaines au V<sup>e</sup> siècle sur l'île aurait entraîné une baisse de prestige de la langue latine vernaculaire, et corollairement un prestige plus élevé de la culture des nouveaux dirigeants germaniques (Higham & Ryan, 2013 : 109-111). Le celtique brittonique a toutefois continué à être parlé<sup>89</sup>.

L'évolution sociolinguistique du gaulois est mieux connue : suite à la conquête romaine, l'usage du latin s'est répandu, les aristocrates gaulois et les commerçants, les villes, puis les campagnes ont adopté le latin après une période de diglossie (Ferguson, 1959) latin – gaulois. Ce dernier n'a apparemment plus été parlé à partir du V<sup>e</sup> siècle (Lambert, 2003 : 10, Walter, 2001 : 46), période où les Germains ont à leur tour conquis de vastes territoires continentaux et insulaires.

### **2.3.2.3 Influences du latin dans les langues des Germains**

Selon qu'on examine la langue parlée ou la langue écrite, l'anglais aujourd'hui ne présente pas la même proportion de mots d'origine latine. Pour l'expliquer, il nous paraît utile de synthétiser certaines étapes historiques en les envisageant sous l'angle des langues utilisées.

Nous avons vu que le latin avait été la langue officielle à l'époque romaine, sans que la question de la pénétration effective de cette langue dans les usages soit tranchée ; les langues brittoniques ont été parlées dans certaines zones pendant encore des siècles, et le gallois l'est encore aujourd'hui. Après le départ des troupes romaines, des Germains<sup>90</sup> ont été appelés à la défense des Brittoromains, autrement dit des insulaires romanisés, contre les incursions de peuples brittoniques du nord de l'île. Ces mercenaires germaniques provenaient de plusieurs zones côtières du continent, et selon Mossé (1958 : 12), « ils parlaient des dialectes à peine différenciés », si bien qu'on peut en déduire qu'il y avait intercompréhension entre ces locuteurs de langues germaniques de l'ouest.

On estime qu'à partir de 441-442 (Le Goff, 2008), des Germains ont migré par milliers et renversé l'aristocratie locale. De 475 à 827, ils ont fondé une vingtaine d'entités de pouvoir

---

89 Le gallois et le breton en sont issus.

90 Par *Germains*, on entend des individus appartenant à des peuples germaniques.

anglo-saxonnes<sup>91</sup> à côté des royaumes celtes. Traditionnellement, on distingue les Angles, les Saxons, les Jutes et les Frisons, quand bien même des fouilles ont mis au jour les mêmes pratiques culturelles sur le continent et dans les différents lieux de l'île où ces peuples se sont implantés.

Vers 430, des Saxons ont été autorisés par des Brittoromains à s'installer dans l'est de l'île, où ils ont fondé plusieurs royaumes. Les Angles, originaires de la région entre le Danemark et l'Allemagne, auraient d'abord émigré vers l'embouchure du Rhin, se mêlant à d'autres populations germaniques. Aux alentours de 440, certains d'entre eux ont migré eux aussi dans l'est de l'île, notamment suite à une dégradation du climat qui avaient lieu depuis un siècle (Gebühr : 2002). Ce sont également des inondations au III<sup>e</sup> siècle (Berglund, 2002 : 10) qui ont poussé les Frisons à quitter leurs terres les plus basses (Ejstrud et al., 2008 : 27). Des établissements frisons sont attestés au V<sup>e</sup> siècle dans la région du Kent. Enfin, les Jutes, originaires du nord de l'actuel Jutland (Bède, 731 : I, 15)<sup>92</sup> et de la Frise-du-Nord, se sont installés à leur tour dans le Kent (450) ; ils ont été dominés par des Saxons au VII<sup>e</sup> siècle.

Ces populations germaniques ont entretenu des contacts avec le continent pendant les V<sup>e</sup> et VI<sup>e</sup> siècles, et une seconde vague d'immigration a même eu lieu au cours du VI<sup>e</sup> siècle depuis l'actuelle Norvège (Carver, 1990 : 117). Il n'y a du reste rien d'étonnant à ce que des relations fréquentes aient eu lieu entre ceux qui vivaient de part et d'autre de la Mer du Nord, à parfois seulement quelques dizaines de lieues marines. Dans ces conditions, on peut supposer que l'intercompréhension s'est maintenue entre ces locuteurs.

Les Angles et les Saxons ont été désignés dès 600 sous le nom anglo-saxon *Engles*, et le sud-ouest s'est appelé England – le pays des Angles. Les dialectes utilisés par les Angles, Saxons, les Jutes et les Frisons sont à l'origine du vieil-anglais (ou anglo-saxon), le nom que porte le germanique insulaire d'environ 450 jusqu'au milieu du XII<sup>e</sup> siècle. Soumis à une variation diatopique dans les quatre dialectes principaux (northumbrien, mercien, kentien et saxon occidental), le vieil-anglais a été de plus en plus codifié à l'écrit jusqu'au X<sup>e</sup> siècle. À la réunion des royaumes anglo-saxons par Alfred le Grand en 878, le saxon occidental tardif de l'école de Winchester est devenu *de facto* la langue de l'administration<sup>93</sup>.

---

91 Voir pour cela la *List of Anglo-Saxon kingdoms* sur la page <https://en.wikipedia.org/wiki/Heptarchy>.

92 « De Iutarum origine sunt Cantuarii et Uictuarii, hoc est ea gens, quae Uectam tenet insulam, et ea, quae usque hodie in prouincia Occidentalium Saxonum Iutarum natio nominatur, posita contra ipsam insulam Uectam. »

93 Il a également été une langue littéraire pendant trois siècles.

Mais le prestige croissant du vieil-anglais n'a pas fait disparaître les nombreux mots latins qui émaillaient déjà les dialectes germaniques.

Selon Streadbeck (1966 : 92), les Germains du V<sup>e</sup> siècle avaient déjà emprunté environ 400 mots latins avant cette époque, pas seulement du fait de relations commerciales : l'Empire Romain recourait à des mercenaires germains, des *foederati*, qui ont été influencés par les coutumes romaines. On sait par exemple que des cavaliers Frisons ont participé à la conquête de la Grande-Bretagne par les Romains (Potter & Johns, 1992 : 190) et s'y étaient même déjà installés. Bacquet (1974 : 29-30) cite des emprunts au latin appartenant à la construction, au commerce, aux mesures, etc.

En second lieu, des mots latins ont été adoptés durablement pendant la période romaine. Bacquet (1974 : 31-32), Bourcier (1978 : 277) et Crépin (1994 : 164) citent des emprunts provenant de l'agriculture, du commerce, de la guerre et de la vie quotidienne<sup>94</sup>. Par ailleurs, Rome jouissait d'un prestige auprès des populations vivant près du *limes*, la frontière du nord de l'empire, si bien que les locuteurs de parlers germaniques eux aussi avaient emprunté nombre de mots. Ces influences interlinguistiques avaient déjà marqué les langues des Germains, si bien que Kersaudy (2001 : 154) écrit des populations brittoromaines : « les Celtes romanisés ont été envahis par des Anglo-saxons en cours de romanisation ».

En Europe de l'Ouest, pendant des siècles, le latin a de plus été variété haute<sup>95</sup> à deux titres, comme langue liturgique véhiculaire et comme langue de la connaissance.

#### **2.3.2.4 *Le latin, langue liturgique véhiculaire et langue de l'administration***

Nous avons dit que le latin était la langue de l'administration romaine. Ce statut de variété haute du latin s'est poursuivi sous l'influence de l'Église dont la langue était le latin dans l'ouest de l'Europe. La foi catholique trinitaire a été déclarée religion officielle et obligatoire en 380 par les empereurs Gratien et Théodose 1<sup>er</sup>. Sur l'île, une présence chrétienne est attestée au début du IV<sup>e</sup> siècle<sup>96</sup>, mais la conversion aurait véritablement pris de l'ampleur au début du VII<sup>e</sup> siècle<sup>97</sup>.

---

94 Des mots qui ont donné (eng) *cat*, *cock*, etc.

95 Au sujet de la critique de 'variété haute' et de 'variété basse', nous renvoyons à Tabouret-Keller (2006).

96 Trois évêques de la *Provincia Britannia* ont participé au concile d'Arles en 314.

97 Le premier roi anglo-saxon à se convertir au christianisme est Æthelberht de Kent, en 597.

Dès lors, le latin est devenu la langue liturgique véhiculaire, et nombreux sont les mots qui ont été empruntés à cette langue des deux côtés de la Manche : (eng) / (nld) / (deu) angel / engel / Engel ; bishop / bisshop / Bischof ; monk / monnik / Mönch, etc.

Étudiant les nouveautés et les particularités du vieux haut-allemand, le tout premier état de langue du haut-allemand après la seconde mutation consonantique<sup>98</sup>, Schmidt (1969 : 72) note que « la plus grande influence sur le stock de mots haut-allemand a été exercée par la christianisation »<sup>99</sup>.

Néanmoins, cette langue liturgique est devenue si éloignée de la langue parlée par le peuple que le culte n'était pas compris, même sur le continent, où la langue vernaculaire provenait pourtant du latin. Pour y remédier, le Concile de Tours de 813 a prescrit que les homélies soient faites dans la langue du peuple. Par ailleurs, en Angleterre, le roi Alfred a fait traduire en (vieil-)anglais des œuvres latines et fait rédiger en (vieil-)anglais des Annales et des textes de loi. La Réforme et le rejet de la papauté ont fait de l'anglais une langue liturgique, avec la première traduction de la Bible en 1537.

Par ailleurs, les Carolingiens ont encouragé la création d'ateliers de copistes, de bibliothèques et d'écoles où le latin était enseigné aux clercs, puis la réforme bénédictine a renouvelé dans toute l'Europe les études latines à la fin du X<sup>e</sup> (Mossé : 39). En conséquence, en sus de son utilisation liturgique, le latin adapté du latin ecclésiastique est devenu une variété haute (Mills, 2004 : XV) : les writs, écrits envoyés par le roi d'Angleterre à certains officiers, étaient rédigés en vieil-anglais jusqu'en 1070, date à laquelle les Anglo-Normands imposent le latin comme langue de l'administration (Genet, 2016 : 51).

### **2.3.2.5 *Le latin, langue de la connaissance***

L'enseignement à l'université et les échanges entre les scientifiques et les philosophes ont eu lieu en latin pendant des siècles. En effet, les lettrés ont appris le latin non seulement pour les besoins de la liturgie, mais aussi parce que l'Église a aussi fondé des écoles, « d'où se diffusait le savoir de l'époque, séculier autant que sacré » (Bourcier, 1978 : 277)<sup>100</sup>. Il en a découlé des emprunts au latin en vieil-anglais, par exemple dans les domaines suivants : la faune (*tiger*, etc.), la flore (*palm*, *pear*, etc.), l'écrit (*paper*, *part*, etc.) (Bacquet, *ibid.* : 35). En effet, selon

---

98 Entre les IV<sup>e</sup> et VIII<sup>e</sup> siècles.

99 Notre traduction.

100 Bourcier donne des exemples de mots anglais relevant de la religion (*abbot*, etc.) et d'autres domaines : *master*, *school*, *box*, etc.

Mossé (1947 : 31) « le latin, langue de l'Église, était aussi la langue de la science, de la philosophie, de toute la culture médiévale ». Walter (2014 : 84) cite de son côté des emprunts au latin qui appartiennent aujourd'hui à la langue de tous les jours : (eng) *butter, cheese, kitchen, short, street, etc.*

Sur le continent, les dialectes de langue d'oïl ont été relatinisés lors de la Renaissance carolingienne, vers le IX<sup>e</sup> siècle. En témoignent encore aujourd'hui en français actuel des mots des domaines juridiques<sup>101</sup>, religieux et savants.

À la Renaissance italienne, les humanistes se sont tournés vers les auteurs antiques et ont à la fois incité à lire dans le texte, en latin et en grec, et à traduire dans les langues vernaculaires. Au sujet des emprunts au grec, Bacquet (1974 : 81-82) souligne « l'œuvre considérable des traducteurs, qui, de manière oblique, poussent les lecteurs cultivés à se reporter à l'original ». Sont venus du grec, souvent par le latin ou le français (Mossé, 1947 : 81) les mots anglais suivants, qui sont des exemples des 2000 mots anglais fréquents de notre étude lexicologique que nous comparons avec leurs équivalents en espagnol, français, italien et portugais : *criterion, idea, parasite, scheme, theory, clinic, energy, rhythm, system, etc.*

Des sentiments ambivalents ont été notés quant à ce métissage : selon Bacquet (1974 : 73-75), de nombreux lettrés ont reproché une certaine mode du latin. Pourtant, les emprunts au latin se sont poursuivis sous leur forme originale : les humanistes anglais et leurs successeurs ont adopté des mots tels quels. *Area, datum, focus, major, medium, series, species*, ou encore *stimulus*, en sont des exemples que nous trouvons dans nos données. De l'italien proviennent de cette époque-là des centaines de mots des domaines de l'architecture, de la musique, de la littérature et du commerce, par exemple (ita) *villa, piano, novel, bank*.

Afin d'affirmer que l'anglais pouvait être une langue de culture, d'autres mots latins ont été anglicisés, parfois sans grand succès. Nous retrouvons là ce que nous avons constaté dans nos données : certains mots ont été créés dans une langue mais l'usage ne les a pas (encore) consacrés (cf. Variation diachronique et processus de néologisation en cours, p.184).

---

101 Le latin est resté la langue du droit en France jusqu'à l'ordonnance de Villers-Cotterêts, en 1539.

### 2.3.2.6 *Emprunts lexicaux au latin et au grec en biologie et en médecine*

Nous traitons ici des emprunts lexicaux au grec, et au latin dans le domaine de spécialité du corpus dont nous avons extrait des données. La section Analogies du lexique de la santé et de la chimie, p.56, a déjà abordé rapidement comment fonctionne la néologie dans ce domaine.

Le latin a emprunté des mots grecs très tôt, dès le VI<sup>e</sup> siècle avant notre ère, sous l'effet conjugué d'échanges commerciaux avec le monde hellénique et de la présence d'esclaves grecs (Grimal, 1984 : 173, Deléani & Vermander, 1975 : 5) À Rome, où le grec jouissait d'un grand prestige, cette langue était enseignée dans les écoles, et l'influence du grec sur le latin a entraîné des centaines d'emprunts d'origine grecque, à l'instar de (lat) *alphabetum*, *geographia*, etc.

Dans le domaine médical particulièrement, les exemples de termes latins provenant du grec abondent : (lat) *anatomia*, *arteria*, *epiderma*, *stomachus*, *thorax*<sup>102</sup>, etc. Les emprunts au grec en français dans le domaine de la santé sont légions : (fra) *pharynx*, *phénomène*, par exemple, ou encore les termes latins *examen*, *foetus*, *infarctus*, *pancreas*, *uterus*, *virus*<sup>103</sup> dont l'orthographe a été légèrement adaptés. D'autres appartiennent à la langue scientifique transdisciplinaire : *processus*, etc.

Dans les publications à caractères professionnel, l'anglais est aujourd'hui la langue la plus utilisée dans le domaine de la santé, si bien que de nombreux emprunts ont lieu de cette langue dans ce domaine. Le lexique des sous-domaines de la biologie et de la chimie, que nous avons inclus dans le lexique de la santé comme nous le précisons dans la section Analogies du lexique de la santé et de la chimie p.56, sont largement d'origine latine et grecque. En effet, au XVIII<sup>e</sup> siècle, le naturaliste Linné a cherché à ordonner le vivant dans une classification des espèces qui s'appuie sur ces deux langues (Walter, 2003 : 121-128). On peut mettre sur le même plan les termes de la chimie aujourd'hui utilisés internationalement (Walter, 2014 : 99-100), dont la forme a été empruntée tardivement au latin ou au grec, parfois par l'intermédiaire du français : dans nos données figurent par exemple (eng) *calcium*, *carbon*, *molecule*, *nucleus*, *oxygen*, *sodium*.

---

102 Ces exemples sont empruntés à Walter (2014 : 41-44).

103 Nous avons sélectionné ces mots chez Walter (2014 : 63) parce que leurs équivalents anglais figurent parmi les 2000 mots anglais fréquents de ScienText.

### 2.3.3 Influences des langues romanes

#### 2.3.3.1 Influences normandes et françaises sur l'anglais

La part de lexique d'origine française en anglais ne fait pas consensus parmi les chercheurs. Selon Bielenia-Grajewska (2009 : 11), 25 % du lexique de l'anglais d'aujourd'hui provient du français. Pour Martinet, « l'anglais, langue germanique, a emprunté la moitié de son vocabulaire au français, langue romane »<sup>104</sup>.

La place importante du lexique du français s'explique par le poids des locuteurs de français dans les classes dirigeantes de l'Angleterre pendant plusieurs siècles. Au XI<sup>e</sup> siècle : vainquant les nobles anglais en 1066, Guillaume le Conquérant a imposé ses barons normands ainsi que des prélats du continent. Tous parlaient le normand, l'un des dialectes de l'ancien-français. En conséquence, le français a été la langue de la cour d'Angleterre du milieu du XI<sup>e</sup> au début du XV<sup>e</sup>. Pourtant, « jamais il n'y eut de politique tendant à imposer la langue des conquérants » (Mossé, 1947 : 47). Les langues de prestige étaient alors le latin et le français (Crépin : 36). Pour Bacquet (1974 : 67), avant la Renaissance, « le français est senti comme langue de culture et de civilisation raffinée au même titre ou presque que le latin », et toute une littérature en français s'est développée en Angleterre : la littérature anglo-normande.

Les effets d'un bilinguisme français-anglais ont entraîné un profond métissage de l'anglais. Des études (dont Baugh & Cable, 1994 ; Strang, 1970 ; Lass, 1987 ; Pyles & Algeo, 1970 ; Trevelyan, 1949) ont montré l'influence du français de l'époque médiévale sur la langue anglaise. Katamba (1994 : 144) rappelle cependant une estimation de Lass : un maximum de 20 % de la population aurait su parler français, pour l'essentiel la cour, la noblesse et le haut clergé. Par une stratégie de mariages, les Normands ont été intégrés à la société anglaise dès le milieu du XII<sup>e</sup> siècle.

Environ 10 000 mots français auraient été empruntés pendant le Moyen Âge, essentiellement par la voie de l'écrit. La perte de la Normandie et la guerre de cent ans ont cependant mis un frein à l'usage du français à la cour d'Angleterre. Ainsi, en 1362, l'anglais est devenue officiellement la langue des tribunaux, et dès 1385, l'enseignement s'est fait en anglais dans les *grammar schools*. Le français est toutefois demeuré la langue de la correspondance, des actes officiels et des épitaphes jusqu'au XV<sup>e</sup>.

---

104 Préface à *L'aventure des langues en Occident*, Walter (1994), cité par Robert (2009A : 85).



L'adoption du normand comme langue de la noblesse a eu un impact considérable, dont nous voyons encore aujourd'hui les traces dans l'orthographe anglaise. Le lexique de l'anglais a été renouvelé : des mots du vieil-anglais comme (ango) *ēā*, *ieldra* et *weald*, ont laissé progressivement place dans l'usage aux mots (eng) *river*, *parent* et *forest* qui sont empruntés au français, avec des mots des domaines du droit et de la vie domestique, tels que (eng) *chair*, *judge*, *marry*, *pay*, etc. (Mossé, 1947 : 68-69). C'est à l'époque anglo-normande que tombe en désuétude le vocabulaire abstrait anglo-saxon. « En même temps affluent les termes abstraits. C'est tout le vocabulaire de la vie intellectuelle et morale que l'on emprunte au français et il est très étendu » (Mossé, *ibid.* : 69).

Katamba (146) écrit que pendant cette période, « l'influence des mots français sur l'anglais eut très peu d'effets sur la structure de l'anglais, excepté dans la phonologie où il entraîna [f] ~ [v] et [s] ~ [z], qui avaient été des allophones du même phonème, à se diviser en deux phonèmes séparés »<sup>105</sup>. Mais surtout, ce sont les graphies héritées de l'anglo-normand qui ont influencé celle de l'anglais (Mossé : 57)<sup>106</sup>.

Comme l'influence du latin sur le français et sur les autres langues romanes s'est poursuivie pendant toute l'époque médiévale, le français a contribué à l'adoption massive par l'anglais de mots latins et grecs (Bacquet, 1974 : 80 ; 88). Le normand a apporté aussi à l'anglais des mots d'origine germanique, qu'il avait lui-même empruntés au francique ou à l'allemand, à l'instar de (eng) *pocket*, *wait* ou encore *supper*.

Certains mots de vieil-anglais ont subsisté alors que d'autres ont été empruntés au scandinave ou au français, avec une différenciation sémantique. Cette originalité fait partie de la richesse actuelle de l'anglais, qui a des doublets ou des triplets là où il n'existait qu'un mot en vieil-anglais.

### **2.3.3.2 L'influence européenne du français du XVI<sup>e</sup> au XIX<sup>e</sup> siècle**

Le français a exercé après la première révolution parlementaire anglaise (1642-1651) une influence encore perceptible aujourd'hui du fait que les mots empruntés au français à l'époque du Classicisme l'ont été également dans d'autres langues européennes. Walter (2001 : 321) signale que « pendant près d'un siècle le français restera la langue scientifique par excellence,

---

105 La diphtongue /oi/ vient également du français, et des suffixes comme (eng) -age, -ance, -able, -ry.

106 Par exemple, curieusement pour un francophone d'aujourd'hui, Mossé (*ibidem* : 1958) note que ce sont les scribes anglo-normands qui ont multiplié l'usage de la lettre k.

au même titre que le latin », et que son influence s'effondre à partir de 1815, au profit de l'anglais.

Pour les autres domaines, Bacquet explique la place d'intermédiaire que joue le français : de nombreux mots, par exemple espagnols, arabes, persans, russes et italiens, ont été empruntés à d'autres langues qui « passent dans la plupart des cas par le canal du français » (1974 : 94). Il en est ainsi de (spa) *malaria*, *caramel* (ibid. : 92 ; 107), (ara) *cotton*, *zero* (ibid. : 94), (fas) *arsenic* (ibid. : 95), (rus) *steppe* (ibid. : 108), (ita) *attitude*, *risk* (ibid. : 107 ; 91).

## 2.3.4 Influences des langues germaniques

### 2.3.4.1 L'influence de la langue des Vikings

Les incursions sur les côtes franques et l'implantation sur le sol anglais de Danois, Norvégiens et Suédois pendant l'époque viking, vers 800-1050 n'ont pas eu les mêmes conséquences linguistiques sur l'ancien-français et le vieil-anglais.

Les deux dialectes vieux-norrois des Vikings, installés dans le Northumberland et l'Est-Anglie, puis dans toute la région du Danelaw à partir de 886, ont laissé des traces autrement plus durables en Angleterre qu'en France.

Quand l'ancien-français n'a guère emprunté que des termes liés à la navigation, des centaines de mots scandinaves ont remplacé des mots de vieil-anglais : des actuels mots tels que (eng) *are* (forme de *be*), *birth*, *die*, *wrong* et *call* en attestent (Bacquet, 1974 : 38-39 ; Pyles & Algeo, 2010 : 253-254, Ó Muirthe, 2010). Chia Kuo Ku (2009 : 38-45 ; 68-73) a envisagé les contacts entre des locuteurs de vieil-anglais et des locuteurs de norrois et propose la catégorie englobante des *mots parallèles*<sup>107</sup>, pour regrouper les mots qui sont des cognats (cf. Mesurer la similitude / l'analogie interlinguistiques, p.35) en norrois et en vieil-anglais, ou bien des emprunts à l'autre langue.

Illustrons la présence de similitudes au moyen de la pierre de l'image 5<sup>108</sup>. L'inscription en vieux-norrois en est la suivante : « *Alli let ræisa stæin þenna æftir sik sialfan. Hann tok Knuts giald a Ænglandi. Guð hialpi hans and* »<sup>109</sup>, autrement dit<sup>110</sup> : Alli fit ériger cette pierre (de lui-même). Il prit le tribut de Knut en Angleterre. Dieu ait son âme.

---

107 En anglais *parallel words*

108 [https://sv.wikipedia.org/wiki/Upplands\\_runinskrifter\\_194](https://sv.wikipedia.org/wiki/Upplands_runinskrifter_194).

109 [https://en.wikipedia.org/wiki/England\\_runestones](https://en.wikipedia.org/wiki/England_runestones)

110 Notre traduction sur la base de la page en anglais [https://en.wikipedia.org/wiki/England\\_runestones](https://en.wikipedia.org/wiki/England_runestones)

Une comparaison entre le texte en vieux-norrois et la traduction en français permet d'identifier des mots analogues entre le vieux-norrois (non) et l'anglais actuel (eng)<sup>111</sup> :

(non) / (eng) : *let* / *let* ; *ræisa* / *rise* ; *stæin* / *stone* ; *tok* / *took* ; *Ænglandi* / *England* ; *Guð* / *God* ; *hialpi* / *help*.

L'équivalence formelle d'autres mots est perceptible bien qu'ils ne soient plus idiomatiques dans ce contexte : *æftir* / *after* ; *sialfan* / *self*.



*Image 5 : Photographie de la pierre runique d'Uppland n°194, Väsby (Suède), début XIe. Crédit : Berig CC BY 2.5*

Mossé (1947 : 32) indique en outre que « la parenté entre le vieux-norrois et le vieil-anglais était si proche que bien des mots (comme *wife*, *burn*, *thick*, *folk*, *come*, *stand*, *sit*, *summer*) avaient à peu près la même forme ».

Une subsistance des deux mots, le mot vieil anglais et le mot scandinave, « a abouti à la création d'un micro-système de doublets » (Bacquet, *ibid.* : 46), avec spécialisation du sens.

---

<sup>111</sup> Et de l'allemand actuel : *Stein*, *sich selbst*, *England*, *Geld*, *Gott*, *helfen*.

Sur le continent, la langue des Vikings a eu peu d'influences sur l'ancien-français, parce que des terres leur ont été attribuées par Charles le Simple en 911 et que des mariages mixtes ont été conclus. Walter (1988 : 80) affirme : « au bout de trois générations, l'intégration des Normands est définitive ».

#### **2.3.4.2 *L'influence du bas-allemand sur le roman et l'anglais***

Les Francs et d'autres peuples germaniques tels les Burgondes, les Wisigoths et les Alamans ont influencé le latin vernaculaire employé sur les territoires européens où est parlé actuellement le français (Walter, 88 : 50). Nous passerons sous silence l'influence régionale des trois derniers peuples et concentrerons notre attention sur l'influence de la langue des Francs, le francique, au nord de la Loire.

Les Francs, des Germains originaires des rives du Rhin, se sont déplacés vers le sud à partir du III<sup>e</sup> siècle, et se sont installés dans une zone peuplée par des Celtes romanisés, occasionnant une « nouvelle source de contacts culturels et linguistiques » (Gadet & Ludwig, 2015 : 26).

Des interactions entre ces locuteurs est probablement née une langue hybride, c'est-à-dire « l'intrication de deux langues ou plus, possiblement à tous les niveaux linguistiques, du phonique au sémantico-pragmatique » (ibidem : 93), une langue qui « a intégré dans son système une proportion importante d'éléments d'une ou plusieurs langues » (ibid. : 53). Au sud du Rhin, là où la romanisation était plus récente, Gadet & Ludwig (ibidem : 52) pensent qu'il y a eu une diglossie emboîtée, autrement dit une situation avec plusieurs niveaux de diglossie. Si nous considérons que ces populations gauloises ont été romanisées à partir du début de notre ère seulement<sup>112</sup>, et que les Francs se sont installés vers le début du III<sup>e</sup> siècle, il n'est pas aberrant d'en déduire que les dialectes celtes de Gaule belgique pouvaient encore être parlés un siècle et demi après la romanisation, au moment où des locuteurs de dialectes germaniques sont entrés en contact avec les populations qui les ont précédés.

Pourtant, les emprunts au lexique du francique dans le latin vernaculaire des Gallo-Romains n'ont pas été nombreux, en dépit de leur domination politique jusqu'à la Loire. Bien plus, au moins la moitié des 400 mots français d'origine germanique ont été adoptés avant l'arrivée des Francs (Guinet, 1982 : 197). À la différence de ce qui s'est produit dans la Britannia, où les

---

112 Une présence romaine est attestée à Tongres, Bonn et Mayence quelques années avant notre ère ; la cité de Trèves a été fondée en 16 de notre ère.

Brittoromains ont fini par parler les dialectes germaniques des conquérants, ce sont les Francs qui ont adopté les dialectes romans utilisés par ceux qu'ils ont dominés sur les territoires gaulois. Le faible nombre de Francs par rapport à celui des Gallo-Romains pourrait expliquer que le francique soit devenu un superstrat par rapport au français<sup>113</sup>. Le francique a cependant profondément modifié la prononciation du latin au nord de la Loire du fait de l'accentuation tonique de la première syllabe (cf. également la section 1.5.1.2, Incidence phonéticohistorique de l'accent à l'initiale des langues germaniques, p.53).

Bacquet (1974 : 28) donne comme exemples d'emprunts au francique les noms courants de couleurs (fra) *blanc, bleu, brun, gris*, Walter (1988 : 57) des mots de la guerre, de la vie rurale, de la vie domestique, etc.

Par ailleurs, les dialectes bas-allemand ont exercé peu d'influences sur l'île : des contacts du fait du commerce avec les Pays-Bas ont entraîné quelques emprunts par le moyen-anglais, notamment dans les domaines maritime, et technique. C'est dans l'autre sens qu'ont eu lieu le plus de contacts. Enfin, des mots haut-allemands ont été empruntés au XVIII<sup>e</sup> siècle dans les domaines militaire et scientifique (notamment en minéralogie et en géologie).

### **2.3.4.3 Mondialisation des emprunts à l'anglais**

Robert précise la distinction entre langues voisines et langues proches. Il rapproche la notion de langues voisines de celle de langues collatérales (Éloy, 2004 : 10) : les langues voisines, seraient des « variétés proches aux plans linguistiques, historiques, sociolinguistiques et socioculturels alors que les langues proches n'offriraient qu'une parenté linguistique. D'un côté intercompréhension très forte et similitudes culturelles, de l'autre intercompréhension moins marquée et différences culturelles plus accentuées » (Robert, 2004).

La porosité des langues aux influences des langues voisines s'est étendue à des langues non voisines du fait d'échanges économiques et scientifiques. Sans entrer dans le détail des évolutions techniques qui ont facilité par exemple la navigation à longue distance ou encore l'impression de textes, permettant ainsi d'intensifier des contacts commerciaux ou intellectuels avec d'autres régions, nous noterons que des éléments du lexique ont été empruntés par les langues germaniques et les langues romanes à d'autres rameaux de la famille des langues indoeuropéennes, ou encore à d'autres familles de langues. Ainsi, du russe ont été adoptés et

---

113 Les Francs auraient représenté 5 % des Gallo-Romains (Walter, 2001 : 57).

adaptés des mots tels que *chamane, morse, steppe*<sup>114</sup>, de l'arabe *alcool, algèbre, zénith*, du turc *caviar, tulipe, yaourt*, du nahuatl *cacao, chocolat, tomate*, etc.

De l'examen de mots de onze langues qui avaient le même étymon (Walter, 2001 : 327 – 332), il est ressorti que 1225 mots ont une vocation internationale qui a dépassé les langues romanes et les langues germaniques puisque le finnois et le grec en font partie. En outre, examinant les composantes linguistiques du langage des sciences, Walter (ibidem : 314) écrit : « ce qui apparaît très nettement dès que l'on examine la nomenclature d'un certain nombre de disciplines scientifiques, c'est l'omniprésence des langues qui ont été sans conteste les langues de la science depuis plus de deux millénaires : le grec, et surtout le latin ».

Du fait de l'influence politique, économique, etc. de l'Empire britannique, puis des États-Unis, la langue anglaise jouit depuis le XVIII<sup>e</sup> siècle d'un prestige toujours accru, et « les emprunts faits à l'anglais pénètrent dans les langues romanes à un rythme de plus en plus accéléré, sous une forme parfois autochtonisée » (Reinheimer & Tasmowski, 1997 : 76). Aussi les échanges internationaux sont-ils conduits de manière privilégiée dans cette langue dans certains domaines de l'activité humaine. Logiquement, la plupart des emprunts se font à partir de l'anglais, ce qui contribue à une internationalisation du lexique.

---

114 Nous donnons les exemples en français.

## 2.4 Conclusion du chapitre 2

Dans ce chapitre, nous avons expliqué pourquoi l'anglais - et d'autres langues germaniques - et le français - et d'autres langues romanes - présentent des analogies. Quelle est la part de l'origine commune, quelle est celle des influences interlinguistiques liées aux contacts de langues ? Posée ainsi, cette question rappelle les théories opposant un déterminisme génétique à une influence du milieu. La quête d'un peuple qui aurait parlé une langue qui serait l'indoeuropéen évoque une fuite en avant en ce que l'ADN extrait d'un squelette ne pourra pas renseigner sur la ou les langues que l'individu a parlée(s), il permettra seulement de mieux comprendre les mouvements de population qui ont eu lieu. S'il est aberrant d'associer une langue à une ethnie et d'associer une ethnie à une langue, on peut en revanche privilégier une approche sociolinguistique et justifier les analogies interlinguistiques par des contacts entre des locuteurs de langues et de variétés de langues différentes. Il n'est besoin que d'observer aujourd'hui les emprunts à l'anglais, et pas uniquement ceux dont témoigne l'interlexique que nous avons élaboré.

Nous nous sommes appuyé sur l'Histoire européenne pour mettre en relief les contacts successifs entre des locuteurs de brittonique et de latin britannique, de gaulois et de latin vernaculaire parlé en Gaule, de francique et de bas-latin, de vieil-anglais et de norrois, d'ancien français et de norrois, d'anglais et de français. Si nous sommes remonté dans le temps, c'est parce que les analogies interlinguistiques entre les langues romanes et les langues germaniques sont anciennes et qu'elles sont si nombreuses qu'elles justifient qu'on envisage de les valoriser dans des dispositifs didactiques afin d'aider à la mémorisation du lexique. La latinité du lexique, et sa grécité dans une moindre mesure (cf. également Latinité et grécité de mots anglais de spécialité : approche hypothético-déductive empruntée à l'intercompréhension, p.165), proviennent de strates successives et imbriquées. Cependant, l'apport de ces deux langues anciennes est considérable, et un enseignement des langues vivantes, romanes ou germaniques, pourrait davantage montrer l'influence essentielle du latin et du grec, tout comme l'influence du chinois sur le coréen, le japonais et le vietnamien est incontournable, ou encore celle du sanscrit sur des langues comme le khmer et le lao.

L'anglais, le français et l'espagnol, qui comptent parmi les langues étrangères les plus apprises actuellement, sont des langues indoeuropéennes. En conséquence, nous supposons que dans de nombreux contextes d'enseignement, des analogies qui existent entre une langue indoeuropéenne du répertoire de l'apprenant et une autre langue cible indoeuropéenne

pourraient être valorisées. Les analogies qu'on observe entre l'anglais et le français, par exemple, peuvent être précieuses dans l'optique d'un enseignement du français à un public qui connaît déjà l'anglais, et peut par conséquent reconnaître dans des mots français des mots anglais qu'il a appris. Certains courants s'y emploient, et des outils développés par la linguistique informatique s'y emploient. C'est l'objet du chapitre 3.



# Chapitre 3 - Valorisation des analogies interlinguistiques par l'intercompréhension et les approches sur corpus

## Sommaire du chapitre 3

3.1 Introduction du chapitre 3.....	104
3.2 L'analogie en didactique des langues : risques et promesses.....	106
3.2.1 Importance de l'analogie dans l'appropriation du savoir.....	106
3.2.2 Du risque d'interférences.....	106
3.2.3 Des influences des langues connues dans l'apprentissage d'une L3.....	106
3.3 Des approches novatrices qui développent l'intercompréhension.....	113
3.3.1 Ancrage et définition.....	113
3.3.2 But et principes de l'intercompréhension.....	113
3.3.3 Stratégies et techniques.....	114
3.3.4 Les langues romanes, des langues 'pluricentriques' ?.....	115
3.4 Apport des approches sur corpus pour la connaissance des analogies linguistiques.....	118
3.4.1 Définitions d'une discipline et d'un outil.....	118
3.4.2 Distinction entre corpus comparables et corpus parallèles.....	118
3.4.3 Vers une compréhension plus fine du discours de spécialité.....	120
3.5 Présentation du corpus ScienText utilisé pour l'étude lexicologique.....	124
3.5.1 Origine et principes de constitution de ScienText.....	124
3.5.2 Comment à requêter partir de ScienQuest.....	124
3.6 Conclusion du chapitre 3.....	126

### 3.1 Introduction du chapitre 3

Nous examinons dans ce chapitre 3 comment les analogies interlinguistiques sont valorisées par des approches innovantes qui contribuent au développement du plurilinguisme et/ou renouvellent la connaissance du lexique.

Nous voyons tout d'abord l'importance des analogies interlinguistiques en didactique des langues, le risque d'interférences qu'elles représentent, l'influence de l'ordre d'apprentissage des langues et le concept de langue passerelle. C'est la section 3.2, p.106.

Les approches qui se réclament de l'intercompréhension font l'objet de la section Des approches novatrices qui développent l'intercompréhension, p.113. Elles renouvellent l'apprentissage d'une langue en valorisant les analogies lexicales, syntaxiques, et morphosyntaxiques que nous avons étudiées dans la section 1.4, p.48, et en mettant l'accent sur la compétence métalinguistique des apprenants.

Parce qu'elle permet une approche systématique des analogies linguistiques en contexte, la linguistique de corpus présente également un potentiel énorme pour la phraséologie, la didactique d'une langue et la didactique du plurilinguisme. Nous le voyons dans la section 3.4, p.118.

## **3.2 L'analogie en didactique des langues : risques et promesses**

### **3.2.1 Importance de l'analogie dans l'appropriation du savoir**

Le cerveau humain use d'un raisonnement analogique. Frei (1929) a défini le phénomène d'assimilation mémorielle, qui consiste à « modifier ou créer un élément par imitation d'un modèle logé hors du discours, dans la conscience linguistique ». Il a appelé cela « l'instinct analogique ». Lorsque celui qui explique une notion abstraite recourt à la comparaison, il permet à son interlocuteur d'associer la nouveauté à ce qu'il connaît déjà, c'est l'objectivation (Moscovici, 1976), un processus essentiel dans l'acquisition (Nautier & Allieu-Mary, 2008). De cette façon, la mémorisation s'opère à partir de l'univers de références de l'apprenant, qui construit de nouveaux acquis. L'objectivation peut être rapprochée du concept de transfert.

Perrenoud (1997) définit le transfert comme « la capacité d'un sujet de réinvestir ses acquis cognitifs – au sens large – dans de nouvelles situations », et se demande comment le sujet parvient à penser à la fois les différences et les similitudes. Il reprend la définition du transfert de connaissances de Mendelsohn (1996 : 15), à savoir « le mécanisme qui permet à un sujet d'utiliser dans un nouveau contexte des connaissances acquises antérieurement ».

La question du processus de transfert des connaissances se pose bien entendu dans l'acquisition d'une langue, que ce soit la L1 ou les langues ultérieurement apprises. Les transferts entre les langues du répertoire de l'apprenant, mais aussi éventuellement entre des langues voisines apprises simultanément se heurtent au risque d'interférences linguistiques.

### **3.2.2 Du risque d'interférences**

Selon certains linguistes (par exemple Fries, 1945), ce qui dans la L1 est différent de la langue étrangère induirait des problèmes d'apprentissage et des erreurs, tandis que ce qui est identique ou similaire serait facile à apprendre. Ainsi, plus le degré de différence est grand entre la L1 et la L2, plus grande serait la difficulté à apprendre la L2. Les transferts de la L1 à la L2 engendreraient des interférences et il faudrait faire perdre de mauvaises habitudes (*bad habits*) à l'apprenant par des mesures didactiques ciblées.

Riemer (2002 : 62) critique cette hypothèse behavioriste car on sait par ailleurs qu'il y a une différence selon qu'une langue donnée est la L1 ou la L2 : les apprenants d'une L1 qui apprennent une L2 ne rencontrent pas les mêmes problèmes que des apprenants de cette L2

qui apprennent cette L1, les apprenants commettent aussi des erreurs dans la L2 qui ne peuvent pas s'expliquer par des transferts de la L1, les apprenants qui ont la même L1 ne commettent pas les mêmes erreurs dans une L2 donnée, et les apprenants font naturellement des erreurs dans la L2 parce qu'il leur manque tout simplement le savoir dans cette langue.

Mais la priorité dans l'enseignement d'une langue n'est pas tant de prévenir les interférences que de créer des conditions qui stimulent les rapprochements bénéfiques. Selon Neuner (2003 : 25) :

« le plus important, c'est la question du *transfert* : comment faire le lien entre l'apprentissage d'une langue étrangère et ce qui est déjà acquis, en termes de connaissances linguistiques, d'expérience linguistique fondamentale et d'*apprentissage* des langues pour leur permettre de se développer ? ».

Neuner en déduit que le transfert porte avant tout sur deux dimensions dans le concept plurilingue.

Premièrement, le renforcement de l'acquis linguistique, en construisant des *passerelles de transfert* (Meißner, 2000), qui vont déclencher des *transferts de reconnaissance* :

« Pour la construction de ces « passerelles de transfert », ce sont tout d'abord les similarités entre les types de langues L1 – L2 – L3 qui jouent un rôle prépondérant et, dans le domaine du vocabulaire, ce qui est important, c'est le degré d'intensité du contact linguistique (apport d'internationalismes et d'emprunts) » (Neuner, *ibidem* : 25).

Deuxièmement, le développement de la compréhension de l'apprentissage des langues par la réflexion sur les processus et les expériences de l'apprentissage des langues. Pour Neuner (*ibid.* : 25-26) :

« il faut prendre en compte non seulement la comparaison des langues au niveau des formes linguistiques (signification identique / similaire ou « faux amis »), mais aussi, avec tout autant d'attention, la façon dont est constitué le lexique mental et comment se fait l'apprentissage du vocabulaire (apport, stockage et activation des unités sémantiques).

Le transfert consiste ici à reprendre, activer et, le cas échéant, développer et différencier chez l'apprenant, l'expérience d'apprentissage linguistique acquise dans sa langue maternelle, et – encore plus important – dans la première langue étrangère ».

Nous souhaitons idéalement susciter des transferts dans l'enseignement d'une langue étrangère, le français, en activant l'expérience linguistique acquise dans la première langue étrangère qu'est l'anglais pour de très nombreux apprenants de français. En effet, les interférences entre les langues font partie d'influences plus larges liées au répertoire langagier de l'apprenant et à l'ordre d'apprentissage des langues.

### **3.2.3 Des influences des langues connues dans l'apprentissage d'une L3**

#### ***3.2.3.1 Influences plus marquées de la L2***

Au cours de l'apprentissage d'une L3, les langues du répertoire de l'apprenant influencent l'acquisition de la L3. Les influences de la L1 et de la L2 sur l'apprentissage d'une L3 ont été étudiées depuis les années 1970 : Stedje (1976) a montré que l'apprentissage de l'allemand L3 par des locuteurs natifs de finlandais était influencée par leurs acquis en suédois L2. Hufeisen (2004) a récapitulé les différents modèles, hypothèses et théories qui ont cours sur l'apprentissage des langues, et a conclu que l'acquisition de la L1, de la L2, de la L3, etc. ne se déroule pas de la même manière. Les phénomènes de transferts et d'interactions entre les langues dépendent de facteurs aussi divers que le statut de la L2, les relations entre les langues, leur complexité, la conscience métalinguistique, les stratégies d'apprentissage etc. Nous ne reprendrons pas plus avant cette littérature. Nous noterons cependant que selon la perspective du facteur L2 (Hammerberg, 2001 : 23, Bono, 2008), il semblerait qu'il y ait une tendance générale à activer davantage la L2 que la L1 lors de l'apprentissage d'une L3. De même, selon Forlot (2009), « les locuteurs d'une L1 sans relation avec leur L2 et/ou leur L3 ont souvent recours à des transferts provenant de leur L2 ».

Nous pensons qu'il serait utile d'éprouver cette hypothèse dans le domaine du FOS lors d'expérimentations où l'enseignement du français L3 serait construit sur les connaissances des apprenants dans une langue passerelle, par exemple l'anglais.

### 3.2.3.2 *L'intérêt d'une L2 passerelle pour apprendre une L3*

Dans sa thèse, Aparicio (2003 : 262-263) conduit plusieurs tests sur la reconnaissance visuelle de mots chez des trilingues afin de comparer le coût du changement de langues. Au sujet du rôle que peuvent jouer les analogies interlinguistiques, il note :

« Les participants ne sont pas nécessairement habitués à effectuer des changements de langue de la L2 vers la L3 et réciproquement. Le traitement consécutif de ces deux langues, supposées moins bien représentées au sein du lexique multilingue car relativement peu utilisées, est donc susceptible de révéler une asymétrie très forte se caractérisant par un ralentissement considérable de la lecture de mots lors d'un changement entre ces deux langues. Toutefois, d'autres passerelles peuvent s'établir, notamment sur la base de ressemblances formelles entre les mots. »

L'existence d'analogies interlinguistiques – syntaxiques, morphologiques, lexicales – autorise d'envisager que des acquis dans une L2 puissent favoriser des transferts en vue de l'acquisition d'une L3. La langue dont des éléments sont transférables est désignée comme langue passerelle, ou langue pont (*Brückensprache, bridge language*).

Les recherches sur le développement de compétences de compréhension écrite dans une langue romane sur la base d'analogies interlinguistiques ne sont pas récentes. Walter (1991) a étudié l'acquisition de compétences de compréhension écrite du portugais à des étudiants germanophones qui connaissent l'espagnol. Pour l'apprentissage du portugais par des anglophones, Montrul & al. (2011), ont montré que des transferts syntaxiques ont lieu surtout de l'espagnol et partiellement de l'anglais.

Des dispositifs didactiques ont été fondés sur l'idée d'une langue passerelle : partir des connaissances de l'apprenant dans des langues a été l'idée novatrice d'EuroCom-online<sup>115</sup>, qui s'adresse aux étudiants ayant l'allemand comme L1 et le français comme L2, et qui désirent apprendre une langue du même groupe linguistique (l'italien, l'espagnol ou le roumain). Le français est considéré comme la « langue dépôt » :

« Pendant la lecture du texte authentique, l'étudiant reconnaît les interlexèmes et l'étudiant germanophone fait des associations interlinguistiques avec l'allemand et l'anglais (premier tamis). Il utilise sa langue dépôt romane, le français, faisant l'association interlinguistique panromane » (Klein, 2004).

---

115 <http://www.kom.tu-darmstadt.de/eurocom/>

Le programme d'intercompréhension EuroCom a ensuite été élargi :

« Pour la compréhension textuelle de n'importe quelle langue romane, une bonne connaissance d'une des langues romanes, de préférence du français, représente le cas idéal. Mais la réalité de la distribution des connaissances linguistiques en Europe nous force à prendre en considération l'anglais. » (Klein, 2008).

La possibilité a donc été étudiée d'envisager l'anglais comme une langue passerelle dans l'intercompréhension des langues romanes (Grzega, 2005 ; Klein & Reissner, 2006, Hemming *et al*, 2011).

### ***3.2.3.3 L'anglais, une langue passerelle vers les langues romanes***

Malgré la présence de nombreux sosies (cf. Définition habituelle de faux amis et failles de cette définition selon Kauffer, p.43), l'anglais et le français sont maintenant considérées par certains chercheurs comme des langues amies plutôt que comme des ennemies (Castagne, 2008) : les multiples analogies entre l'anglais et les langues romanes aux niveaux du lexique, de la syntaxe et de la morphosyntaxe (cf. Principales analogies et dissemblances entre l'anglais et les langues romanes, p.52), justifient à nos yeux que soient explorées le rôle de l'anglais comme langue passerelle vers ces langues.

Des recherches rendent compte d'une utilisation de l'anglais comme langue pont vers l'espagnol L3 pour des apprenants qui ont comme L1 l'allemand (Reissner, 2012 ; Grünwald, 2014). Par ailleurs, Muftah (2014) a constaté que lorsque les étudiants arabophones ont l'anglais en L2, l'accès au français L3 est favorisé grâce à des passerelles linguistiques. Le degré d'intercompréhension qui existe entre l'anglais et le français fait l'objet de recherches (Robert, 2008 ; Forlot 2009). Robert (2009A : 94) nuance cependant :

« l'utilisation de l'anglais comme langue dépôt, comme base de transfert, pour l'acquisition d'une compétence de compréhension passive en français est possible, mais reste limitée à un public bien précis : locuteurs natifs ou apprenants avec un bon niveau dans cette langue ».

Des limites au recours à l'anglais sont pointées par d'autres études. D'après Bailly *et al*. (2009), l'anglais peut être perçu comme une aide par certains apprenants de FLE, mais cela dépend du niveau du cours et du domaine d'apprentissage. D'autre part, Cuet (2011)

mentionne la crainte des apprenants chinois de confondre l'anglais et le français. Ces obstacles cognitifs liés aux interférences interlinguistiques, mais aussi aux traditions d'apprentissage, posent un défi à un enseignement d'une L3 qui chercherait à valoriser les acquis en L2. Il est peut-être possible d'atténuer cette insécurité linguistique en présentant les deux langues comme des langues qui se ressemblent à bien des égards, et en annonçant clairement que la pédagogie du français va s'appuyer au départ sur les analogies afin d'aider à comprendre certains traits essentiels du français. Ce faisant, il ne s'agirait pas de nier les différences entre les deux langues, mais de faire savoir à l'apprenant que ce procédé qui les fera paraître artificiellement proches lui permettra d'entrer plus rapidement dans la L3.

#### ***3.2.3.4 Psychotypologie et apprentissage d'une langue***

Selon la perspective psychotypologique (Kellerman, 1979 ; Ringbom, 2007 ; Singleton, 1999, Ó Laoire & Singleton, 2009), la conscience d'une proximité typologique entre les langues est déterminante : un apprenant construirait plus facilement son apprentissage d'une L3 s'il perçoit des analogies entre les langues de son répertoire et la L3. Les résultats obtenus par Cenoz (2001 : 16) montrent que « la perception de la distance linguistique et la perception de la 'transférabilité' peuvent être plus importantes que la distance linguistique »<sup>116</sup>.

Nous cherchons à créer des conditions d'apprentissage qui valorisent les analogies interlinguistiques et sensibilisent à la proximité entre certaines langues afin de favoriser les transferts positifs entre une L2 pont et plusieurs L3 apparentées entre elles mais non apparentées à cette L2. Cette approche devrait être possible parce que les analogies interlinguistiques entraînent à un certain degré une intelligibilité entre des locuteurs de plusieurs langues, mais aussi parce que la recherche didactique explore des voies pour que des apprenants de différentes langues apprennent à se comprendre sans apprendre à parler en même temps.

#### ***3.2.3.5 Intelligibilité inhérente et intelligibilité acquise***

La distance entre les langues est souvent estimée au moyen de la méthode lexicostatistique, qui mesure la part du lexique commune à deux langues. Comparant les langues utilisées au

---

116 Notre traduction.



Epele *et al.* (2012 :27)<sup>117</sup> estiment que deux langues ou variétés présentent une intelligibilité inhérente lorsque plus de 85 % d'une liste standard de mots montrent une similarité à la fois dans la forme et dans le sens, autrement dit sont analogues à 85 %. Les contacts des locuteurs de langues et de dialectes différents permettent l'intelligibilité. L'intelligibilité est dite acquise lorsqu'elle procède d'une exposition ou d'un apprentissage de l'autre langue. Les approches didactiques de l'intercompréhension peuvent y contribuer.

---

117 <http://www-01.sil.org/asia/nepal/downloads/ethnologuenepaltextnoprnt.pdf>

### 3.3 Des approches novatrices qui développent l'intercompréhension

#### 3.3.1 Ancrage et définition

L'intercompréhension compte parmi les approches plurielles des langues et des cultures. Selon le Centre européen pour les langues vivantes<sup>118</sup>, « Les Approches plurielles des langues et des cultures reposent sur des activités d'enseignement-apprentissage qui impliquent plusieurs (= plus d'une) variétés linguistiques et culturelles »<sup>119</sup>. Le Cadre de Référence pour les Approches Plurielles des langues et des cultures (CARAP, Candelier *et al.*, 2011) formalise des compétences et des ressources communes aux approches plurielles.

Le terme *intercompréhension* a été employé pour la première fois par Ronjat (1913) pour désigner le phénomène qui existe dans le continuum (cf. Le concept de continuum, p.83) des dialectes de l'occident. Degache (2009 : 85) définit l'intercompréhension comme « une forme de communication plurilingue où chacun s'efforce de comprendre la langue de l'autre et s'emploie à se faire comprendre ». Les analogies entre le connu et ce qu'on croit reconnaître participent de l'intercompréhension dans une situation de communication : un locuteur / lecteur développe des compétences partielles de compréhension en décodant plus ou moins consciemment les analogies qui existent entre sa L1 / L2 / L3 et la langue qu'utilise son interlocuteur. Dans cette approche communicative, la langue cible n'est pas envisagée comme un objet dont l'apprenant devrait apprendre à maîtriser les compétences de compréhension et de production.

#### 3.3.2 But et principes de l'intercompréhension

Le but de l'intercompréhension est avant tout fonctionnel. Selon Caddéo et Jamet (2013 : 50-64), l'intercompréhension repose sur plusieurs principes, dont nous reprenons ci-après ceux qui semblent pertinents<sup>120</sup> dans une approche du français par la passerelle qu'est l'anglais.

- Le principe de proximité, que cette proximité soit linguistique ou qu'elle découle « de tout ce qui a trait à l'expérience », c'est-à-dire par exemple du type de document, de la connaissance préalable du thème, etc. ;

---

118 <http://www.ecml.at/Home/tabid/59/language/fr-FR/Default.aspx>

119 <http://carap.ecml.at/CARAP/tabid/2332/language/fr-FR/Default.aspx>

120 Nous écartons la place de la L1 dans l'apprentissage et l'intercommunication / l'interproduction.

- Le principe de la simultanéité des langues. Pour les auteures (ibid. : 53), « Cela veut dire que plusieurs langues peuvent figurer ensemble dans la même séquence pédagogique. Ainsi l'apprenant reste-t-il concentré sur l'objectif de compréhension : en changeant la langue cible dans une même séance de travail qui vise la compréhension, l'apprenant doit maintenir son attention sur le savoir-faire à transférer et le sens à construire sans s'interroger sur sa légitimité à travailler sur ces langues » ;
- Le principe d'immersion : « il faut entendre immersion comme le contact immédiat avec la langue authentique à travers les documents et sans qu'il y ait eu apprentissage préalable » (ibid. : 57).

Nous avons déjà parlé de la proximité entre l'anglais et le français. Nous entendons l'utiliser avec des apprenants comme nous le détaillons dans le chapitre Perspectives : vers un dispositif didactique plurilingue d'enseignement du FOS sur la base d'analogies interlinguistiques, p.247. Le recours simultané à des exemples provenant de plusieurs langues romanes pourrait faciliter l'enseignement du FLE. Pareillement, le recours à des exemples tirés du français pourrait faciliter l'enseignement des autres langues romanes. Pour ce qui est de l'immersion, les apprenants du FOS de la santé seront mis face à des écrits authentiques de leur spécialité.

### **3.3.3 Stratégies et techniques**

Pour développer l'intercompréhension à l'écrit, plusieurs stratégies et techniques sont appliquées (Caddéo & Jamet, 2013 : 65-81) :

- apprendre à lire et à comprendre en langue étrangère ;
- stimuler la perception de la proximité ;
- encourager les inférences ;
- utiliser certains types de documents.

Afin de stimuler la perception de la proximité, certains traits font l'objet d'une comparaison. Or on sait que les approches contrastives des langues présentent un intérêt pour la mémorisation et la réduction de l'insécurité linguistique (Blanchet, Clerc & Rispaill, 2014). Les approches plurilingues ne visent pas l'appropriation parfaite, idéale (idéalisée) d'une langue mais une compétence de compréhension étendue. Selon Castellotti & Moore (2002),

« Des enseignements ainsi axés sur une optimisation didactique de la proximité interlinguistique doivent en particulier permettre de valoriser l'expertise des apprenants en leur donnant l'occasion de se servir de leurs compétences linguistiques, et de mettre à l'épreuve leurs intuitions sur le langage et les moyens de transférer leurs compétences d'un contexte à l'autre. L'objectif est la mise en valeur et la construction de compétences transversales, qui peuvent être réactivées quelles que soient les langues apprises ».

Parce qu'elle est réceptive (Ollivier & Strasser, 2013 : 42-44), l'intercompréhension présente des atouts pour le plurilinguisme (Bär, 2004) : l'apprenant n'a pas à produire dans la langue étrangère, ce qui diminue la charge cognitive, il peut alors concentrer ses efforts sur la compréhension et apprendre à comprendre plusieurs langues.

Plusieurs méthodes ont été élaborées à cette fin : EuroComRom (Klein, 2004), Galatea (Dabène, 2003), EuRom5 (Bonvino et al., 2011), ICE (InterCompréhension Européenne, Castagne, 2007B : 466-468), « comprendre les langues romanes » (Teyssier, 2012), etc. Certains chercheurs développent l'intercompréhension par interactions en ligne, par exemple anciennement sur le portail galanet.eu (Martin Kostomaroff, 2010), et depuis 2015 sur miriadi.net (Garbarino, 2015).

Nous réfléchissons à une manière d'aborder la langue sur objectifs spécifiques en incorporant certains apports de l'intercompréhension : les analogies linguistiques entre l'anglais et le français serviraient alors de catalyseur, pour emprunter une métaphore à la langue de la chimie. Dans cette approche, nous ferions appel simultanément à d'autres langues romanes si l'autre langue romane permet de dépasser l'opacité du français avec l'anglais.

La didactique des langues reste majoritairement orientée vers des approches monolingues, c'est-à-dire des approches qui visent l'enseignement d'une seule langue. Présenter plusieurs langues en parallèle revient à refuser cette restriction. Les langues qui ne sont pas ciblées au départ entrent en jeu comme des variétés d'un même continuum.

### **3.3.4 Les langues romanes, des langues 'pluricentriques' ?**

On peut se demander si l'on pourrait considérer les langues romanes comme des langues pluricentriques nées du latin, des langues qu'on pourrait enseigner comme autant de variétés d'une seule langue.

Selon Clyne (1992),

« Le terme pluricentrique a été employé par Kloss (1978 II : 66-67) pour décrire des langues avec plusieurs centres interagissants, chacun fournissant une variété nationale avec au moins certaines de ses normes (codifiées) ». <sup>121</sup>

Les approches plurilingues des langues romanes (ou de langues appartenant à d'autres familles) ne visent pas une seule langue à atteindre dans l'apprentissage. Dans ce travail, nous n'évitons malheureusement pas le biais du poids attribué aux langues selon leur prestige international ou leur nombre de locuteurs, puisque nous avons traduit les données dans les quatre langues romanes les plus utilisées dans le monde actuellement. Le fait que nous ayons choisi ces quatre langues romanes et que nous nous focalisions sur l'enseignement du français ne signifie cependant pas que nous leur accordions un prestige supérieur à d'autres, et nous sommes d'accord avec la position suivante de Mira Mateus et Villalva (2006 : 27) <sup>122</sup> :

« Ce n'est pas d'être parlée par plus de personnes, dans plus de pays ou dans plus d'institutions internationales qu'une langue gagne en valeur intrinsèque. Du point de vue linguistique, le nombre de locuteurs d'une langue ou le prestige international qu'elle peut avoir sont des critères de comparaison absolument vides de signification ».

Nous avons remarqué lors de communications que nous avons données à des colloques que la question de la cible des activités didactiques que nous prévoyons revenait fréquemment : qui serait donc intéressé par des activités qui incorporent plusieurs langues ? Ne serait-il pas plus réaliste de se concentrer sur l'enseignement d'une langue, dans notre cas le FLE ? Ce serait à notre sens adopter une approche restrictive, celle que nous déplorons dans les approches monolingues, et renoncer aux avantages qu'offre l'existence d'analogies interlinguistiques pour apprendre à des apprenants à tisser des liens dans la construction de leur compréhension d'une langue, et de ce que sont les langues.

Apprendre à comprendre le français en découvrant des analogies avec l'anglais et des langues romanes revient d'une certaine manière à dédramatiser l'apprentissage des langues, à

---

121 « The term *pluricentric* was employed by Kloss (1978 II : 66-67) to describe languages with several interacting centres, each providing a national variety with at least some of its own (codified) norms. », notre traduction.

122 « Não é por ser falada por mais pessoas, em mais países ou em mais instituições internacionais que uma língua ganha valor intrínseco. Do ponto de vista linguístico, o número de falantes de uma língua ou o prestígio internacional que ela possa ter são critérios de comparação absolutamente vazios de significado », notre traduction.

envisager que ces langues sont autant de variétés d'une langue pluricentrique, malgré leurs indéniables différences.

Nous avons montré dans la section Le latin comme dénominateur commun, p.86, la place centrale du latin dans la construction progressive des langues utilisées dans l'Ouest de l'Europe. Du rôle qu'à joué le latin, Escudé & Janin (2010 : 73) affirment :

« Le latin a longtemps joué un rôle structurant entre locuteurs et langues de la Romania. Dans une grande partie de l'Europe, la législation, la culture, la religion, la représentation scientifique et imagée du monde, tout témoigne d'une romanité commune. Or, en l'absence désormais irrémédiable d'une connaissance fine et partagée de la langue latine, plus aucun élément structurant ne vient donner aux centaines de millions de locuteurs des langues apparentées que sont les langues romanes la conscience que ces langues sont toutes des variantes d'une même famille, en un mot, les dialectes d'une même langue ».

Il devrait être possible, dans des dispositifs plurilingues, de donner une visibilité à ces éléments structurants, et pas seulement pour les étudiants spécialistes des langues. Les approches sur corpus semblent le faciliter.

## **3.4 Apport des approches sur corpus pour la connaissance des analogies linguistiques**

### **3.4.1 Définitions d'une discipline et d'un outil**

Selon Williams (1999, introduction du chapitre I), la linguistique de corpus est « le mariage entre la linguistique appliquée, surtout dans les domaines de l'enseignement des langues de spécialité et la lexicographie, et l'informatique ». Ce mariage présente au moins deux intérêts. Un intérêt social : permettre que des textes souvent réunis par leur contenu thématique soient interrogeables par un public donné. Un intérêt linguistique : permettre de déterminer les particularités d'une langue de spécialité, de la langue scientifique, etc.

Rastier (2002) définit ainsi le mot *corpus* :

« Un corpus est un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés : (i) de manière théorique réflexive en tenant compte des discours et des genres, et (ii) de manière pratique en vue d'une gamme d'applications. »

Nous avons utilisé dans notre recherche plusieurs corpus que nous présentons plus loin. Les approches sur corpus, que nous avons découvertes pendant cette thèse, ont remis en question notre façon d'aborder une langue par les possibilités immenses qu'elles permettent dans une perspective lexicographique et/ou didactique. Les corpus, qui sont constitués aujourd'hui couramment de dizaines de millions de mots, et au-delà, permettent des comparaisons intralinguistiques, et interlinguistiques lorsqu'ils sont multilingues.

### **3.4.2 Distinction entre corpus comparables et corpus parallèles**

Parmi les corpus multilingues, on peut distinguer les corpus comparables et les corpus parallèles.

Selon Teubert (1996 : 245),

« Des « corpus comparables » sont des corpus en deux langues ou plus composés de façon identique ou similaire. Les textes qu'ils contiennent peuvent être classés selon une

variété de traits intralinguistiques et extralinguistiques. Le domaine, par exemple, peut être une caractéristique pertinente pour la composition du corpus. »

Ainsi, bien des documents qui constituent les corpus comparables sont des « on-topic documents » (Fung & Cheung, 2004 : 1052), autrement dit des documents qui traitent du même thème. Ces textes ne sont pas une traduction les uns des autres. Au contraire, un corpus parallèle est un « corpus de phrases alignées qui contient des traductions bilingues du même document ». Dans un corpus parallèle, chaque texte est segmenté en phrases, les phrases de la langue source correspondent plus ou moins avec leur traduction dans la langue cible. Le corpus est dit parallèle justement du fait de cet alignement phrastique, dont l'image 6 fournit un exemple<sup>123</sup>.

---

Hubo complicaciones que no pudimos tratar.

There were complications we couldn't deal with.

---

*Image 6 : Capture d'écran d'un exemple de phrases alignées sur Reverso context*

Parmi les trois types de phrases alignées que distingue Đõ (2011 : 67), nous nous intéressons aux phrases comparables, celles qui ne sont pas exactement la traduction l'une de l'autre mais qui contiennent des fragments parallèles : nous les recherchons dans des corpus multialignés (Kraif, 2015). Un corpus multialigné est un corpus multiparallèle formaté de manière à ce que chaque unité (souvent, il s'agit de phrases) soit alignée dans plusieurs langues en même temps. Les corpus alignés permettent d'étudier le lexique et le fonctionnement syntaxique de plusieurs langues en même temps.

Les corpus et les outils dont on se sert pour les interroger permettent un enseignement des langues qui tire sa légitimité de la fréquence d'emploi dans des documents authentiques. Ils conduisent également à mieux comprendre le fonctionnement du discours, par exemple dans notre cas le discours de spécialité.

---

123 <http://context.reverso.net/traducao/espanhol-ingles/tratar>. Reverso Context est un concordancier.



### 3.4.3 Vers une compréhension plus fine du discours de spécialité

#### 3.4.3.1 *Le lexique témoin d'une communauté de discours*

Nous avons utilisé un corpus, ScienText, que nous décrivons plus loin. Nous nous sommes intéressé à ses mots fréquents et à la dimension phraséologique du discours scientifique (Williams, 1999, Tutin, 2008, etc.), dont les usages sont ancrés dans une communauté de discours.

Swales (1990 : 24-27) a défini des critères qui permettent de considérer qu'un ensemble de personnes constituent une communauté de discours. Au sujet des usages en vigueur au sein d'une communauté de discours, il pointe la double nature du lexique de spécialité, tantôt emprunté à la langue courante, tantôt hyper spécialisé :

« Cette spécialisation peut impliquer l'utilisation d'unités lexicales connues de la communauté de discours plus large, comme dans le discours des communautés de la technologie de l'information, ou l'utilisation d'une terminologie très technique comme dans les communautés médicales »<sup>124</sup> (ibid. : 26).

Les frontières dans les usages ne sont plus là géographiques (cf. 2.3.1.2, p. 83), mais sociales, et on peut étudier des particularités d'un lexique qui reflète les usages communs aux disciplines scientifiques.

#### 3.4.3.2 *Classer le lexique des écrits scientifiques*

Dans ses travaux, le CREDIF (Centre de Recherches et d'Études pour la Diffusion du Français) désigne le lexique utilisé dans les écrits scientifiques comme le Vocabulaire Général d'Orientation Scientifique (VGOS). Selon Phal (1971, 9) :

« Le vocabulaire scientifique général est [...] commun à toutes les spécialités. Il sert à exprimer les notions élémentaires dont elles ont toutes également besoin (mesure, poids, rapport, vitesse, etc.) et les opérations intellectuelles que suppose toute démarche méthodique de la pensée (hypothèse, mise en relation, déduction et induction, etc.) ».

---

124 « This specialization may involve using lexical items known to the wider speech community in special and technical ways, as in information technology discourse communities, or using highly technical terminology as in medical communities »

Pecman (2005) va au-delà du seul vocabulaire et définit la langue scientifique générale (LSG) comme la « pratique langagière spécifique à une communauté de discours composée de chercheurs en sciences exactes et sciences humaines dont les objectifs communicatifs émanent de préoccupations partagées par des scientifiques à travers le monde et indépendamment de leurs spécificités disciplinaires ». Granger (2010) pointe la dimension universitaire de cette communauté, avec l'appellation *general academic vocabulary*.

Les trois auteurs mentionnés ont en commun de désigner ce lexique ou cette langue scientifique au moyen de l'adjectif *général*. Tutin (2007 : 7) envisage dans un texte scientifique « plusieurs strates lexicales plus ou moins étanches allant du lexique général à la terminologie du domaine », et utilise à la suite de Drouin (2007) une autre désignation : le lexique scientifique transdisciplinaire (LST) : « le lexique scientifique transdisciplinaire ne renvoie pas aux objets scientifiques des domaines de spécialité mais au discours sur les objets et les procédures scientifiques » (Tutin, *ibid.* : 6). Dans le texte scientifique qu'elle examine, Tutin distingue le lexique propre aux écrits scientifiques, le lexique abstrait non spécialisé, du lexique méthodologique disciplinaire, le lexique terminologique et le lexique de la langue « générale » ou « commune ».

Hatier (2016 : 16) revient sur ces lexiques de l'écrit scientifique et souligne :

« les frontières sont complexes à tracer entre ces différents lexiques, particulièrement entre le LST et un ensemble lexical composé de mots abstraits également mobilisé dans la langue générale ».

Nous avons raisonné dans notre recherche au niveau de l'acception et non au niveau de la lexie. En butte à des difficultés de classement de certaines acceptions, nous avons distingué deux catégories sémantiques : le lexique de spécialité de la santé (abrégé dans les tableaux par lexique de la santé) et le lexique scientifique transdisciplinaire (LST), que nous entendons néanmoins différemment de Tutin :

- le lexique de spécialité de la santé regroupe les acceptions du lexique terminologique de la santé et celles du lexique méthodologique disciplinaire de la santé ;
- le LST regroupe toutes les acceptions qui ne relèvent pas du lexique de spécialité de la santé : il comprend le lexique sur les objets et les procédures scientifiques, le lexique abstrait non spécialisé et le lexique général.

Nous incluons dans le lexique spécifique de la santé le lexique méthodologique disciplinaire parce qu'il renvoie lui aussi au discours sur les objets et les procédures, quand bien même il réfère aux méthodes propres à la discipline - ici, la santé.

Nous avons gardé l'appellation LST parce qu'elle exprime l'idée d'un dénominateur commun aux textes scientifiques et les organisant indépendamment de la spécialité dont ils traitent. Nous incluons dans le LST le lexique abstrait non spécialisé parce que la distinction entre un lexique qui serait propre aux écrits scientifiques et un lexique abstrait non spécialisé nous paraît artificielle : puisque le LST est transdisciplinaire, il n'est justement pas spécialisé. Il devrait par conséquent inclure les mots du discours sur les objets et les procédures scientifiques qui introduisent les causes de la recherche (*poser un problème, la difficulté joue, influence*)<sup>125</sup>, les mots qui permettent la description de l'objet (*élément*), et encore les enchaînements logiques (*conduire à*). Nous incluons également dans le LST le lexique de la langue 'générale' ou 'commune', parce qu'à notre sens, le lexique commun aux textes scientifiques, bien qu'il soit « plutôt un lexique abstrait particulièrement représenté dans les écrits scientifiques » (Tutin, 2007 : 8), contient une dimension très concrète.

Nous avons classé les acceptions du LST en sept sous-catégories : la connaissance, les dynamiques sociales, les orientations de la recherche, la description des caractéristiques, la démarche scientifique, la mesure, l'interprétation. La section Sous-catégories du LST, p.156, donne des exemples de mots anglais que nous avons classés dans chacune. Pour dresser cette liste, nous avons trié 2008 acceptions, comme nous l'expliquons section 4.5, p.152. Les mots triés par sous-catégories figurent dans les annexes Annexe 5 : Liste du lexique spécifique de la santé, p. 398 et Annexe 6 : liste du lexique scientifique transdisciplinaire, p.403.

Les éléments de ce lexique fréquent peuvent être étudiés en texte afin de faire émerger des séquences collocatives propres à ce type de discours et les incorporer dans une didactisation qui mette en valeur la dimension phraséologique de ces écrits.

### **3.4.3.3 Une phraséologie des séquences collocatives analogues**

González Rey (2008 : 5) définit la phraséologie ainsi :

« Placée à mi-chemin entre le lexique et la syntaxe, la phraséologie s'occupe des signes polylexicaux composés d'au moins deux mots, stables, répétés et souvent figurés. »

---

<sup>125</sup> Ces exemples sont tirés de l'article de Tutin (2007).

L'article scientifique, tout en rendant publiques de nouvelles données recueillies et analysées, emprunte aux textes appartenant au même discours des mots et des tournures parfois si répandues qu'elles s'apparentent au figement. Certaines de ces tournures figées se retrouvent dans plusieurs langues.

González Rey (2007) propose de classer les expressions figées en trois catégories : les collocations, les expressions idiomatiques et les parémies. Nous nous intéressons uniquement aux collocations, c'est-à-dire aux associations préférentielles entre certains mots. Si les mots qui composent les collocations (nous préférons l'expression séquences collocatives) sont analogues, nous parlons de séquences collocatives analogues.

À l'issue de l'élaboration de l'interlexique, nous avons cherché dans des corpus multialignés des séquences collocatives analogues dans les cinq langues, ou dans au moins un couple de langue anglais – langue romane. Les séquences collocatives analogues dans plusieurs langues romanes nous intéressent dans la perspective de concevoir des activités didactiques plurilingues car nous pensons qu'elles peuvent faciliter l'accès au sens du français dans les cas où le français ne présente pas d'analogie avec l'anglais.

Dans sa méthode de phraséodidactique de FLE, González Rey (2012) propose d'enseigner le FLE tel qu'on le parle. À notre tour, nous suggérons dans la section Perspectives : vers un dispositif didactique plurilingue d'enseignement du FOS sur la base d'analogies interlinguistiques, p.247, des pistes qui aillent dans le sens d'une prise en compte des usages réels du domaine de spécialité de la santé, à l'écrit. Nous nous appuyons pour cela sur les connaissances supposées des apprenants en anglais, sur un questionnement métalinguistique et sur les analogies interlinguistiques identifiées dans le corpus ScienText.

## 3.5 Présentation du corpus ScienText utilisé pour l'étude lexicologique

### 3.5.1 Origine et principes de constitution de ScienText

Le corpus que nous avons utilisé, qui regroupe des articles dont les thématiques sont liées à la santé, est en réalité un sous-corpus. Il fait partie d'un projet plus vaste : ScienText.

Le projet ANR ScienText a eu lieu de 2007 à 2010. Des équipes pluridisciplinaires de chercheurs en linguistique, informatique, et didactique de six laboratoires ont élaboré un corpus de textes scientifiques constitué de quatre (sous-)corpus. Ces quatre corpus regroupent des textes scientifiques en français et en anglais : le corpus français de textes scientifiques, le corpus anglais de textes scientifiques, le corpus d'apprenants d'anglais, et le corpus d'évaluations de communications<sup>126</sup>. Le logiciel Syntex (Didier & al., 2005) a annoté structurellement ces textes au niveau de l'entête, de la structure, de la morphologie et de la syntaxe.

L'objectif du projet ScienText était double : l'étude des marques du positionnement de l'auteur et « la constitution d'un corpus représentatif d'écrits scientifiques annotés linguistiquement et d'un ensemble d'outils permettant d'interroger en ligne ce corpus » (Falaise et al., 2011: 105). Une interface facile à utiliser a été conçue pour les non-spécialistes du TAL, qui sont souvent rebutés par le langage de requête complexe utilisé par les spécialistes. Une enquête de besoins a été menée, un prototype élaboré et évalué à plusieurs reprises par des usagers.

Pour la présente recherche, nous avons recouru à l'un des quatre sous-corpus, le corpus de textes scientifiques anglais de ScienText, qui regroupe 35,2 millions de mots. Le laboratoire LiCorn de l'Université de Bretagne Sud (Williams & Millon, 2009) y a rassemblé 7564 textes libres de droits tirés de 136 revues publiées par l'éditeur BioMedCentral et relevant du domaine de la santé : médecine, biologie et génétique.

### 3.5.2 Comment requêter à partir de ScienQuest

Nous avons adopté une approche *corpus-driven*, à savoir « une démarche inductive partant du corpus pour aller vers une éventuelle modélisation des régularités émergentes des données authentiques traitées » (Tognini-Bonelli, 2001 : 13).

---

<sup>126</sup> Pour une présentation plus détaillée des sous-corpus, voir Grossmann et Tutin (2013 : 16-17).

La plateforme web ScienQuest<sup>127</sup>, développée par Falaise (Falaise, Tutin & Kraif, 2011), est l'outil qui permet d'interroger grâce à son interface guidée les quatre corpus élaborés pendant le projet ScienText, mais aussi certains autres corpus. Il s'appuie sur le moteur de recherche ConcQuest développé par Olivier Kraif (2008), et donnait accès en 2016 à la version 1.8. de ScienText.

Sur la plateforme ScienQuest, une requête portant sur le corpus ScienText s'effectue en trois ou quatre étapes. Nous détaillons ci-après la démarche parce que l'abord de ces outils est malaisé pour un novice.

1) Dans la page d'accueil, l'utilisateur choisit le corpus.

2) Éventuellement, il peut dans le menu de gauche (Textes, puis Critères) sélectionner certaines parties textuelles (introduction, etc.). Il peut cocher certains textes seulement dans la rubrique Liste, qui mentionne pour chaque texte son titre, ses auteurs, son sous-domaine, le nombre de mots qu'elle contient et le type de licence légale sous laquelle l'œuvre est mise à disposition.

3) Dans Recherche, il choisit une forme lemmatisée et/ou un mot, et/ou une catégorie grammaticale, éventuellement sous forme d'expression régulière, en association ou non avec d'autres mots. La recherche peut être restreinte à un nombre donné de mots du corpus et le contexte limité à un nombre de mots avant et après le mot-pivot (keyword). Un tableau de concordance est généré qui permet de visualiser le mot-pivot et son contexte (KWIC, pour KeyWord in Context). Cliquer sur une ligne permet l'affichage du contexte dans le texte source et l'arbre syntaxique. Le tableau est exportable sous les formats CSV, HTML et XLSX. ScienQuest offre aussi la possibilité d'une recherche en mode sémantique ou en mode avancé<sup>128</sup>.

4) Le bouton Résultats conduit à l'affichage des statistiques calculées par le concordancier par forme, mot, catégorie, traits morphosyntaxiques et partie textuelle. L'utilisateur peut rapatrier ces résultats et les retravailler dans son tableur.<sup>129</sup> Grâce à cette fonctionnalité, les données lexicales peuvent être travaillées par fréquences.

---

127 <http://corpora.aiakide.net/scientext18>

128 [http://corpora.aiakide.net/scientext18/static/doc/Scientext\\_grammaires\\_v5.pdf](http://corpora.aiakide.net/scientext18/static/doc/Scientext_grammaires_v5.pdf) explique la syntaxe des grammaires utilisables pour le mode avancé. Cette syntaxe a été développée par Achille Falaise à partir de la syntaxe de ConcQuest mise au point par Olivier Kraif (2008).

129 Un tutoriel a été rédigé par Achille Falaise et Agnès Tutin : [http://corpora.aiakide.net/scientext18/static/doc/mde\\_sciquest\\_2016.pdf](http://corpora.aiakide.net/scientext18/static/doc/mde_sciquest_2016.pdf). On peut aussi se référer à Falaise O. et al. (2011 : 118-123).

### 3.6 Conclusion du chapitre 3

Nous avons vu dans ce chapitre comment les analogies interlinguistiques sont valorisées par des approches didactiques innovantes qui contribuent au développement du plurilinguisme et/ou renouvellent la connaissance du lexique.

Au lieu de voir dans les analogies interlinguistiques le risque d'interférences linguistiques, l'enseignement gagnerait à créer des conditions qui stimulent les rapprochements entre les préacquis d'un individu et la langue cible. Les transferts peuvent procéder d'analogies entre la langue cible et la L2 (et L3), qui sert alors de langue passerelle, surtout si l'apprenant est amené à percevoir ces analogies. Des approches contribuent à accentuer l'intelligibilité entre des langues voisines, l'initiation au français après l'anglais pourrait entrer dans ce cadre.

La proximité linguistique des langues, le contact avec la langue authentique et la découverte simultanée de plusieurs langues constituent l'originalité des approches de l'intercompréhension, qui mettent en lumière les analogies lexicales, syntaxiques, et morphosyntaxiques en développant la compétence métalinguistique des apprenants. L'initiation au français (FLE ou FOS) par une approche qui soulignerait systématiquement ses analogies avec d'autres langues, dont l'anglais et les langues romanes, replacerait un système dans un intersystème.

Par son approche systématique des analogies linguistiques en contexte, la linguistique de corpus présente également un potentiel énorme pour la didactique d'une langue et la didactique du plurilinguisme. D'une part, l'étude de la langue dans sa dimension phraséologique permet à la fois de formuler un usage général et d'approcher des traits (sémantiques, syntaxiques etc.) mal connus. D'autre part, le travail sur corpus alignés donne l'impression d'une certaine égalité entre les langues, mises en parallèle dans leur fonctionnement (leur richesse, leur imprécision, leur manque).

Aujourd'hui, des documents authentiques sont accessibles en ligne dans potentiellement toutes les langues<sup>130</sup>, si bien que des enseignants formés à ces approches pourraient aborder avec leurs apprenants des langues qui présentent des analogies avec la langue de scolarité et/ou la langue cible et des systèmes linguistiques qui traduisent autrement le monde et le rapport au monde.

---

130 Par exemple dans les versions de l'encyclopédie collaborative en ligne [Wikipédia](#).

Dans une classe, une approche contrastive permet en effet de prendre en compte le répertoire langagier et symbolique des apprenants, et de l'élargir à d'autres langues et cultures afin de développer un type de plurilinguisme ouvert aux autres. C'est l'esprit des programmes de l'école et du collège du Ministère français de l'Éducation nationale, de l'enseignement supérieur et de la recherche :

« L'apport d'informations, la comparaison, la mise en perspective, l'expérience des échanges et de la mobilité permettent à l'élève de découvrir les spécificités liées aux langues apprises, d'en percevoir les diversités et en les rapprochant entre elles et avec sa propre culture, d'en appréhender les convergences de façon à favoriser une meilleure intercompréhension entre les langues et les cultures »<sup>131</sup>.

Au niveau universitaire, nous pensons que l'enseignement du FOS pourrait tirer partie des analogies qui existent entre le français et l'anglais étant donné que les apprenants de français ont généralement appris l'anglais précédemment.

Nous avons utilisé le corpus ScienText afin d'obtenir le lexique anglais fréquent que nous avons traduit en espagnol, français, italien et portugais. Nous avons ensuite comparé la forme des mots pour en tirer l'interlexique. La méthode suivie est détaillée dans la deuxième partie.

---

131 <http://eduscol.education.fr/pid36086/la-competece-culturelle-en-langue-vivante.html>



## **Deuxième partie - Analogies des mots fréquents du discours scientifique de la santé : méthodologie d'élaboration d'un interlexique en 5 langues à partir du corpus ScienText**

L'interlexique anglais-espagnol-français-italien-portugais, dont nous retraçons dans cette partie les étapes de l'élaboration, reflète des analogies lexicales qui existent entre l'anglais et ces quatre langues romanes dans le discours scientifique de la santé. Cependant, seuls les mots les plus fréquents du corpus de la santé de ScienText ont été examinés, si bien que de très nombreux mots (des milliers) ne font pas partie de cette étude. En effet, non seulement le temps dans lequel cette recherche devait avoir lieu imposait de restreindre ses ambitions, mais surtout l'objectif lexicologique était avant tout de délimiter, puis de traduire des mots qu'un apprenant d'anglais sur objectifs spécifiques de la santé devrait avoir mémorisé parce qu'il les rencontre régulièrement dans son domaine de spécialité. C'est sur ces connaissances présumées que devraient être à l'avenir conçues et expérimentées des activités plurilingues de découverte du FLE.

La méthodologie à laquelle nous avons recouru dans ce chapitre est en quelque sorte hybride. Il nous a été nécessaire de croiser des approches complémentaires parce que les corpus reflètent les usages d'une communauté de discours. Or, nous ne sommes pas spécialiste de la santé, si bien que nous manquons de références pour asseoir une interprétation. Par conséquent, l'approche suivie est tout autant empirique qu'hypothético-déductive. Elle se fonde d'une part sur des descriptions issues d'ouvrages de référence, d'autre part sur l'introspection lorsque le discernement a été nécessaire pour résoudre certains cas.

Deux chapitres constituent cette deuxième partie. Le chapitre 4, p.130, traite du lexique anglais fréquent de la santé dans le discours scientifique. Nous y exposons la méthodologie suivie afin de délimiter les 2000 mots fréquents du corpus ScienText (p.135), cerner leur sens (p.143) et répartir en deux catégories sémantiques les 2208 acceptions identifiées (p.152).

Le chapitre 5 détaille les étapes de traduction (p.162) et de comparaison de la forme des mots dans les cinq langues (p.187), des étapes qui ont mené à d'élaboration de l'interlexique en 5 langues et des quatre lexiques bilingues par paires de langues anglais – langue romane.

# Chapitre 4 - Le lexique anglais dans le discours scientifique de la santé : délimitation des 2000 adjectifs, noms et verbes fréquents du corpus ScienText et de leurs 2208 acceptions fréquentes dans le corpus

## Sommaire du chapitre 4

4.1 Introduction du chapitre 4.....	130
4.2 Considération préalable : pourquoi ne pas avoir compilé des lexiques existants....	132
4.2.1 Défaut de critères de sélection des mots-vedettes dans les dictionnaires.....	132
4.2.2 Une évolution notoire : des dictionnaires bilingues qui définissent dans la langue source.....	132
4.2.3 Notre approche : le lexique et les acceptions fréquents.....	133
4.3 Constitution d'une liste des adjectifs, noms et verbes anglais fréquents du corpus ScienText.....	135
4.3.1 Plusieurs extractions nécessaires.....	135
4.3.2 Écrémage des mots.....	136
4.3.3 Des erreurs d'étiquetage liées à des particularités linguistiques de l'anglais... ..	138
4.3.4 Sélection des 2000 mots les plus fréquents.....	141
4.4 Détermination du sens des 2000 mots anglais fréquents.....	143
4.4.1 Évolutions de l'approche suivie pour cerner le sens.....	143
4.4.2 Comment définir les 2208 acceptions identifiées.....	148
4.5 Deux catégories sémantiques pour les 2208 acceptions fréquentes.....	152
4.5.1 Un tri manuel en deux catégories : lexique de spécialité et LST.....	152
4.5.2 Sous-catégories du lexique de spécialité de la santé.....	153
4.5.3 Sous-catégories du LST.....	156
4.6 Conclusion du chapitre 4.....	158

## 4.1 Introduction du chapitre 4

Afin d'étudier des analogies interlinguistiques qui existent entre l'anglais, l'espagnol, le français, l'italien et le portugais, plusieurs options s'offraient. Compiler des ressources bilingues / plurilingues<sup>132</sup> existantes pour en tirer des mots dans les langues qui nous

---

132 Par exemple celui de M. Azzaretti (2008).

intéressent, dont nous aurions comparé la forme. Nous avons écarté cette possibilité pour élaborer nous-même des données et étudier le lexique fréquent dans le discours. Le corpus d'écrits scientifiques de la santé de ScienText a été utilisé à cette fin, auquel ont contribué des chercheurs de notre laboratoire, le LIDILEM.

Le présent chapitre 4 comporte quatre subdivisions. La section 4.2 (Considération préalable : pourquoi ne pas avoir compilé des lexiques existants, p.132) montre pourquoi nous avons constitué des données à partir d'un corpus.

Dans la section 4.3 (Constitution d'une liste des adjectifs, noms et verbes anglais fréquents du corpus ScienText, p.135), nous rendons compte de plusieurs extraction successives des adjectifs, des noms et des verbes fréquents dans le souci d'obtenir des données plus fiables, à mesure qu'étaient rendues accessibles les versions améliorées du corpus. Un écrémage s'est imposé afin d'obtenir les 2000 mots fréquents : les données ont été rectifiées et homogénéisées. Ce nettoyage est motivé par les erreurs d'étiquetage dues pour partie à des particularités de la langue anglaise.

Dans la section 4.4 (Détermination du sens des 2000 mots anglais fréquents, p.143), nous expliquons comment le sens a été délimité et les acceptions fréquentes définies.

Enfin, les 2208 acceptions ont été triées selon qu'elles se rattachent au lexique de spécialité de la santé ou au LST. Ce processus fait l'objet de la section 4.5 (Deux catégories sémantiques pour les 2208 acceptions fréquentes, p. 152).

## **4.2 Considération préalable : pourquoi ne pas avoir compilé des lexiques existants**

### **4.2.1 Défaut de critères de sélection des mots-vedettes dans les dictionnaires**

Des dictionnaires unilingues, bilingues ou plurilingues de qualité existent dans le domaine de la santé. Mais certains termes « directement compréhensibles car très proches en anglais et en français » (Duizabo, 2009 : 5) en sont parfois absents, malgré certaines nuances.

La question du choix des mots-vedettes des dictionnaires spécialisés de la santé peut également se poser. Ainsi, on peut s'étonner que des termes comme (eng) *recovery*, (fra) *récupération*, *rétablissement* ; (eng) *regulatory*, (fra) *régulatoire* ; (eng) *release*, (fra) *libération* ; (eng) *removal*, (fra) *prélèvement* ; (eng) *responsive*, (fra) *sensible* ; (eng) *sample*, (fra) *échantillonner* ; ou encore (eng) [DNA] *strand*, (fra) *brin* [d'ADN], ne soient pas des mots-vedettes du Dorland (Maniez, 2009) dans sa version bilingue.

Une évolution s'est cependant produite qui estompe la dichotomie entre les dictionnaires unilingues et les dictionnaires bilingues publiés en français.

### **4.2.2 Une évolution notoire : des dictionnaires bilingues qui définissent dans la langue source**

Il est fréquent qu'un usager recoure à un ouvrage différent selon qu'il cherche la traduction d'un mot dans une langue étrangère ou une définition de ce mot dans sa langue. Ainsi, les dictionnaires bilingues sont utilisés dans la perspective de trouver un équivalent de traduction tandis que les dictionnaires spécialisés ou les encyclopédies le sont probablement davantage avec l'intention de vérifier ou d'approfondir sa compréhension d'un point particulier. On retrouve cette situation dans les usuels de la santé. Selon Maniez (2009 : vii) :

« les dictionnaires médicaux destinés aux francophones continuent cependant de se diviser en deux catégories bien distinctes : ceux dont le rôle est essentiellement de donner à l'utilisateur une définition des termes choisis, et ceux qui en fournissent la traduction dans une langue donnée ».

Dans la première catégorie, on compte notamment le *Dictionnaire de médecine Flammarion* (Kernbaum, 2208), le *Dictionnaire médical de poche* (Quevauvilliers, 2007), le *Dictionnaire*

*illustré des termes de médecine* (Garnier et al., 2012), la partie mini-dictionnaire de l'*Anglais médical* (Gross, 2009 : 217-243), qui proposent une traduction après le mot-vedette.

Par contraste, le *Dorland Dictionnaire médical bilingue français-anglais / anglais-français* (Maniez, *ibid.*) a ceci de particulier qu'il combine les deux approches : l'utilisateur consulte au même endroit la définition d'un mot-vedette dans la langue de ce terme et sa traduction dans la langue cible. Aussi, un mot-vedette anglais est défini en anglais et traduit en français. Un mot-vedette français est défini en français et traduit en anglais. Toutes les définitions sont tirées de ou traduites de la 28<sup>e</sup> édition du dictionnaire de poche Dorland. Ainsi, l'apprenant peut «se familiariser non seulement avec la terminologie, mais également avec la phraséologie bilingue du domaine biomédical » (*ibid.* : vii).

Nous avons pensé qu'il était utile que les mots anglais fréquents dont nous avons établi une liste soient systématiquement définis. La définition, en anglais, est la plupart du temps empruntée à un dictionnaire généraliste collaboratif, le *Wiktionary*. Nous expliquons pourquoi dans la section Le choix du Wiktionary, p.148. Nous avons aussi consulté l'*Oxford English Dictionary*<sup>133</sup> dans les cas où une acception manquait dans le *Wiktionary*.

### **4.2.3 Notre approche : le lexique et les acceptions fréquents**

Nos données reposent sur un type de discours donné : le discours scientifique de la santé écrit en anglais. La notion d'usage implique l'idée que certains mots et certaines tournures seraient employées de préférence à d'autres. Dans un domaine de spécialité, il va de soi qu'une terminologie spécifique à la discipline est employée. En outre, on peut observer une variation diastratique et certaines séquences collocatives propres au discours de spécialité.

Pour cerner le lexique en usage, lexique de spécialité ou non, nous avons recouru au critère de la fréquence à deux étapes de notre recherche.

Premièrement, lors de la délimitation d'un ensemble de 2000 mots anglais - adjectifs, noms et verbes. Deuxièmement, lorsque nous en avons délimité les acceptions fréquentes.

Il nous faut à ce stade justifier pourquoi, dans l'établissement de nos données, nous nous sommes limité à la catégorie des mots lexicaux (Teyssier, 2012 : 58), sans toutefois incorporer les adverbes.

---

133 <http://www.oed.com/>

À la suite de Vietri (2004 : 159), nous avons considéré que

« les éléments lexicaux dotés de signification pleine ne sont pas seulement les verbes, simples ou composés, mais également les noms et les adjectifs »<sup>134</sup>.

Si la notion de signification pleine peut être discutée, l'adverbe induit souvent des « degrés d'adhésion subjective à un jugement » (Gosselin, 2010 : 168), or la modalité n'est pas l'objet de notre étude.

La part des trois catégories grammaticales des 2000 adjectifs, noms et verbes anglais fréquents du corpus ScienText est examinée dans les résultats, dans la section Catégories grammaticales des 2000 mots anglais fréquents, p.203.

Nous avons également appliqué le critère de la fréquence lors de la délimitation des acceptions des 2000 mots en contexte (cf. Détermination du sens des 2000 mots anglais fréquents, p.143). Comme on peut s'y attendre, certains mots sont polysémiques ; comment alors savoir dans laquelle / lesquelles de ses acceptions un mot est employé ? Pour cela, nous avons examiné le profil combinatoire de chacun des 2000 mots (cf. Seconde tentative : déterminer le sens par le profil combinatoire, p.144). Les 2208 acceptions fréquentes filtrées ont alors été définies et nous avons cherché les équivalents de traduction des 2000 mots fréquents pour leurs 2208 acceptions fréquentes. Cette procédure est détaillée p.162 (Traduction des 2000 mots anglais fréquents en espagnol, français, italien et portugais).

À l'évidence, la liste obtenue n'est pas exhaustive, et dans certains sous-domaines, on pourra objecter qu'elle n'a pas une meilleure couverture que les dictionnaires de la santé. De surcroît, elle sera rapidement datée vu les avancées de la science dans ce domaine. Les 2000 unités retenues dans cette étude sont donc seulement les plus fréquentes dans un type de discours, à un moment donné et dans un corpus de spécialité donné. Leurs 2208 acceptions sont les plus fréquentes que nous avons identifiées.

---

134 Notre traduction.

## **4.3 Constitution d'une liste des adjectifs, noms et verbes anglais fréquents du corpus ScienText**

### **4.3.1 Plusieurs extractions nécessaires**

#### ***4.3.1.1 Adaptation à un corpus évolutif***

Dans ScienQuest, les fichiers de données peuvent être obtenues par catégories grammaticales. Nous avons d'abord extrait les fichiers des adjectifs, des noms et des verbes à partir de la version 1.4. de ScienText. Cependant, afin d'améliorer la fiabilité des données, nous avons jugé nécessaire de reconsidérer à deux reprises les résultats que nous avons obtenus.

Nous avons recommencé le processus d'extraction et d'écrouissage des mots une première fois, lorsque la version 1.8. de ScienText a été interrogeable sur le site ScienQuest. Puis nous sommes repassé par les mêmes étapes une seconde fois lorsque nous avons constaté que le site de ScienQuest n'interrogeait pas encore l'intégralité des textes du corpus. En effet, nous souhaitons que notre étude reflète au plus près l'usage des mots anglais dans le discours scientifique de la santé. En conséquence, il était plus consistant de faire porter notre étude sur l'ensemble des 35,2 millions d'occurrences du corpus. La répétition des opérations portant sur ces différentes versions a été particulièrement coûteuse en temps.

Certains fichiers de travail élaborés lors de phases successives seront accessibles sur le portail hal, dans les documents de l'auteur de cette thèse.

#### ***4.3.1.2 Première vérification de l'étiquetage des mots à 1 000 occurrences***

Nous sommes parti des données brutes extraites qui correspondent aux lemmes étiquetés comme adjectifs, noms et verbes. Ces données comportent de nombreuses erreurs en raison de l'étiquetage automatique par TreeTagger (Schmid, 1995). Ainsi, il existerait 93 864 mots différents qui sont des adjectifs - un nombre considérable. Bien entendu, nous ne pouvions pas rectifier toutes les erreurs d'étiquetage, mais il fallait une base aussi fiable que possible des mots fréquents.

Nous avons tout d'abord vérifié l'étiquetage des lemmes d'adjectifs, de noms et de verbes dont le nombre d'occurrences s'élève au moins à 1 000.



L'opération de vérification s'est effectuée, pour les mots courants, à partir de nos connaissances en anglais. L'appartenance des textes du corpus au champ de la santé nous a cependant parfois conduit à penser que certains avaient davantage de probabilités d'appartenir à une autre catégorie grammaticale. Des requêtes dans le corpus ont permis de vérifier les mots qui nous étaient inconnus.

Une deuxième étape a consisté à ne conserver que les 2 000 mots les plus fréquents, indépendamment de leur catégorie grammaticale, en prenant en compte leur nombre d'occurrences (cf. Sélection des 2000 mots les plus fréquents, p.141).

## 4.3.2 Écrémage des mots

### 4.3.2.1 Écrémage des adjectifs

Voici un extrait d'un document de travail, la liste des adjectifs en cours d'élaboration. Elle contient deux mots qui avaient été étiquetés comme adjectifs et que nous avons écrémés : *24* et *ii*.

following	6528
able	6459
24	6426
structural	6419
mutant	6310
random	6280
vascular	6238
central	6229
ii	6215
appropriate	6208

*Tableau 3 : Extrait de la liste des adjectifs en cours de nettoyage*

Parmi les mots étiquetés comme adjectifs ont été écartés :

- les chiffres et nombres ;
- les déterminants numéraux cardinaux (*seven, million, etc.*) ;

- les autres déterminants et pronoms ;
- les combinaisons de chiffre(s) et de lettre(s) (*brca1, p27, etc.*) ;
- les suites de lettres utilisées comme abréviations ou symboles (*iii, ca, etc.*) ;
- les noms (*patient, baseline, etc.*) ;
- les formes verbales et adverbes
- les autres catégories : préfixes, prépositions.

#### 4.3.2.2 *Écrémage des noms*

Parmi les mots étiquetés comme noms, nous avons écrémé les mots qui étaient en réalité :

- des chiffres, nombres et combinaisons de chiffres et de lettres ;
- des suites de lettres qui ne sont pas des noms attestés (*kda, etc.*) ;
- des adjectifs ou pronoms incorrectement étiquetés, et des éléments préfixés (*multi, meta, etc.*) ;
- des noms propres étiquetés comme noms communs (noms de marques, par exemple).

En revanche, nous avons retenu les sigles usuels tels que *DNA*, ainsi que les abréviations d'unités de mesure internationales (*kg, ml, etc.*). Nous signalons en outre que nous avons rectifié certaines erreurs telles que *genomic*, qui renvoie en réalité au nom *genomics*.

#### 4.3.2.3 *Écrémage des verbes*

Parmi les mots étiquetés comme verbes, certains étaient en réalité :

- des noms ;
- des suites de lettres qui ne sont pas des verbes attestés ;
- des formes lemmatisées d'un verbe ;
- des adjectifs ;
- des adverbes ;
- des suffixes (*-like, -loaded, etc.*).

#### 4.3.2.4 *Écrémage de mots ambigus ou redondants*

Certains mots réfèrent en réalité à des mots distincts. Ainsi, le lemme *mister* réfère aux occurrences de *Mr*; mais aussi à celles de *MR* (Magnetic Resonance). Nous avons supprimé ce type de lemmes. En effet, nous ignorons si le nombre d'occurrences de chacun des mots suffit pour qu'ils figurent parmi les mots fréquents.

En outre, certains lemmes ont été supprimés en raison de leur étiquetage erroné. Ainsi, on trouvait le lemme *sera*. Les formes *sera* et *serum* renvoient dans le corpus à deux lemmes distincts alors qu'il ne s'agit que d'un mot, *sera* étant le pluriel de *serum*. Nous n'avons gardé qu'un lemme, le singulier *serum*, nous conformant ainsi à l'usage des dictionnaires, qui utilisent le singulier comme mot-vedette, et nous avons additionné les occurrences de *serum* à celles de *sera*.

### 4.3.3 Des erreurs d'étiquetage liées à des particularités linguistiques de l'anglais

#### 4.3.3.1 *Erreurs liées au faible nombre de flexions en anglais*

La morphosyntaxe de la langue anglaise est probablement à l'origine de nombreuses erreurs d'étiquetage par TreeTagger dans le corpus ScienText. En anglais, le nombre de flexions est réduit (cf. Analogies et dissemblances morphosyntaxiques des verbes de l'anglais et des langues romanes, p.59), à la différence des langues romanes, ce qui a pour conséquence des erreurs dans l'étiquetage syntaxique. Nous avons vu que dans les langues romanes, l'adjectif prend la marque du genre et du nombre, le nom a un genre, et que le verbe subit une flexion en fonction du mode, du temps et de la personne. Toutes ces marques entraînent l'existence de formes distinctes. En anglais, en revanche, l'adjectif est invariable, le nom n'a pas de genre, et un verbe régulier ne peut prendre que quatre formes (*look, looks, looked, looking*). Cette caractéristique de l'anglais rend également caduque toute tentative de vérification de l'étiquetage d'un adjectif en utilisant la fonctionnalité de recherche libre par forme sur ScienQuest : un adjectif anglais n'admet qu'une seule forme, contrairement aux deux<sup>135</sup> ou quatre formes qui se rencontrent en puissance pour un adjectif dans des corpus équivalents écrits dans l'une des langues romanes : (eng) *small*, (fra) *petit, petite, petits, petites* ; (spa) *pequeño, pequeña, pequeños, pequeñas*.

<sup>135</sup> Deux formes (le singulier et le pluriel) pour les adjectifs épiciens tels (ita) *utile* (singulier), *utili* (pluriel), (por) *forte* (singulier), *fortes* (pluriel).

La 'flexion zéro' de l'anglais facilite le transfert de catégorie grammaticale d'un mot, ce que Tesnière appelle *translation* (1959 : 364) et Crépin (1994 : 154) *conversion fonctionnelle*, un processus extrêmement fécond dans la néologie.

#### 4.3.3.2 *Erreurs d'étiquetage dues aux homographes en anglais*

Le très grand nombre d'homographes en anglais est à l'origine d'autres erreurs d'étiquetage. Ainsi, (eng) *patient* était étiqueté comme adjectif alors que ses occurrences réfèrent au nom. Le mot (eng) *certain* était étiqueté comme adjectif alors qu'il est en réalité déterminant. Le mot (eng) *fat* était étiqueté comme adjectif, dont l'équivalent français serait *gras*, alors qu'une vérification en contexte, au moyen d'une requête dans ScienQuest, montre qu'il s'agit majoritairement du nom *fat*, c'est-à-dire (fra) *graisse*.

Parfois, la même forme lemmatisée figurait dans les trois catégories grammaticales : (eng) *binding* apparaît trois fois, comme adjectif, nom et verbe. L'interprétation d'un adjectif qui serait *binding* a posé des difficultés : le mot a en effet la forme des adjectifs verbaux anglais du type *interesting*, *following*, *remaining*, etc. Cependant, une recherche avancée permet de trouver que les six occurrences de *is binding* et les cinq occurrences de *was binding* du corpus réfèrent toutes au verbe *bind* à la forme progressive et qu'il n'y a apparemment pas d'occurrences dans lesquelles *binding* est utilisé de manière attributive. Un examen complémentaire conduit à la possibilité d'interpréter *binding* comme un nom dans tous les cas que nous avons envisagés. Par ailleurs, la présence de *binding* parmi les verbes dénotait une erreur. Nous avons donc regroupé toutes les occurrences de *binding* dans la catégorie des noms.

Le lemme du nom est par convention la forme du nom au singulier. Parfois, il a semblé opportun de faire une requête de la forme du mot au pluriel afin de donner du sens à la présence de nombreuses occurrences du nom au pluriel : nous avons ainsi repéré l'importance de la locution (eng) *in terms of* que donne la requête de la forme *terms*. Les langues romanes utilisent une locution morphologiquement et sémantiquement équivalente : (spa) *en términos de*, (fra) *en termes de*, (ita) *in termini di*, (por) *em termos de*. Par ailleurs, l'anglais *term* se traduit par les formes au singulier (spa) *término*, (fra) *terme*, (ita) *termine* et (por) *termo*. En conséquence, nous avons traduit (eng) *term* par ces équivalents au singulier, mais deux entrées distinctes pourraient figurer dans notre liste : *term* et *in terms of*.

Lorsque nous savions qu'un mot étiqueté comme verbe admet une forme homonyme parmi les adjectifs ou les noms, nous avons procédé à une recherche par formes afin de vérifier si le verbe que nous jugions douteux dans la liste extraite avait été correctement étiqueté.

Prenons l'exemple de *copy* et de *head*, qui peuvent être verbes ou noms. Les verbes *copy* et *head* peuvent prendre les formes fléchies *copies*, *copied*, *heads* et *headed*. Une requête des formes *copied* et *headed* a permis d'écartier des résultats ce qui pourrait être des occurrences des noms *copy* et *head* au singulier et au pluriel.

Lorsque la confusion risquait de se rencontrer entre des occurrences d'adjectifs et de noms homonymes, nous avons cherché la forme au pluriel. 4 040 occurrences du mot (eng) *standard* sont étiquetées comme des noms et 11 432 occurrences comme des adjectifs. Une recherche de la forme *standards* permet de s'assurer que le nom *standard* devrait effectivement faire partie des mots fréquents du corpus, avec 1 464 occurrences de cette forme : le nombre minimal d'occurrences pour faire partie des mots fréquents étant 1 556 (voir ci-après Sélection des 2000 mots les plus fréquents, p.141), on peut supposer qu'au moins 92 occurrences du mot sont employées au singulier dans le corpus. La forme *standard* réfère donc à deux entrées distinctes, un nom et un adjectif.

Comme nous l'avons constaté, en cas d'étiquetage automatique, l'utilisation des corpus à des fins d'analyse lexicologique et/ou didactique requiert une vigilance particulière pour les mots potentiellement homographes. L'étude du profil combinatoire (cf. infra section Seconde tentative : déterminer le sens par le profil combinatoire, p.144) permet de lever le doute sur des mots comme *light*, qui peut être étiqueté à tort comme adjectif alors qu'il est nom. Un enseignant d'anglais peut aussi, admettons-le, essayer de faire feu de tout bois et considérer que les erreurs d'étiquetage des homographes anglais se prêtent à l'invention de diverses tâches de discrimination.

Il est un cas fréquent dans lequel on gagnera à améliorer les réglages des étiqueteurs syntaxiques, tant ce cas est fréquent et dépend de l'état de lexicalisation d'une séquence collocative : comment distinguer d'un côté l'utilisation attributive du substantif, par exemple dans *cell survival* et *cell isolation procedure*, de l'autre le mot composé, par exemple *cell culture*, *cell multiplication*, etc. ?

### 4.3.3 Homogénéisation de la graphie

Au sujet de la forme écrite des mots anglais, nous avons opté pour :

- la généralisation de la graphie américaine, dont les apprenants asiatiques sont probablement plus familiers : *color* au lieu de *colour*, *defense* et non *defence*, *fiber* et non *fibre*, etc.<sup>136</sup> ;
- l'ajout d'éléments figés indissociables qui constituent dans les faits des polylexicaux (Gross, 1996 : 155) : *amino* n'existe pas seul, *acid* lui a été ajouté (*amino acid*), *Saccharomyces* vient inmanquablement après *cerevisiae*, *in* devant *vitro* et *vivo*, etc. ;
- la rectification de l'orthographe de quelques mots : *inhibitor* pour *inhibiter*, *lineage* pour *linage*, etc.

Cependant, certaines erreurs n'auront pu être rectifiées en raison de l'ampleur de la tâche. Les parasyntétiques tels que *long-term*, par exemple, sont parfois analysés comme deux mots par l'étiqueteur syntaxique.

### 4.3.4 Sélection des 2000 mots les plus fréquents

Une fois les trois listes nettoyées, nous avons conservé les adjectifs, noms et verbes qui comptent parmi les 2000 plus fréquents du corpus scientifique anglais de ScienText. Pour cela, nous avons examiné leur nombre d'occurrences. Le nombre minimum d'occurrences s'élève à 1 556, autrement dit chacun des 2 000 mots retenus apparaît au moins 1 556 fois, et le nombre maximum d'occurrences est de 510 052. L'écart est considérable, néanmoins, nous avons opté pour ce vaste échantillon afin de pouvoir en tirer des enseignements quant à la catégorie grammaticale des mots, leur catégorie sémantique, leur similarité avec leur équivalent de traduction dans les langues romanes, certaines différences entre les plus fréquents et les moins fréquents, etc.

23,8 % des mots sont des adjectifs, 57,5 % des noms et 18,7 % des verbes ; nous traitons de leur représentativité dans la section 6.2.2.1 (Résultats globaux : les catégories grammaticales des 2000 mots anglais fréquents), p.203.

Leur liste brute figure dans l'Annexe 1 : les 2000 lemmes anglais fréquents du corpus ScienText, p.291. Ces mots seront par ailleurs consultables en ligne sur le portail hal parmi les documents de l'auteur de cette thèse<sup>137</sup>.

---

136 Une quinzaine de mots sont concernés.

137 <https://hal.archives-ouvertes.fr/>

Un mot monosémique dans le corpus apparaît en bleu clair. Un mot polysémique à une traduction apparaît en vert clair (cf. image 7). Enfin, un mot polysémique à plus d'une traduction apparaît en bleu foncé. Pour la distinction entre ces trois catégories, voir *Similarité : trois catégories de mots traduits*, p.195.

V	30367	induce	inducir	induire	indurre	induzir	To cause, bring about, lead to
V	29062	contain	contener	contenir	contenere	conter	To hold inside
V	28674	provide	fornecer / proveer	fournir / pourvoir	fornire / provvedere	fornecer / prover	To give what is needed or desired (...)
V	27691	increase	acrescentar	s'accroître	accrescersi	acrescentar-se	To become larger
V	27219	perform	cumplir / ejecutar	exécuter	compiere / eseguire	cumprir / executar	To do something; to execute
V	26322	express	1) expresar 2) expresar / exprimer		esprimere	1) expresar 2) expresar	1) express: (biochemistry) To translate messenger RNA i
V	25896	describe	describir	décrire	descrivere	descrever	To represent in words
V	25275	include	incluir	inclure	includere	inchuir	To contain, as parts of a whole; to comprehend
V	24697	determine	determinar	déterminer	determinare	determinar	To ascertain definitely; to figure out, find out, or conclude
V	24066	follow	seguir	suivre	seguire	seguir	1) follow : To go or come after in a sequence 2) follow :
V	23810	see	ver	voir	vedere	ver	1) see : To perceive or detect with the eyes, or as if by si
V	21996	give	dar	donner	dare	dar	To move, shift, provide something (...)
V	23330	obtain	obtener	obtenir	ottenere	obter	To get hold of; to gain possession of, to procure; to acqui
V	21978	observe	observar	observer	osservare	observar	To notice or view, especially carefully or with attention to
V	21597	indicate	indicar	indiquer	indicare	indicar	To point out; to discover; to direct to a knowledge of; to s

*Image 7 : Extrait de la feuille contenant les 374 verbes*

## **4.4 Détermination du sens des 2000 mots anglais fréquents**

Nous nous sommes demandé comment procéder étant donné le nombre d'occurrences des mots (entre 1556 fois 510 052 fois), comment traiter les mots polysémiques, et quelle(s) acception(s) d'un mot polysémique devai(en)t primer sur les autres.

Cette section s'organise en trois temps. D'abord, nous montrons comment a évolué notre méthodologie, d'un examen manuel à un traitement automatique des données afin de tendre vers plus de consistance. Puis, nous expliquons comment ont été définies les 2208 acceptions fréquentes en utilisant le profil combinatoire de chaque mot. Enfin, nous précisons comment nous avons trié les 2208 acceptions fréquentes que nous avons pu identifier en deux catégories sémantiques.

### **4.4.1 Évolutions de l'approche suivie pour cerner le sens**

#### **4.4.1.1 *Un premier examen manuel des 100 premières occurrences***

Initialement, nous avons l'intention d'identifier l'acception dans laquelle un lecteur de textes scientifiques s'attend qu'un mot soit employé. Nous aurions ensuite trouvé les équivalents sémantiques qui traduisent cette acception et aurions comparé la forme de ces mots. Pour trouver à quoi peut s'attendre un lecteur dans ce genre de textes, il nous fallait cerner le sens que prennent les 2000 mots identifiés dans les textes du corpus.

Notre première orientation méthodologique se justifiait par les contingences car l'examen manuel exhaustif des 14 millions d'occurrences serait impossible. Elle découlait surtout de notre ignorance de certains outils du TAL.

Nous avons examiné manuellement les 100 premières occurrences du mot grâce au concordancier de ScienQuest afin de trouver l'acception dominante en fonction de ses cooccurrents de droite et de gauche. Pour limiter l'examen à des proportions raisonnables, nous avons concentré notre attention sur le sens que prennent les 100 premières occurrences de chaque mot anglais fréquent. Concrètement, dans l'outil de recherche de ScienQuest, nous avons réglé nos préférences sur 100 occurrences, 10 cooccurrents de droite et 10 cooccurrents de gauche. Nous avons alors effectué des requêtes pour examiner les cooccurrents de droite et de gauche de ces 100 occurrences du mot afin de cerner le sens qu'il prend dans ces contextes.



#### **4.4.1.2 Insuffisances de ce procédé**

Cette démarche présentait des insuffisances dont nous nous sommes rapidement rendu compte. Ainsi, nous ne tenions pas compte de la distribution du mot dans le corpus. Or certains mots sont utilisés à de multiples reprises dans un même texte, si bien que les 100 occurrences examinées pouvaient ne référer qu'à très peu de textes. Le mot *length*, par exemple, apparaît 33 fois dans le texte n°71. Dans le cas extrême où un mot serait apparu 100 fois dans un même texte, nous risquions de prendre comme représentative une acception de ce mot qui ne se rencontre que dans un seul des textes du corpus.

En outre se posait la question de la polysémie : comment s'assurer qu'une acception fréquente dans un échantillon de 100 occurrences est une acception fréquente dans l'ensemble du corpus ? Et du reste, pourquoi s'en tenir à l'acception la plus fréquente ? Il nous fallait changer d'échelle et passer d'un tâtonnement intuitif à une prise en compte effective des dimensions du corpus. Pour ce faire, des outils existent, que nous avons découverts tout au long de cette recherche.

La méthodologie que nous avons adoptée a connu des évolutions à mesure que nous entrions en discussion avec nos directeurs de thèse, avec des doctorants, ainsi qu'avec d'autres chercheurs lors de colloques. Les séminaires du laboratoire LIDILEM auxquels nous avons eu part, mais aussi les échanges que nous avons eus sur les plateformes d'échanges plurilingues en ligne dans le cadre du programme de recherche MIRIADI (Mutualisation et Innovation dans un Réseau d'Intercompréhension A Distance) ont heureusement influencé notre approche. Surtout, les outils de TAL découverts pendant des formations ont rendu possible le renouvellement de notre façon d'appréhender à la fois le lexique plurilingue, l'enseignement du FLE et d'une langue en général.

Nous avons utilisé un outil statistique empruntée à la textométrie : le profil combinatoire.

#### **4.4.1.3 Seconde tentative : déterminer le sens par le profil combinatoire**

Le sens que prend un mot dans un contexte dépend des mots qui le précèdent, de ceux qui le suivent et des relations syntaxiques qui existent entre ces mots. L'expression *profil combinatoire* réfère à « la structure schématique du voisinage syntaxique et sémantique d'un mot telle qu'elle se manifeste dans un vaste corpus » (Blumenthal, 2002 : 116). Nous avons

utilisé le Lexicoscope<sup>138</sup> (Kraif & Diwersy, 2012 : 399-406), un outil en ligne qui permet d'établir le profil combinatoire d'un mot.

Le profil combinatoire d'un mot pivot montre quels sont ses cooccurrents. Les plus fréquents des cooccurrents peuvent être considérés comme des collocatifs. C'est pourquoi nous pouvons dire que nous avons déterminé les acceptions fréquentes des adjectifs, noms et verbes fréquents du corpus ScienText en nous appuyant sur les adjectifs, noms, et verbes qui sont leurs collocatifs dans ce même corpus. Dans le Lexicoscope, nous avons recherché ces collocatifs en fixant un seuil de significativité : sont considérés comme collocatifs les cooccurrents qui ont une fréquence de 50 et un seuil de spécificité de 10,83.

The screenshot shows the Lexicoscope interface with the 'Cooccurrences (lexicogrammes)' tab selected. The interface is divided into three main sections: 'Concordances', 'Cooccurrences (lexicogrammes)', and 'Expressions polylexicales'. The 'Cooccurrences (lexicogrammes)' section contains the following settings:

- Type de profil combinatoire: collocation
- Définition des Collocatifs : lemme  Forme   
 Catégories  Traits
- Fusionner les relations pour un même collocatif : Oui  Non
- Ancrer les relations sur le pivot (#1) seulement : Oui  Non
- Inclure les tokens non instanciés de la requête (lemme=.\*): Oui  Non
- Utilisation du cache des résultats : Oui  Non
- Seuil de fréquence : >= 50
- Seuil de dispersion : >= 1
- Seuil de spécificité : log-likelihood >= 10.83

Image 8 : Copie d'écran du Lexicoscope

Nous n'analyserons pas les collocatifs des mots fréquents de ScienText parce que nous les avons étudiés ici dans un but principalement définitoire. Néanmoins, nous pensons que le profil combinatoire présente un grand intérêt pour l'étude de la langue de spécialité de la santé et son enseignement, mais aussi pour la recherche de séquences collocatives analogues dans les langues romanes. Nous nous en tiendrons ci-après à quelques points saillants.

Une extraction des collocatifs a été faite à partir de la liste des 2000 mots fréquents. Ces collocatifs présentent des disparités remarquables. Ainsi, certains mots n'admettent qu'un ou deux collocatifs (adjectif, nom ou verbe), nous pouvons alors probablement parler de

138 Lexicoscope : <https://frama.link/qCzuF29W>

figements. Ainsi, *crucial* est associé dans ScienText à un seul collocatif, *role*. Le mot *consistent* a deux collocatifs, *pattern* et *result*. Il en est de même de *distal*, avec *end* et *radius*, de *format*, avec *file* et *use* (verbe), de *intensive*, avec *care* et *therapy*, de *necrosis* avec *factor* et *tumor*. D'autres mots fréquents, en revanche, en ont plusieurs dizaines, à l'instar du verbe *mediate*, qui en a 46.

The screenshot shows the Lexicoscope interface with the title "Lexicoscope - Lexicoscope : Résultat d'une requête -". Below the title, there is a section for "Lexicogrammes" with a sub-section "Lexicogrammes : Tableau de contingence et mesures d'association". The interface includes a search bar, a "Show 25 entries" dropdown, and a table of results. The table has columns for I1, I2, f.deprels, f, f1, f2, N, f.disp, am.log.likelihood, and r.log.likelihood. The table displays five entries, all starting with "intravenous\_A" in the I1 column and various medical terms in the I2 column. Below the table, there are input fields for I1, I2, f.deprels, f, f1, f2, N, f.disp, am.log.likelihood, and r.log.likelihood, and a "Télécharger : [.csv]" button.

I1	I2	f.deprels	f	f1	f2	N	f.disp	am.log.likelihood	r.log.likelihood
intravenous_A	infusion_N	~NN NOMPREP	134	5054	4802	47778862	1	1234,4767	1
intravenous_A	injection_N	~NN	113	5054	9645	47778862	1	843,7076	2
intravenous_A	administration_N	~NN OBJ	113	5054	12037	47778862	1	793,8682	3
intravenous_A	immunoglobulin_N	~NN	53	5054	2090	47778862	1	477,1700	4
intravenous_A	drug_N	~NN	66	5054	29058	47778862	1	279,8840	5

Image 9 : Copie d'écran d'un extrait de lexicogramme dans le Lexicoscope

Le profil combinatoire a conduit à des surprises : *member* n'est pas fréquemment utilisé dans l'acception '*Part of an animal capable of performing a distinct office*', comme on pourrait s'y attendre dans des articles liés à la santé. La seule acception qui se rencontre plus de 50 fois avec des collocatifs est '*One who officially belongs to a group*'. Aussi en déduit-on que dans ce corpus de spécialité, l'acception relevant de la sociologie est plus utilisée que l'acception spécifique à la langue de spécialité.

Au bout du compte, nous avons renoncé à sélectionner pour un mot polysémique l'acception qui serait dominante dans le corpus car pour la déterminer avec rigueur, il faudrait additionner les occurrences de chacun des collocatifs après avoir rattaché chaque couple lemme – cooccurrent à une acception, une tâche d'une ampleur excessive.

Nous avons retenu toutes les acceptions fréquentes que suggèrent les collocatifs qui apparaissent 50 fois et ont un seuil de spécificité minimum de 10,83.

La somme des acceptions fréquentes des 2000 mots anglais fréquents du corpus scientifique anglais de ScienText que nous avons identifiées par cette méthode s'élève à 2208 ; c'est notre base de travail pour élaborer l'interlexique.

#### **4.4.1.4 Détermination de la monosémie / polysémie d'un mot du corpus ScienText**

Dans le cadre de ce travail, un mot anglais est considéré comme monosémique s'il admet une seule acception fréquente dans le corpus ScienText. Un mot est polysémique s'il admet au moins deux acceptions fréquentes dans le corpus ScienText.

Nous nous sommes appuyé sur deux critères : les collocatifs dans le corpus et les acceptions données dans le *Wiktionary*.

Ainsi, comme le profil combinatoire du mot *solution* fait apparaître qu'il a pour collocatifs : *add, aqueous, balanced, block, buffer (N), buffer (V), contain, free, good, hybridization, ml, obtain, possible, prepare, provide, saline, salt, stock, structure, substrate* ;

nous en déduisons qu'il est utilisé dans deux acceptions du *Wiktionary*.

Premièrement, presque tous les collocatifs incitent à sélectionner l'acception valable en (bio)chimie :

*Solution* : 'A homogeneous mixture, which may be liquid, gas or solid, formed by dissolving one or more substances'.<sup>139</sup>

Deuxièmement, les collocatifs *good, possible, provide* suggèrent :

*Solution* : 'An act, plan or other means, used or proposed, to solve a problem'.

---

<sup>139</sup> Les définitions en anglais proviennent du *Wiktionary*.

## 4.4.2 Comment définir les 2208 acceptions identifiées

### 4.4.2.1 Le choix du Wiktionary

Au lieu d'utiliser prioritairement un dictionnaire de spécialité, nous avons fait le choix de puiser les définitions des acceptions dans le *Wiktionary*<sup>140</sup>, en dépit des inévitables imprécisions de certains articles.

En effet, non limité à un nombre de pages pour une version papier, un dictionnaire collaboratif généraliste en ligne comme le *Wiktionary* peut définir virtuellement tous les mots, et toutes les acceptions d'un mot, qu'elles relèvent de la langue générale ou de la langue de spécialité. En outre, nous considérons comme un réel progrès qu'au moyen d'un dictionnaire continuellement mis à jour, le grand public accède en ligne à la signification de mots qui appartiennent à un domaine aussi essentiel pour l'être humain que la santé.

Certaines définitions du Wiktionary, trop longues, ne figurent pas *in extenso* dans les données. À chaque fois, nous avons copié la partie de la définition qui devrait permettre de saisir l'acception dans laquelle on rencontre le mot dans le corpus ScienText, sans consigner les indications de nature grammaticale telles que *countable*, *transitive*, etc. De la définition, nous n'avons parfois gardé que le début. Le cas échéant, nous avons indiqué par des points de suspension entre parenthèses (...) que la définition a été tronquée par nos soins, l'objectif étant de faire reconnaître l'acception à laquelle réfère la glose.

Les définitions relevées pendant l'élaboration des traductions, entre octobre 2014 et août 2015, sont une sorte d'instantané en ce qu'elles reflètent l'état du *Wiktionary* pendant cette période-là. Elles auront déjà été amendées et actualisées et le seront à l'avenir, à mesure que des internautes les amélioreront et que l'usage ajoutera de nouvelles significations à ces mots, ceci tant que le *Wiktionary* rencontrera du succès.

Nous avons été confronté à la question de la qualité parfois insuffisante des définitions du *Wiktionary*. Ainsi, la définition de *combination* énoncée par le *Wiktionary* comme '*The act of combining, the state of being combined or the result of combining*' n'explicite en rien ce mot. Cependant, nous avons laissé telles quelles ces définitions qui admettent un mot similaire dans les langues romanes, avec un signifiant très large.

Il en est tout à fait autrement des acceptions relevant de la langue de spécialité de la santé puisque le dispositif didactique que nous visons repose en partie sur ce lexique spécifique, et

---

140 [https://en.wiktionary.org/wiki/Wiktionary:Main\\_Page](https://en.wiktionary.org/wiki/Wiktionary:Main_Page)

que nous nous devons de trouver des équivalents de traduction précis pour ces acceptions valables dans le domaine de la santé. Nous avons donc contribué au *Wiktionary* en ajoutant des définitions pour des acceptions qui faisaient défaut (cf ci-après Ajout de définitions dans le Wiktionary).

#### 4.4.2.2 *Acceptions présentes / absentes du Wiktionary*

Que nous connaissions la signification d'un mot ou non, nous avons systématiquement fait des va-et-vient entre les occurrences de ce mot dans le corpus ScienText, son profil combinatoire dans le Lexicoscope, et les définitions qu'en propose le *Wiktionary*.

Dans la tentative de cerner le sens des occurrences d'un mot du corpus, nous pouvons distinguer selon que :

- A. Le mot n'est pas défini dans le *Wiktionary*.
- B. Le mot est défini dans le *Wiktionary*.

Pour B, nous distinguons deux cas, à leur tour subdivisibles :

- B.1. Le *Wiktionary* définit une seule acception pour le mot.
  - B.1.1. L'acception définie dans le *Wiktionary* est celle rencontrée dans le corpus.
  - B.1.2. L'acception définie dans le *Wiktionary* n'est pas celle rencontrée dans le corpus.
- B.2. Le *Wiktionary* définit plusieurs acceptions pour le mot.
  - B.2.1. L'une des acceptions définies dans le *Wiktionary* est celle rencontrée dans le corpus.
  - B.2.2. Aucune des acceptions définies dans le *Wiktionary* n'est celle rencontrée dans le corpus<sup>141</sup>.

Dans trois cas, A., B.1.2. et B.2.2., il manquait une définition.

---

<sup>141</sup> S'ajoute ici le cas des mots polysémiques du corpus dont l'une des acceptions ne figure pas parmi les acceptions définies dans le *Wiktionary*.

Nous n'avons pas trouvé de cas A, ce qui signifie que les 2000 mots les plus fréquents font tous l'objet d'une entrée dans le *Wiktionary*, autrement dit qu'au moins l'une des acceptions qu'ils admettent y est définie, mais pas nécessairement l'acception qu'on rencontre dans le corpus ScienText.

Cette constatation nous conduit à ranger le *Wiktionary* parmi les outils qui peuvent compter pour le grand public anglophone, particulièrement si l'on considère que certains mots très spécialisés n'étaient pas définis dans les dictionnaires unilingues généraux en langues romanes au moment où nous avons mené notre recherche, ni même traduits dans certains des dictionnaires bilingues anglais-français consultés (cf. section Traduction des 2000 mots anglais fréquents en espagnol, français, italien et portugais, p.162), alors qu'ils l'étaient dans le *Wiktionary*.

Les mots espagnol et italien équivalents à (eng) *haplotype*, par exemple, ne figurent pas parmi les mots-vedettes du *Diccionario de la Lengua Española*<sup>142</sup>, ni dans le *Vocabolario de la Treccani*<sup>143</sup>. Les noms (fra) *orthologue*, (ita) *ortologo* et (por) *ortólogo*, équivalents du nom (eng) *ortholog*, ne sont définis ni dans le TLFi<sup>144</sup>, ni dans le *Vocabolario*<sup>145</sup>, ni dans le *Dicio*<sup>146</sup>. Nous avons même rencontré des mots anglais dont aucun équivalent dans les quatre langues romanes n'est défini dans les dictionnaires généralistes unilingues sus-nommés : l'adjectif (eng) *antisense* est défini dans le *Wiktionary*, mais ses équivalents dans les quatre langues romanes n'ont pas fait l'objet d'une entrée dans les dictionnaires consultés.

Nous avons traité de la même façon les cas B.1.2. et B.2.2., à savoir l'absence de définition pour une acception présente dans le corpus ScienText. La section qui suit détaille ce point.

#### **4.4.2.3 Ajout de définitions dans le Wiktionary**

Dans le cas des acceptions non définies dans le *Wiktionary*, nous avons formulé une définition et l'avons ajoutée à ce site.

119 mots sont concernés pour lesquels 124 définitions ont été (re)formulées. Il n'est probablement pas surprenant que 96 acceptions relèvent de la santé : en effet, un spécialiste

---

142 <http://www.rae.es/recursos/diccionarios/drae>

143 <http://www.treccani.it/vocabolario/>

144 <http://atilf.atilf.fr/dendien/scripts/tlfiv4/showps.exe?p=combi.htm;java=no> ;

145 Ibidem

146 <http://www.dicio.com.br/>

du domaine préférera peut-être contribuer à des ressources spécialisées telles que le Portail Terminologique de Santé de l'Université de Rouen<sup>147</sup>. Nous n'analyserons pas plus avant les mots qui ont nécessité d'être définis ; nous renvoyons à l'Annexe 3 : Définitions à valider des experts, p.388, qui recense ces définitions, qui devront être validées par des experts.

Pour ajouter ces acceptions au *Wiktionary*, nous nous sommes appuyé la plupart du temps sur une ou plusieurs pages de *Wikipédia*.

1) Quand la définition d'un mot manquait dans le *Wiktionary* et qu'il était défini dans une page de *Wikipédia*, nous en avons recopié la formulation, parfois en l'adaptant.

2) Quand la définition d'un mot manquait dans le *Wiktionary* et qu'aucune page de *Wikipédia* ne lui était consacrée, nous avons consulté d'autres pages de cette ressource où le mot apparaît et avons formulé une définition. Comme le *Wiktionary* est un dictionnaire collaboratif, les éventuelles erreurs que nous y aurons introduites seront rectifiées par des internautes experts du domaine puisque la révision est au cœur du processus éditorial.

En cherchant des définitions, nous nous sommes heurté au fait que *Wikipédia* est le fruit de contributeurs de tous horizons, si bien que la qualité des articles présente des disparités évidentes. Nous avons été attentif aux articles signalés par certains internautes comme requérant une réécriture en raison de leur qualité jugée insuffisante.

Comme le menu de gauche de *Wikipédia* renvoie à des articles qui traitent la même entrée (ou parfois une entrée au contenu proche) dans les autres langues du projet, nous avons pu comparer le traitement d'une même entrée en d'autres langues que l'anglais afin de confirmer ou rectifier la définition d'un mot à ajouter au *Wiktionary*. Il a été parfois nécessaire de recourir à un article de *Wikipédia* dans d'autres langues que celles que nous traitons ici<sup>148</sup> si sa formulation permettait de cerner plus aisément la signification d'un mot. Nous avons adapté ces définitions en anglais.

---

147 <http://www.hetop.eu/hetop/>

148 Notamment en allemand pour certains termes appartenant à la génétique.



## **4.5 Deux catégories sémantiques pour les 2208 acceptions fréquentes**

### **4.5.1 Un tri manuel en deux catégories : lexique de spécialité et LST**

Trier les mots anglais selon leur sens permet une approche contrastive du lexique anglais fréquent du discours de spécialité et du lexique anglais fréquent transdisciplinaire. En outre, dans une perspective didactique, nous pensons que le lexique spécialisé peut fournir une passerelle de transfert vers la langue utilisée dans le discours écrit de la santé. Nous avons donc trié les acceptions selon deux catégories sémantiques : le lexique de spécialité de la santé et le LST.

Pecman note que se pose la question de la définition des contours de l'échantillon du lexique de la langue scientifique générale. Comme elle, nous nous sommes demandé sur quels critères classer un mot dans une catégorie qui dépasserait les frontières (imaginaires) entre les disciplines scientifiques.

Pour établir un lexique, Pecman a comparé les mots présents dans des textes scientifiques se rattachant à plusieurs domaines scientifiques. Elle a rassemblé un nombre équilibré de textes relevant de chacun de ces domaines. À la différence de ce qu'a fait cette chercheuse, nous nous sommes fondé sur un corpus préélaboré, regroupant des textes appartenant à un même grand domaine, la santé. Nous n'avons pas eu de prise sur le nombre de textes relevant des différents sous-domaines représentés (génomique, biochimie, médecine, etc.), et la question de la représentativité de chacun peut se poser.

Le tri par catégories sémantiques s'appuie par conséquent sur l'ensemble des 35,2 M de mots des textes du corpus scientifique anglais de ScienText, dont nous avons délimité les 2208 acceptions fréquentes sur la base de leur profil combinatoire. Il a été effectué manuellement, en déterminant en premier lieu le lexique de spécialité, en second lieu le LST.

Sont rangés dans la catégorie du lexique de spécialité les acceptions liées à l'anatomie (générale, comparée), à la pathologie humaine (anatomopathologie générale, biologie médicale) et à la thérapeutique (traitement, génomique, biologie interventionnelle). Les mots les plus faciles à trier ont été les plus spécialisés, ceux qui ne présentent aucun doute sur la spécificité.

Les unités lexicales qui figurent dans cette catégorie sémantique peuvent appartenir à la langue *générale*, si l'acception dans laquelle on les rencontre dans le corpus témoigne d'une spécialisation sémantique dans le sens où l'entend Zwanenburg (1983 : 34) : « la spécialisation sémantique elle-même présuppose que l'élément spécialisé figure dans le lexique ». Ainsi, *air* est rangé parmi le lexique de spécialité parce qu'il se rencontre dans des contextes où il est question de la saturation du sang en oxygène.

Les mots du LST regroupent tous les autres mots, ceux qu'on pourrait trouver dans d'autres discours scientifique : des mots liés à la connaissance, aux dynamiques sociales, aux orientations de la recherche, aux caractéristiques, à la démarche, à la mesure et à l'interprétation. Nous avons considéré que tous les mots qui ne font pas partie du lexique de spécialité font par défaut partie du LST.

Nous n'avons pas procédé à des calculs de spécificité en comparant nos données avec d'autres corpus, c'est pourquoi ranger certains mots dans l'une ou l'autre de ces deux catégories a posé des problèmes récurrents, tant la spécialisation de mots courants comme *dog* ou *left* ne va pas de soi. Si nous avons classé ces deux mots parmi le lexique de spécialité, c'est que dans le corpus anglais de ScienText, le mot *dog* réfère à l'animal exploité lors d'expérimentations sur les animaux, et *left* est utilisé pour caractériser un organe ou une partie d'organe à gauche du plan sagittal médian<sup>149</sup>. La classification de certaines acceptions pourra à l'avenir être comparée à des listes de LST.

## 4.5.2 Sous-catégories du lexique de spécialité de la santé

Les sous-catégories que nous traitons ci-après correspondent à des sous-domaines de spécialité de la santé : l'anatomie générale, l'anatomie comparée, la pathologie humaine sous ses deux principales subdivisions, et la thérapeutique.

### 4.5.2.1 Anatomie générale

L'anatomie générale est définie par l'Académie de Médecine comme la « branche de l'anatomie qui étudie les structures fondamentales des organes et l'appartenance de ceux-ci à des systèmes de même structure et à des appareils constitués en vue d'une fonction »<sup>150</sup>.

---

149 Voir aussi <http://dictionnaire.academie-medecine.fr/?q=gauche>

150 <http://dictionnaire.academie-medecine.fr/?q=anatomie>

En plus de termes généraux (*body, organ, etc.*), il s'agit là des mots qui expriment :

- les positions anatomiques (*side, lateral, etc.*), les plans de coupe (*longitudinal, etc.*), les directions (*left, etc.*), la situation par rapport à la surface du corps (*internal, etc.*) ;
- le système ostéo-articulaire (*bone, cartilage, etc.*) ;
- le système cardio-vasculaire (*cardiac, ventricular, etc.*) ;
- le système respiratoire (*ventilation, respiratory, etc.*) ;
- le système nerveux (*brain, cerebral, etc.*) ;
- le système digestif (*nutrition, food, etc.*) ;
- le système génito-urinaire (*kidney, renal, etc.*) ;
- le système endocrinien (*secretion, gland, etc.*) ;
- le système immunitaire (*immunity, affinity, etc.*).

#### **4.5.2.2 Anatomie comparée**

Selon le Dictionnaire médical de l'Académie de Médecine, «l'anatomie comparée étudie les rapports de concordance et de divergence existant entre les structures homologues des êtres vivants pour en tirer des arguments de classification et d'évolution des espèces »<sup>151</sup>.

Les termes concernés peuvent se rattacher à :

- l'ontogenèse (*development, differentiation, etc.*) ;
- la phylogénèse (*evolution, compete, etc.*).

#### **4.5.2.3 Pathologie humaine : anatomopathologie générale**

La pathologie humaine comprend l'anatomopathologie, définie par le MeSH comme l' «étude des causes des maladies comme expressions des changements structuraux des cellules ou des tissus, provoqués par la maladie »<sup>152</sup>, et la biologie médicale.

---

151 <http://dictionnaire.academie-medecine.fr/?q=anatomie>

152 <http://www.chu-rouen.fr/page/anatomopathologie>

Nous avons classé dans l'anatomopathologie générale des termes qui présentent une portée générale, tels que *condition*, *injury*, etc., mais aussi ceux qui entretiennent un lien sémantique avec :

- les inflammations (*sepsis*, *osteoarthritis*, etc.) ;
- la cancérologie (*cancer*, *tumor*, etc.) ;
- les troubles vasculaires (*hypertension*, *infarction*, etc.) ;
- les altérations cellulaires, la réparation et la nécrose (*lesion*, *toxicity*, etc.) ;
- les troubles psychiatriques (*dependence*, *depression* etc.) ;
- l'épidémiologie (*pathogen*, *parasite*, etc.).

#### **4.5.2.4 Pathologie humaine : biologie médicale**

Dans la sous-catégorie de la biologie médicale, nous avons classé ce qui concerne :

- la chimie (description de la matière : *substance*, *fluid*, etc.) ;
- les réactions (bio)chimiques (*oxidation*, *product* etc.) ;
- la biochimie clinique (*biopsy*, *in vitro*, etc.).

#### **4.5.2.5 Thérapeutique**

La thérapeutique est définie comme suit par l'Académie de Médecine : « discipline médicale qui se consacre aux moyens pouvant être utilisés pour supprimer une maladie ou corriger les lésions produites par un traumatisme et prévenir leurs séquelles »<sup>153</sup>. Nous incluons ici les disciplines de la génomique et de la biologie interventionnelle, dont la raison d'être est de contribuer au traitement des patients.

Se rattachent à la thérapeutique les termes qui concernent :

- le traitement (*practitioner*, *diagnosis*, etc.) ;
- la génomique (*phenotype*, *chromosome*, etc.) ;
- la biologie interventionnelle, avec la thérapie cellulaire et la thérapie génétique (*clone*, *tag*, etc.).

---

153 <http://dictionnaire.academie-medecine.fr/?q=th%C3%A9rapeutique>

### **4.5.3 Sous-catégories du LST**

Nous avons dit dans la section 4.5.1 que tous les mots qui ne font pas partie du lexique de spécialité ont été classés parmi le LST. Nous avons pu distinguer les sept sous-catégories suivantes, communes au discours scientifique : la connaissance, les dynamiques sociales, les orientations de la recherche, la description des caractéristiques, la démarche scientifique, la mesure et l'interprétation.

#### **4.5.3.1 *La connaissance***

C'est l'objectif premier de la science. Parmi les mots de la connaissance, nous avons rangé les mots de :

- la cognition (« idea », « concept », etc.) ;
- la diffusion (« publish », « map », etc.).

#### **4.5.3.2 *Les dynamiques sociales***

Le déploiement de la science dépend de dynamiques sociales qui la dépassent mais qui conditionnent son action. Ces mots sont par exemple :

- des personnes et des entités (« member », « government », etc.) ;
- des lieux (« home », « office », etc.) ;
- des besoins (« provide », « suitable », etc.).

#### **4.5.3.3 *Les orientations de la recherche***

Les mots référant aux orientations de la recherche dénotent une inspiration théorique et la direction suivie, c'est pourquoi nous avons trié dans cette sous-catégorie :

- les objectifs (« explore », « purpose », etc.) ;
- les approches (« model », « preference », etc.).

#### **4.5.3.4 La description des caractéristiques**

Il s'agit d'un vocabulaire plutôt concret qui réfère aux caractéristiques :

- la composition (« constitute », « segment », etc.) ;
- l'apparence (« blue », « aspect », etc.) ;
- l'espace (« point », « central », etc.).

#### **4.5.3.5 La démarche scientifique**

Les termes de la démarche scientifique rassemblent :

- les outils (« testing », « automated », etc.) ;
- le déroulement de la recherche (« phase », « avoid », etc.).

#### **4.5.3.6 La mesure**

La sous-catégorie de la mesure est assez riche. Elle comprend :

- le temps (« daily », « second », etc.) ;
- les outils mathématiques (« proportion », « measurement », etc.).

#### **4.5.3.7 L'interprétation**

L'interprétation comprend les termes de :

- la comparaison et la vérification (« standard », « correct », etc.) ;
- la discussion (« cause », « responsible », etc.) ;
- la validation (« evident », « confirm », etc.).

## 4.6 Conclusion du chapitre 4

Dans ce chapitre, nous avons expliqué comment nous avons élaboré un lexique des adjectifs, noms et verbes fréquents dans le corpus scientifique anglais de ScienText. Au moyen d'outils du TAL, nous avons constitué une liste des 2000 lemmes fréquents et délimité le sens de 2208 acceptions fréquentes au moyen du Lexicoscope, qui établit le profil combinatoire des mots.

Nous n'avons pas envisagé les acceptions moins fréquentes de ce corpus. Pourtant, un mot peut avoir plusieurs collocatifs qui indiquent qu'il est employé dans une acception dont nous n'avons pas tenu compte simplement parce que les occurrences de ses collocatifs n'atteignent pas le seuil de fréquence fixé. Or l'addition des occurrences de ces collocatifs pourrait révéler, avec un calcul différent, qu'elles ont aussi leur place parmi les acceptions fréquentes. Il faudrait alors trier tous les collocatifs en fonction de l'acception qu'ils suggèrent, additionner les occurrences de l'une et l'autre acception, etc. Un tel travail conviendrait davantage à une équipe qui construirait une ontologie.

Le sens de référence associé à chaque lemme est donc le résultat de choix quantitatifs, et de notre compréhension du rapport de sens entre le lemme et ses collocatifs. L'outil EmoProf (Cavalla *et al.*, 2013) permet à un apprenant de mieux saisir ce rapport de sens. Un enseignant peut l'utiliser pour construire des séquences didactiques destinées à affiner la compréhension du lexique des émotions d'une langue au moyen de la base de donnée EmoBase. L'apprenant peut croiser les éléments lexicaux avec leurs collocatifs et élaborer des cartes mentales multicritères. Nous pensons que ce type d'outil pourrait se prêter à l'apprentissage du français L3 fondé sur les connaissances de l'apprenant en anglais L2.

Le lexique fréquent dans le discours scientifique de la santé est probablement le lexique qu'il est judicieux de connaître pour comprendre des textes de ce type de discours. Il n'est toutefois pas nécessaire que chacun des 2000 mots anglais fréquents obtenus dans notre étude soit connu d'un apprenant. En effet, l'opportunité de rencontrer ces mots dépend étroitement de son domaine de spécialité : biochimie clinique, médecine, génétique, etc. Nous ne ferons pas preuve d'originalité en estimant que ce sont probablement les mots les plus fréquents qu'il est pertinent de faire apprendre en priorité. S'il s'appuie sur les listes que nous avons élaborées, un enseignement de l'anglais pourra prendre en compte le nombre d'occurrences de chaque

mot fréquent dans le corpus. Le tri sémantique par sous-catégories renseigne plus précisément sur les termes fréquents du sous-domaine de spécialité des étudiants.



# Chapitre 5 - Traduction des 2000 mots anglais fréquents du corpus ScienText en espagnol, français, italien et portugais et mesures de similarité

## Sommaire du chapitre 5

5.1 Introduction du chapitre 5.....	160
5.2 Traduction des 2000 mots anglais fréquents en espagnol, français, italien et portugais.....	162
5.2.1 Une approche prioritairement cibliste.....	162
5.2.2 Trois cas rencontrés lors de la traduction.....	162
5.2.3 Exemple du processus de traduction : le nom <i>assay</i> .....	170
5.2.4 Conséquences des différences lexicosémantiques interlinguistiques.....	173
5.2.5 Conséquences des divergences lexicosyntaxiques.....	176
5.2.6 Convergence sémantique interlexicale.....	179
5.2.7 Conséquences de la variation intralinguistique.....	182
5.2.8 Forme et ordre des mots traduits dans les tableaux de données.....	186
5.3 Mesures de similarité : vers un interlexique et quatre lexiques bilingues.....	187
5.3.1 Un interlexique au niveau des acceptions et non des mots.....	187
5.3.2 Neutralisation des diacritiques des langues romanes.....	188
5.3.3 Mesure de similarité par le script CompareCols.csv.pl.....	189
5.3.4 Étapes suivies pour mesurer la similarité formelle.....	191
5.3.5 Mesure de similarité des mots de moins de 4 lettres.....	193
5.3.6 Similarité : trois catégories de mots traduits.....	195
5.3.7 Étapes pour obtenir quatre lexiques bilingues anglais-langue romane des mots analogues.....	196
5.4 Conclusion du chapitre 5.....	198

## 5.1 Introduction du chapitre 5

Dans le chapitre précédent, nous avons présenté la méthodologie suivie en vue d'extraire les 2000 mots anglais les plus fréquents du corpus anglais de la santé de ScienText, de définir leurs acceptions fréquentes et de trier ces acceptions en fonction de leur sens. Le chapitre 4 était donc focalisé sur le lexique anglais fréquent de la santé.

Dans le chapitre 5, nous expliquons comment ces 2000 mots anglais fréquents ont été traduits en espagnol, français, italien et portugais. Puis, au moyen d'un programme, nous avons mesuré la similarité formelle des mots de ces langues romanes avec les mots anglais qu'ils traduisent. Ces mots qui ont le même sens et une forme similaire sont des mots analogues. Les séries de quatre paires analogues (mots anglais – traduction dans une langue romane) constituent l'interlexique.

La recherche d'équivalents de traduction en espagnol, français, italien et portugais pour les 2208 acceptations identifiées est traitée dans la section 5.2, p.162.

La section 5.3, p.187, est consacrée à la mesure de similarité des mots fréquents anglais et de leurs équivalents de traduction dans les quatre langues romanes.

## **5.2 Traduction des 2000 mots anglais fréquents en espagnol, français, italien et portugais**

### **5.2.1 Une approche prioritairement cibliste**

C'est le travail de traduction qui a été le plus long dans cette recherche. Nous avons suivi une approche plutôt cibliste. L'admiral (2014 : 4) précise la distinction qu'il opère entre les traducteurs sourciers et les traducteurs ciblistes :

« J'appelle « sourciers » ceux qui, en traduction (et particulièrement, en théorie de la traduction), s'attachent au signifiant de la langue du texte-source qu'il s'agit de traduire ; alors que les « ciblistes » entendent respecter le signifié (ou, plus exactement, le sens et la « valeur ») d'une parole qui doit advenir dans la langue-cible. »

C'est en effet le jeu des signifiés présents dans les définitions des acceptions valables dans le corpus ScienText qui a influencé le choix de l'équivalent de traduction d'un mot. Il aurait pourtant été séduisant d'opter parmi les traductions possibles pour le plus semblable à l'anglais, nous aurions alors adopté une approche sourcière. Mais un tel choix conduirait à faire apparaître les langues de notre étude comme plus proches qu'elles ne le sont.

Ou bien nous pensions connaître l'équivalent de traduction d'un mot, ou bien nous ne connaissions pas sa traduction mais la morphologie du mot anglais évoquait une origine grecque ou latine, ou bien encore nous ignorions quelle pouvait être sa traduction et sa morphologie n'évoquait pas d'origine grecque ou latine. Dans le deuxième cas, nous avons appliqué une stratégie utilisée par l'intercompréhension (cf. 3.3.3, p.114).

### **5.2.2 Trois cas rencontrés lors de la traduction**

#### ***5.2.2.1 Équivalents de traduction connus***

Si nous pensions connaître l'équivalent du mot anglais dans une ou plusieurs des langues romanes, nous avons vérifié la pertinence du mot qui nous venait à l'esprit en consultant les

dictionnaire unilingues suivants : le *Diccionario de la Lengua Española*<sup>154</sup>, le *Trésor de la Langue Française informatisé*<sup>155</sup>, le *Vocabolario*<sup>156</sup> de la Treccani et le *Dicio*<sup>157</sup>.

Prenons le mot fréquent (eng) *ask*, qui appartient à la langue courante. Nous pensions en connaître la traduction dans les quatre langues romanes. Le profil combinatoire du verbe *ask* fait apparaître les collocations suivants : *participant, patient, question, respondent, subject, woman*, c'est-à-dire que 5 collocations sur 6 désignent des personnes. Nous déduisons des collocations *participant, question, respondent* et *subject* que ce verbe prend manifestement de sens de poser une question, comme le fait un enquêteur dans un questionnaire. Le rôle du soignant qui fait des choix pour porter assistance au patient peut aussi suggérer qu'il s'agit d'une invitation d'agir dans une certaine direction imprimée par le soignant, et une requête avancée pourrait nous permettre d'identifier les structures valenciennes (*ask someone to V*, etc.) pour le confirmer ou l'infirmier. Le comportement syntaxique des pivots n'est cependant pas pris en compte afin de ne pas alourdir les recherches. Il y a là sujet à d'autres recherches à l'avenir en syntaxe / sémantique comparatives.

Sur les déductions précédentes, nous avons retenu la première définition du *Wiktionary* :

*Ask : To request (information, or an answer to a question)*

Nous avons alors vérifié dans les dictionnaires unilingues si cette acception est exprimée dans l'une des définitions de l'équivalent de traduction qui nous venait à l'esprit. Nous hésitions entre (spa) *preguntar / pedir*, (ita) *chiedere / perguntare*, (por) *pedir / preguntar*.

Il résulte de la combinaison de certains signifiants utilisés dans la définition des dictionnaires de langues romanes que l'un des deux verbes implique plus nettement une question.

Ainsi, parmi les deux acceptions de (spa) *pedir* dans le *Diccionario de la Lengua Española*, la définition suivante<sup>158</sup> :

*Pedir : Expresar a alguien la necesidad o el deseo de algo para que lo satisfaga*

exprime plutôt une requête, une demande d'agir. En cela, *pedir* exprime une acception différente, c'est pourquoi nous l'écartons.

Le verbe *preguntar*<sup>159</sup>, dans l'acception suivante :

154 <http://www.rae.es/recursos/diccionarios/drae>

155 <http://atilf.atilf.fr/dendien/scripts/tlfiv4/showps.exe?p=combi.htm;java=no> ;

156 <http://www.treccani.it/vocabolario/chiedere/>

157 <http://www.dicio.com.br>

158 <http://dle.rae.es/?id=SJF1xVQ>

159 <http://dle.rae.es/?id=Tzhorqm>

*Preguntar : Interrogar o hacer preguntas a alguien para que diga y responda lo que sabe sobre un asunto*

exprime les sèmes véhiculés par les collocatifs de *ask*, qui impliquent une situation où l'on cherche à obtenir quelque chose par le langage plutôt qu'à inciter à l'action. Nous conservons donc *preguntar* comme équivalent de traduction de *ask* pour l'acception identifiée dans le corpus ScienText.

En français, un seul verbe, *demander*, nous vient à l'esprit. Dans le TLFi<sup>160</sup>, sa première signification est :

*Demander : Faire savoir que l'on souhaite connaître quelque chose*

Cette définition est plus large que la définition de *preguntar*, mais *demander* présente l'intérêt d'admettre un objet humain ou non humain, comme *ask*, bien qu'on ne puisse pas traduire *ask a question* par *\*demander une question*.

Quid de *chiedere / domandare* en italien ? Dans le Vocabolario de la Treccani<sup>161</sup>, voici la définition de la deuxième acception de *chiedere* :

*Chiedere : Esprimere il desiderio di avere o di ottenere qualche cosa, sia cose concrete sia cose astratte*

Cette définition coïncide avec celle de (spa) *pedir*, dont nous avons dit qu'il convient moins que (spa) *preguntar*. Qu'en est-il alors de (ita) *domandare* ? L'une des définitions renvoie à *chiedere*, tandis que l'autre<sup>162</sup> :

*Domandare : Rivolgersi con parole a qualcuno per ottenere risposta*

implique une requête à quelqu'un puisque (ita) *parole* signifie (fra) *mot* ou *parole*. Nous savons qu'en italien, on dit plutôt *fare una domanda* pour *ask a question*<sup>163</sup>. Étant donné que le verbe de la langue courante *domandare* s'utilise quand quelqu'un demande quelque chose à un tiers, nous avons cependant avec ce mot un équivalent de traduction approprié.

Enfin, deux verbes portugais sont associés : (por) *pedir* et *perguntar*. Dans le Dicio<sup>164</sup>, la deuxième acception de *perguntar* est définie comme suit :

*Perguntar : Indagar ; questionar através de perguntas, questões*

---

160 TLFi

161 <http://www.treccani.it/vocabolario/chiedere/>

162 <http://www.treccani.it/vocabolario/domandare/>

163 C'est là une limite à l'exercice : il n'est pas rare que l'usage aille à l'encontre de la traduction mot à mot.

164 <https://www.dicio.com.br/perguntar/>

Le mot *questões*, traduction d'un des collocatifs de *ask*, suggère que nous sommes sur la bonne voie.

(por) *pedir*<sup>165</sup>, est défini ainsi :

*Pedir : Solicitar, implorar ; elaborar pedidos*

où *implorar* implique une certaine humilité non contenue dans le sémème de (eng) *ask*. En conséquence, (por) *perguntar* semble le verbe approprié.

On le voit, lors de la phase où nous avons traduit, nous sommes constamment passé d'un dictionnaire unilingue à l'autre, dans le dédale des définitions qu'ont élaborées des générations de lexicographes<sup>166</sup>.

Dans nos données, nous envisageons la polysémie des mots anglais du corpus ScienText, et non celle des équivalents de traduction dans les langues romanes, qui peuvent admettre à leur tour une ou plusieurs acceptions. En outre, lorsqu'un mot anglais a des équivalences multiples pour une même acception, nous les conservons. Pour cette raison, nous indiquons pour certains mots anglais monosémiques plusieurs équivalents possibles qui ne sont pas nécessairement synonymes.

Bien des études pourraient être menées sur la formulation des définitions, l'ordre dans lequel les acceptions sont définies, les nuances culturelles qui se devinent dans les connotations implicitement ou explicitement exprimées. Nous nous limiterons à l'examen des deux autres cas rencontrés.

### ***5.2.2.2 Latinité et grécité de mots anglais de spécialité : approche hypothético-déductive empruntée à l'intercompréhension***

Nous avons suivi une méthode a priori sourcière si la morphologie du mot anglais évoquait une origine grecque ou latine. Le biais potentiellement induit est contrebalancé par la vérification dans des dictionnaires unilingues ainsi que par le fait que près de la moitié des mots appartient au lexique de spécialité de la santé, domaine où les emprunts abondent.

Certains mots nous étaient inconnus, mais leur forme a orienté notre travail de traduction. Nous avons vu dans la première partie (Des analogies interlinguistiques liées aux contacts des langues en Europe de l'Ouest, p.81) que les emprunts mutuels entre l'anglais et les langues

---

165 <https://www.dicio.com.br/pedir/>

166 Le travail des lexicographes de la Treccani est d'une précision remarquable.

romanes sont très nombreux en raison des contacts de langues. La latinité ou la grécité de certains mots est encore accentuée au sein du domaine spécifique qui nous intéresse parce que les professionnels de la santé parlant des langues européennes ont massivement emprunté aux deux langues anciennes de référence. L'intuition de l'étymologie dépend de la conscience linguistique éveillée lors de la formation initiale et développée ensuite par des recherches personnelles. Aussi, dans le cas de mots anglais dont la morphologie évoque une origine latine ou grecque, nous nous attendions à trouver des équivalents de traduction analogues.

Voici un exemple où ce genre d'intuition a fonctionné. L'adjectif anglais *recombinant* a une morphologie qui évoque une racine latine / française. Nous avons cherché si un mot similaire est attesté en espagnol dans le *Diccionario de la Lengua Española*, où figure (spa) *recombinante* avec la même signification. De la même manière, la *Treccani* atteste de l'existence de (ita) *ricombinante*. Notre recherche a été infructueuse dans le *TLFi* et dans le *Wiktionnaire*<sup>167</sup>. Aussi avons-nous cherché dans le *Dictionnaire de l'Académie de Médecine*<sup>168</sup>, où nous avons trouvé (fra) *recombinant*. Enfin, (por) *recombinante* existe dans le dictionnaire *Dicio*.

Voici les outils spécialisés auxquels nous avons recouru :

- pour l'espagnol : le *Diccionario de genotoxicología*, de l'Universidad de Concepción (Chili) ; le *MedlinePlus*<sup>169</sup>, de la U.S. National Library of Medicine ; le *dicciomed*<sup>170</sup> de l'Universidad Salamanca ; le *Glosario Hablado de Términos Genéticos*<sup>171</sup>, du National Human Genome Research Institute (États-Unis) ; le *Diccionario-medico*<sup>172</sup>, de la Clinica Universidad de Navarra ; le *Vocabulario inglés-español de bioquímica y biología molecular*<sup>173</sup> (Saladrigas & Claros, 2002) ;
- pour le français : le *TERMIUM Plus*<sup>174</sup>, la banque de données terminologiques du gouvernement du Canada ; le *Dictionnaire médical de l'Académie de Médecine*<sup>175</sup>

---

167 [https://fr.wiktionary.org/wiki/Wiktionnaire:Page\\_d%E2%80%99accueil](https://fr.wiktionary.org/wiki/Wiktionnaire:Page_d%E2%80%99accueil)

168 <http://dictionnaire.academie-medecine.fr>

169 <https://medlineplus.gov/spanish/encyclopedia.html>

170 <http://dicciomed.eusal.es>

171 <https://www.genome.gov/glossarys/index.cfm?textonly=&search=targeting>

172 <http://www.cun.es/diccionario-medico>

173 [http://www.medtrad.org/panacea/IndiceGeneral/Pana9\\_tradyterm\\_mvsgc.pdf](http://www.medtrad.org/panacea/IndiceGeneral/Pana9_tradyterm_mvsgc.pdf)

174 <http://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-fra.html?lang=fra>

175 <http://dictionnaire.academie-medecine.fr>

(France), le site Cours-Pharmacie.com<sup>176</sup> ; les dictionnaires de iH<sup>177</sup> (Hôpital Lavoisier, France) ; le portail biologie cellulaire<sup>178</sup> et moléculaire de Wikipédia ;

- pour l'italien : les *Lezioni di genetica medica*<sup>179</sup>, de Vincenzo Nigro, professeur à l'Università Vanvitelli (Naples) ; le *Glossario*<sup>180</sup> du site BIOteach, de l'Università degli studi di Milano ; le *Glossario*<sup>181</sup> de Hillis *et al.* (2013) ; et un cours de chimie et de technologie pharmaceutique<sup>182</sup> (Università di Roma 1) ;
- pour le portugais : une ressource sur la biologie moléculaire<sup>183</sup> qui n'est hélas plus en ligne, mais qui était proposée sur le site de l'Universidade Federal de Pernambuco (Brésil), le *DeCS*<sup>184</sup> (Descritores em Ciências da Saúde), du Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde ;
- les ressources multilingues suivantes ont aussi été consultées : l'excellent *HeTOP*<sup>185</sup> (Health Terminology/Oncology Portal) du Centre Hospitalier Universitaire de Rouen, le *United Medical Dictionary*<sup>186</sup> de l'Organisation Mondiale de la Santé (OMS) ; le site *IATE* (Interactive Terminology for Europe), de l'Union Européenne.

Certaines de ces ressources ont servi à comprendre les mots à traduire ou à confirmer des traductions douteuses. Par ailleurs, nous ne mentionnons pas toutes les ressources utilisées au hasard de nos vérifications.

Outre les outils spécialisés, nous avons privilégié les dictionnaires généralistes : nous avons ainsi pu noter quels termes de spécialité n'avaient pas pénétré les dictionnaires généralistes au moment où nous avons mené ce travail. Un astérisque (\*) dans les tableaux de données signale les mots espagnols, français, italiens et portugais non attestés dans les quatre dictionnaires généralistes que nous avons principalement utilisés pour vérifier le sens.

---

176 Par exemple : <http://www.cours-pharmacie.com/biologie-moleculaire/transcription-de-ladn.html>

177 <http://www.informationhospitaliere.com/dicopharma.php>

178 [https://fr.wikipedia.org/wiki/Portail:Biologie\\_cellulaire\\_et\\_mol%C3%A9culaire](https://fr.wikipedia.org/wiki/Portail:Biologie_cellulaire_et_mol%C3%A9culaire)

179 <http://www.vincenzonigro.it/Lezioni%20genmed%20Spec.2014a.pdf>

180 <http://bioteach.ariel.ctu.unimi.it/v3/home/PreviewArea.aspx?name=Glossario>

181 [http://online.universita.zanichelli.it/hillis-files/appendici/Hillis\\_Glossario.pdf](http://online.universita.zanichelli.it/hillis-files/appendici/Hillis_Glossario.pdf)

182 [https://elearning2.uniroma1.it/pluginfile.php/97686/mod\\_resource/content/4/12%20splicing%20RNA.pdf](https://elearning2.uniroma1.it/pluginfile.php/97686/mod_resource/content/4/12%20splicing%20RNA.pdf), le nom de l'auteur(e) n'est pas indiqué.

183 [https://www.ufpe.br/biolmol/aula3\\_RNAtranscri.htm](https://www.ufpe.br/biolmol/aula3_RNAtranscri.htm), mais le lien ne fonctionne plus.

184 <http://decs.bvs.br/cgi-bin/wxis1660.exe/decsserver/>

185 <http://www.hetop.eu/hetop/RESTR>

186 <http://www.emro.who.int/fr/Unified-Medical-Dictionary.html>



Prenons un autre exemple, où l'approche a été moins heureuse et a dû être complétée. Pour trouver les équivalents du nom anglais *genomics*, nous avons vérifié si les équivalents de ce mot dans les langues romanes étaient similaires à l'anglais. Nous trouvons (ita) *genomica*, (fra) *génomique*, et (spa) *genómica*, mais (por) \**genômica* n'est pas attesté dans le *Dicio*, nous avons donc usé d'une approche par le profil combinatoire.

Parmi les collocatifs de (eng) *genomics* dans ScienText figurent selon le Lexicoscope les adjectifs *comparative*, *functional* et *structural*. Ces adjectifs sont traduits en portugais par des mots analogues, respectivement (por) *comparativo*, *funcional* et *estrutural*. Une requête sur google associant les mots (por) \**genômica* et *comparativa*, « *genômica comparativa* », donne de nombreux résultats<sup>187</sup>, nous en déduisons que *genomics* équivaut à *genômica* en portugais.

La vérification de l'existence de certains mots pose problème quand ils sont susceptibles d'avoir la même forme en espagnol et portugais (et éventuellement en italien).

Ainsi, le suffixe *-ine* de (eng) *serine* suggère d'autres mots tels que (eng) *dopamine*, *cysteine*, etc., qui ont des formes analogues en *-ina* dans ces trois langues. *Sérine* existe en français, et *serina* en est l'équivalent en italien et en portugais, mais ce mot est absent du *Diccionario de la Lengua Española*.

Nous avons alors cherché dans le Lexicoscope les collocatifs de (eng) *serine* et les équivalents de traduction de ces derniers. Ce sont (eng) *kinase*, *phosphorylation*, *protease*, *protein*, *proteinase* et *residue*. Nous avons choisi un collocatif analogue : (eng) *residue*, traduit en espagnol par *residuo* et avons vérifié si la séquence collocative (spa) « *residuo de serina* » est attestée sur Internet. Comme des pages à contenu de spécialité en témoignent, (spa) *serina* est probablement l'équivalent de (eng) *serine*. De l'examen des premières pages, on constate cependant que certains résultats sont en portugais du fait de l'homographie de *serina* en espagnol et en portugais mais aussi de la grande similitude entre (por) *resíduo* et (spa) *residuo*. Nous avons donc réitéré la requête en ajoutant au début le déterminant défini espagnol *el*. L'expression « *el residuo de serina* » est attestée, ce qui confirme que (spa) *serina* traduit (eng) *serine*.

---

187 33800 résultats le 04/10/16.

### 5.2.2.3 *Équivalents de traduction manquants : recherches croisées*

Pour les autres mots dont nous ignorions la traduction, nous avons été amené à croiser les recherches en recourant à plusieurs outils selon la complexité des cas.

Nous nous sommes d'abord inspiré des traductions qui figuraient dans le *Wiktionary* : nous avons vérifié la pertinence de ces traductions dans les dictionnaires unilingues cités.

Le *Wiktionary* s'est avéré là encore un outil bienvenu dans une perspective plurilingue : les équivalents de traduction figurent par acception, en parallèle avec les traductions dans les autres langues, ce qui permet au traducteur de s'inspirer des équivalents de traduction des langues traduites qu'il connaît pour confirmer / mettre en doute / rectifier la traduction proposée.

Que la traduction soit présente et correcte, manquante ou erronée, elle a été systématiquement vérifiée dans les dictionnaires unilingues. Une garantie d'exactitude n'existe cependant pour aucun des outils, aussi spécialisés soient-ils. Ainsi, HeTOP donne pour équivalent italien de (fra) *morbidity* (ita) *morbidity*, alors que le terme équivalent est *morbosità*<sup>188</sup>.

### 5.2.2.4 *Termes non traduits dans le Wiktionary : requêtes dans Wikipédia et dans un moteur de recherche*

Lorsqu'un mot anglais n'était pas traduit dans le *Wiktionary*, la plupart du temps un terme au contenu sémantique très spécifique, nous avons cherché s'il faisait l'objet d'une entrée dans l'encyclopédie *Wikipédia* en anglais<sup>189</sup>. S'il y figurait, nous avons choisi par le menu de gauche la langue dont l'équivalent de traduction manquait, afin d'arriver à une page dans cette langue et obtenir le mot équivalent. Si cette page existait, nous avons vérifié ensuite la présence de cet équivalent dans le dictionnaire unilingue. Nous avons surtout recouru à des requêtes dans un moteur de recherche quand ni les dictionnaires, ni les ressources spécialisées consultés n'attestaient de l'acception d'un mot, voire de son existence.

Le domaine d'appartenance des sites sur lesquels apparaît le mot dans la première page est un guide dans le foisonnement des résultats : si le mot apparaît dans des pages de sites qui publient manifestement des données scientifiques, nous avons estimé que le critère de la spécificité de la publication associé à une place dans la première page des résultats

---

188 <https://frama.link/1f2uCF7k> et <http://www.treccani.it/vocabolario/morbosita/>.

189 [https://en.wikipedia.org/wiki/Main\\_Page](https://en.wikipedia.org/wiki/Main_Page)

présentaient une garantie suffisante pour considérer que c'était la traduction recherchée dans la langue romane en question<sup>190</sup>.

## 5.2.3 Exemple du processus de traduction : le nom *assay*

### 5.2.3.1 Les équivalents de traduction du Wiktionary

Les collocatifs du nom (eng) *assay* obtenus par le profil combinatoire (cf. 4.4.1.3, p.144) indiquent qu'il se rencontre dans le corpus ScienText avec l'acception suivante :

*The qualitative or quantitative chemical analysis of something*

Si on affiche l'élément *Translations* de la page du Wiktionary<sup>191</sup>, on observe des équivalents de traduction qui ont été proposés par des internautes. Ces mots apparaissent ensemble, dans toutes les langues déjà traduites, comme on le voit sur cette capture d'écran (image 10) :

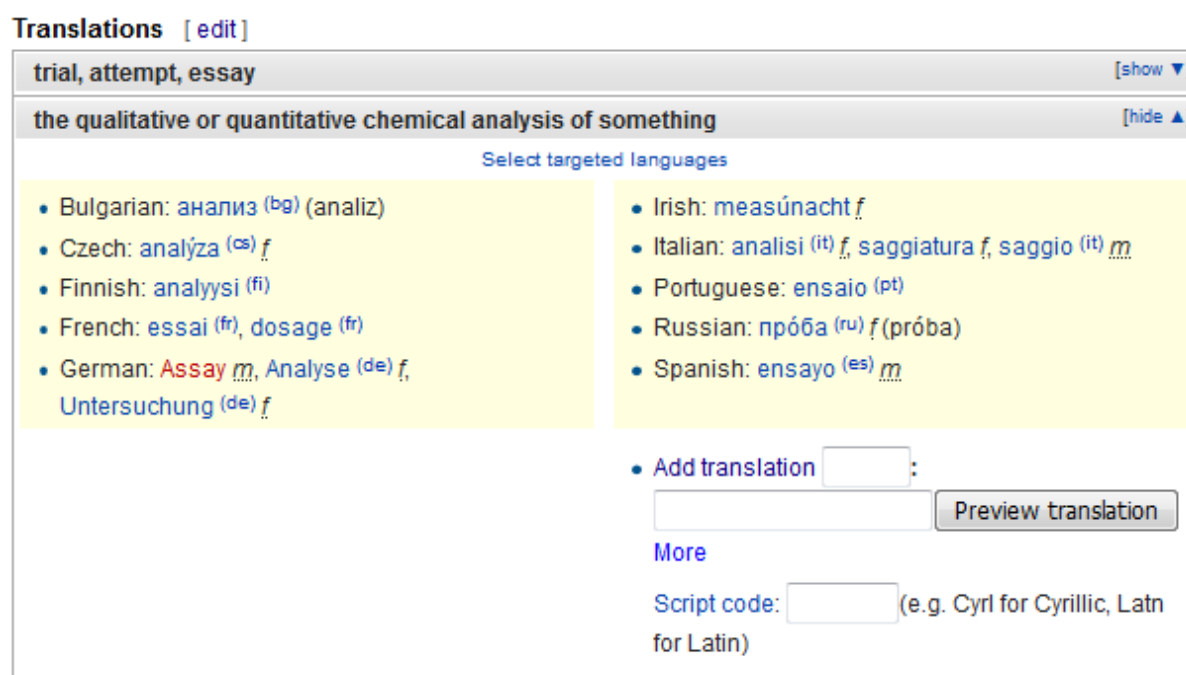


Image 10 : Copie d'écran du cadre contenant les traductions de (eng) *assay* dans le Wiktionary

190 Parfois, un mot recherché n'était attesté dans google que quelques dizaines de fois, mais exclusivement dans des pages à contenu spécialisé. Nous nous sommes risqué à accepter cet équivalent de traduction.

191 Nous avons contribué à cette partie du Wiktionary en y ajoutant des traductions en espagnol, français, italien et portugais.

### 5.2.3.2 Traduction d'assay en français

Deux mots français sont donnés comme équivalents de *assay*. Le mot (fra) *essai* présente une analogie formelle très forte avec (eng) *assay*, mais c'est un mot au sens très large, si bien qu'il appelle à la prudence. Par ailleurs, (fra) *dosage* est défini ainsi dans le TLFi<sup>192</sup> :

*CHIM., PHARM. Action de doser, de déterminer avec précision la quantité d'une substance (y compris dans un mélange)*

Pourtant, cette définition, qui indique que la substance est déjà connue, ne met en avant que l'objectif quantitatif du procédé. Du reste, (eng) *dosage* existe, qui est plus spécifiquement employé dans le langage pharmaceutique. Ce terme est défini ainsi dans le *Wiktionary*<sup>193</sup> :

*The administration of a medication etc, in a measured amount ; dosing*

Mais les collocatifs de (eng) *assay* ne limitent pas son usage à des médicaments. Une recherche dans *Wikipédia*<sup>194</sup> le confirme :

*« An assay is an investigative (analytic) procedure in laboratory medicine, pharmacology, environmental biology and molecular biology for qualitatively assessing or quantitatively measuring the presence or amount or the functional activity of a target entity (the analyte) . »*

Le Dictionnaire médical de l'Académie de Médecine<sup>195</sup>, montre que le terme (fra) *dosage* semble être utilisé dans la plupart des cas. Si nous entrons dans le champ de recherche *assay*, nous obtenons les termes dont la traduction en anglais contient *assay*. Aussi avons-nous :

- dosage radio-immunologique : *radioimmunoassay* : *Épreuve biologique (...)* ;
- dose biologique : *bioassay* : *Méthode de dosage quantitatif d'une substance (...)* ;
- ELISA (technique) : ELISA (test), défini par : *Technique de dosage (...)* ;
- Farr (test de) : *Farr's assay* : *Procédé de recherche et de quantification des (...)* ;
- immunocapture : *sandwich type assay* : *Technique de dosage immunologique (...)* ;
- radio-immunodosage : *radio-immunoassay* : *Technique de dosage (...)* ;

---

192 <http://atilf.atilf.fr/dendien/scripts/tlfiv5/advanced.exe?8;s=651818505> ;

193 <https://en.wiktionary.org/wiki/dosage>

194 <https://en.wikipedia.org/wiki/Assay>

195 <http://dictionnaire.academie-medecine.fr/?q=assay>

- radio-immunoessai : *radioimmunoassay* : *Anglicisme parfois utilisé pour les dosages radio-immunologiques.*

Les mots-vedettes de ce dictionnaire ainsi que leurs définitions indiquent par conséquent que (eng) *assay* est traduit par (fra) *dosage* la plupart du temps. Nous remarquons cependant l'hyperonyme (fra) *épreuve* ainsi que (fra) *essai*, présenté comme un anglicisme. Nous gardons donc *dosage* dans nos listes.

### 5.2.3.3 Traduction d'*assay* en italien

Trois termes italiens sont donnés comme traduisant *assay* :

- (ita) *analisi* paraît a priori trop imprécis, nous l'écartons.

- (ita) *saggiatura*, défini comme :

*L'operazione di saggiare, soprattutto metalli preziosi.*

n'est manifestement pas adapté au contexte de la santé.

- (ita) *saggio*<sup>196</sup> n'est pas éloigné du précédent :

*In chimica, denominazione generica di metodi di analisi, di tipo prevalentemente qualitativo ; in passato il termine fu particolarmente usato in riferimento all'analisi di minerali, di metalli e loro leghe.*

Nous recourons alors au dictionnaire contextuel multilingue<sup>197</sup> Reverso<sup>198</sup> où nous entrons (eng) *assay*. Parmi les résultats, nous retrouvons *saggio* et *analisi*, mais aussi *dosaggio*, qui rappelle celle du français *dosage*. Voici sa définition<sup>199</sup> :

Determinazione della percentuale in cui uno o più componenti si trovano in un composto o in un miscuglio, o in cui un soluto è presente in una soluzione ; in partic., nelle analisi chimico-cliniche, la determinazione della concentrazione di una sostanza per unità di volume (di un liquido biologico) o per unità di peso (di un tessuto biologico).

Il s'agit donc de l'équivalent italien de *assay*, aussi consignons-nous ce terme.

---

196 <http://www.treccani.it/vocabolario/saggio2/>

197 Ils ont été indispensables dans cette recherche. Nous avons également utilisé Linguee,

198 <http://context.reverso.net/traducao/ingles-italiano/assay>

199 <http://www.treccani.it/vocabolario/dosaggio/>

### 5.2.3.4 Traduction d'assay en portugais

Nous vérifions le mot (por) *ensaio* dans le Dicio<sup>200</sup> :

*Química Análise de um produto químico*

Cette définition est encore plus large que celle de (eng) *assay*, ce qui génère un certain désarroi. Nous nous tournons alors vers d'autres outils.

Reverso<sup>201</sup> indique parmi les traductions possibles (por) *ensaio*, mais aussi *doseamento*. En cliquant sur ce dernier, nous obtenons les alignements pour le couple (eng) *assay* – (por) *doseamento*. Les contextes (por) *doseamento de todos os componentes*, *doseamento cromogénico*, *doseamento da substância activa*, *doseamento de tolerância da glucose*, réfèrent à la pharmacologie.

Pour vérifier si le mot est aussi utilisé en biochimie, nous effectuons une requête en combinant dans le moteur de recherche google (por) *doseamento* avec la traduction du collocatif *enzyme* de *assay*, c'est-à-dire (por) *enzima*.

Le groupe nominal (por) « *doseamento de enzimas* » donne 2840 résultats au 29/10/16. Par comparaison, (por) "*ensaio de enzimas*" donne 6 résultats.

Dans un deuxième temps, afin de vérifier si (por) *doseamento* est également usité en biologie interventionnelle, nous menons le même type de recherche avec la traduction de PCR (Polymerase Chain Reaction), le même acronyme dans la langue de Camões. L'expression « *doseamento de PCR* » donne 8 résultats, « *ensaio de PCR* » 20800 résultats.

Suite à ces deux vérifications, nous gardons *doseamento* et *ensaio* comme équivalents de *assay*.

La méthode suivie pour traduire *assay* en espagnol est la même. Nous ne la détaillons pas.

## 5.2.4 Conséquences des différences lexicosémantiques interlinguistiques

### 5.2.4.1 De la difficulté de traduire les mots anglais courants

Il paraîtrait illusoire de vouloir donner un équivalent qui traduise tous les sèmes de certains mots usuels et donc marqués culturellement. Leur traduction en est d'autant plus délicate : (eng) *hand* est si usité qu'on le trouve dans des locutions telles que (eng) *at hand*, *on one*

200 <https://www.dicio.com.br/ensaio/>

201 <http://context.reverso.net/traducao/ingles-portugues/assay>

*hand, in our hand, etc.*, dont il n'existe pas nécessairement un équivalent dans les langues cibles qui contiennent la traduction usuelle de (eng) *hand*.

Les verbes (eng) *be, draw, make, set*, sont si polysémiques que nous les avons traduits par des verbes tout aussi polysémiques et peu précis dans les langues romanes cibles.

De même, (eng) *household* et les équivalents que nous proposons se distinguent probablement aux yeux du droit dans les États qui utilisent ces mots.

Pourtant, certaines différences tendent à s'atténuer en raison d'une convergence culturelle. On lit à ce sujet les influences interlinguistiques dans les indications étymologiques, lorsqu'elles figurent dans les dictionnaires unilingues<sup>202</sup>. Plus généralement, on peut penser que le développement de la traduction conduira à une plus grande homogénéité sémantique, qu'on le déplore ou au contraire qu'on s'en félicite. Il n'en demeure pas moins que certains mots anglais que nous avons examinés se distinguent dans l'extension de leur signification par rapport à leurs équivalents dans les langues romanes.

#### **5.2.4.2 Mots anglais dont l'acception est plus étendue**

Comme toute langue, l'anglais possède des mots dont la signification est plus large que celle du mot de la langue cible qui semble pourtant lui correspondre le plus. Ainsi, le nom (eng) *clinician* englobe tout personnel de soin, alors que le nom (fra) *clinicien* fait majoritairement référence à des médecins.

Un tel mot correspond donc à des équivalents distincts. L'hyperonyme (eng) *facility* en est un bon exemple. Pour ce mot, les langues romanes distinguent l'outillage et le lieu où il est utilisé. Aussi le sens de *facility* est-il rendu tantôt par des mots comme (fra) *matériel* et *équipement*, tantôt par d'autres comme (fra) *installations* (utilisé au pluriel).

D'autres mots l'illustrent, comme (eng) *know*, correspondant à deux mots dans les langues romanes : (fra) *savoir* et *connaître*, (ita) *sapere* et *conoscere* ; (eng) *immune*, traduit en français par deux termes, *immunitaire* et *immun*.

Dans le domaine du LST, il semble d'usage de distinguer pour (eng) *plot* deux termes dans les langues romanes : (spa) / (por) *diagrama*, (fra) *diagramme*, (ita) *diagramma*, et (spa) *gráfica*, (fra) *graphique*, (ita) *grafico*, (por) *gráfico*.

---

202 Pour (por) *familial*, préféré dans le domaine médical à *familiar*, on peut lire <http://www.jmrezende.com.br/familial.htm>.

De la même façon, (eng) *training* se traduit d'au moins deux façons différentes selon la langue, selon qu'il contient l'idée de l'appropriation de nouveaux savoirs, et les langues romanes utilisent pour cela un terme tel que (spa) *formación*, ou bien dans la langue de spécialité, l'idée de réactivation d'une compétence affaiblie par une affection ou une intervention chirurgicale, autrement dit (fra) *rééducation*, (ita) *addestramento*, etc.

On observe qu'un mot adopte un sens plus restreint dans certains contextes. Ainsi, (eng) *use* associé à une substance (médicamenteuse ou non) signifie (fra) *consommation*.

La spécialisation du sens entraîne parfois plusieurs équivalents de traduction :

- le terme (eng) *node* correspond à deux termes dans les langues romanes : (spa) *nódulo*, (fra) *nodule*, (ita) *nodulo*, (por) *nódulo*, d'une part, (spa) / (ita) *ganglio*, (fra) *ganglion* et (por) *gânglio*, d'autre part. À (eng) *record*, on a certes des mots tout aussi généraux que (ita) *cartella*, etc., mais dès qu'il s'agit de (eng) *medical record*, les termes varient : (spa) *historial médico*, (ira) *libretto sanitario*.

- (eng) *disorder* est un terme plus vague dans son extension sémantique que les équivalents que nous proposons dans les quatre langues romanes : (spa) *desorden*, (fra) *désordre*, (ita) *disordine*, et (por) *desordem*, traduisent des symptômes plus sérieux que (spa) *trastorno*, (fra) *trouble*, (ita) *disturbo* et (por) *transtorno*.

- (eng) *cord* n'a pas le même équivalent de traduction selon qu'il s'agit de *spinal cord*, de *umbilical cord* ou de *vocal cord*, nous devons donc faire figurer dans nos données les termes spécifiques qui équivalent à (eng) *cord* dans l'un de ces noms composés : (spa) *cuerva / cordón / médula*, (fra) *corde / cordon / moelle*, (ita) *corda / cordone / midollo*, (por) *corda / cordão / medula*.

Un mot de la langue courante, (eng) *office* apparaît dans le corpus dans l'une de ses acceptions qui désigne de façon elliptique le bureau dans lequel a lieu la consultation du médecin, autrement dit (spa) *consultorio*, (fra) *cabinet*, (ita) *studio* et (por) *consultório*.

Quand il était possible de traduire un même mot anglais par plusieurs mots dans l'une ou l'autre des langues romanes, il nous a paru pertinent de faire figurer plusieurs mots. Nous l'expliquons dans quel ordre nous les faisons apparaître dans la section Ordre des traductions multiples dans les tableaux, p.186.



### 5.2.4.3 *Mots anglais dont l'acception est moins étendue que les langues romanes*

Dans d'autres cas, ce sont les mots anglais qui ont une acception moins étendue que les langues romanes, ce qui entraîne la présence de plusieurs équivalents de traduction identiques là où l'anglais utilise plusieurs mots.

Les mots (spa) / (por) *contexto*, (fra) *contexte*, (ita) *contesto*, traduisent plusieurs mots anglais : *background*, *context* et *setting*.

On retrouve la spécialisation de certains mots, tels que (eng) *yield*, après lequel vient le produit d'une réaction chimique, mais qui n'a pas d'équivalent précis dans les langues romanes à notre connaissance.

Nous ne multiplierons pas les exemples ici. Nous invitons le lecteur à se référer aux feuilles déposées en ligne, sur le portail hal.

## 5.2.5 **Conséquences des divergences lexicosyntaxiques**

### 5.2.5.1 *Traduction des particularités lexicosyntaxiques des adjectifs anglais*

On sait qu'en anglais, on construit à loisir des noms composés, mais également des adjectifs composés, des *compounds*. La langue de spécialité de la santé en forme également de nombreux (Carnet, 2000 : 85-103), mais ils sont difficilement identifiables dans le corpus ScienText, hormis si l'on procède à des requêtes avancées, ce que nous avons évité autant que possible dans cette thèse afin de ne pas encore alourdir le travail.

Ainsi, pour (eng) *dependent*, nous avons constaté que la traduction peut suivre des schèmes différents : (eng) *hormone dependent* est traduit par (fra) *hormono-dépendant*, (eng) *cycline dependent* par (fra) *cycline-dépendant*, mais (eng) *position dependent effect* le serait par (fra) *effet dépendant de sa position*.

Nous devons également envisager la question des adjectifs de sens négatif de l'anglais, les néologismes qui commencent par le préfixe *un-*. Les langues romanes ne sont pas aussi souples, et des adjectifs tels que (eng) *untreated* ont une traduction qui requiert deux mots dans les quatre langues romanes : (spa) *no tratado*, (fra) *non traité*, (ita) *non trattato*, (por) *não tratado*.

Mais cette particularité des adjectifs anglais se retrouve dans d'autres catégories grammaticales dans l'une ou l'autre langue, et d'autres phénomènes féconds s'observent : la

composition au moyen du préfixe *dé-*, par exemple, est utilisée en français pour traduire certaines occurrences de (eng) *removal* dans des noms composés : (fra) *détatouage* pour (eng) *tattoo removal*, (fra) *déleucocytation* pour (eng) *blood leukocyte removal*, etc.

### **5.2.5.2 Multiplicité des constructions syntaxiques pour un même mot**

Certains mots anglais fréquents se rencontrent dans plusieurs constructions syntaxiques, ce qui génère une difficulté à les traduire parce que la liberté syntaxique de ces mots anglais trouve difficilement des équivalents dans les quatre langues romanes.

Si nous prenons l'exemple de l'adjectif (eng) *free*, nous constatons qu'il est employé comme adjectif épithète, par exemple dans *free radicals*, (fra) *radicaux libres*, comme attribut, dans (eng) [*is*] *free of contamination*, (fra) *non contaminé*, ou encore comme suffixe privatif incorrectement étiqueté par TreeTagger comme adjectif dans (eng) *pathogen free facilities*, mis pour (eng) *pathogen-free facilities*, c'est-à-dire (fra) *installations dépourvues d'éléments pathogènes*). Comme nous constatons que cette dernière construction prévaut, nous avons pris le parti de traduire l'adjectif (eng) *free* par des prépositions dans les quatre langues romanes.

### **5.2.5.3 Mots n'entrant que dans des figements**

Que faire en outre des mots qui ne semblent entrer que dans des figements tels que les noms composés ? Nous avons déjà précisé (section Homogénéisation de la graphie, p.141) que des termes tels que (eng) *amino (-acid)* ou (*Escherichia coli*), indissociables d'un autre composé, ont été considérés comme un seul lemme.

D'autres combinaisons de mots s'apparentent à des figements : (eng) *room* est majoritairement employé avec son collocatif (eng) *temperature*, et l'ensemble se traduit dans les quatre langues romanes par un nom et un adjectif : (spa) / (ita) / (por) *temperatura ambiente*, (fra) *température ambiante*.

Faudrait-il alors considérer ces couples comme des mots composés qu'il faudrait alors traduire dans leur globalité, comme un seul mot, eux aussi ? Mais comment trouver les occurrences d'une forme qui appartiennent au nom composé et celles qui n'y appartiennent pas ? Comment distinguer les occurrences de (eng) *force* de celles de (eng) *task force*, *driving force*, *labor force* ?

Nous avons rencontré peu de cas où un mot se rencontre exclusivement en combinaison avec un collocatif. (eng) *nucleic acid* en fait partie, il conviendrait alors ne pas traiter *nucleic* comme un adjectif fréquent, mais *nucleic acid* comme un des noms anglais fréquents.

En règle générale, nous avons traduit un nom anglais par son équivalent, même dans les cas où l'usage privilégie ce nom dans des compounds. Notons cependant les limites de la traduction de mots isolés : le mot (eng) *control* est traduit par (spa) *control*, (fra) *contrôle*, (ita) *controllo*, (por) *controle* alors que dans la plupart des cas, il se traduirait par un groupe nominal : (spa) *de control*, (fra) *de contrôle*, etc.

Enfin, certains noms composés anglais donnent lieu dans les langues romanes à un seul terme : (eng) *adverse drug monitoring*, (fra) *pharmacovigilance*, (ita) *farmacovigilanza*.

#### **5.2.5.4 *Autres cas de divergences syntaxiques***

Nous aborderons très peu ici la possibilité de traduire un mot appartenant à une classe grammaticale par un mot appartenant à une autre classe. L'anglais *fat*, souvent étiqueté comme adjectif alors qu'il est en réalité un nom, peut se traduire par un adjectif dans les langues romanes.

La signification superlative de quelques mots se pose également : le mot (eng) *major* est davantage usité que le mot (fra) *majeur* : s'il est employé dans le sens de (eng) *great*, de *greater* ou de *greatest*, il est plus naturellement traduit au moyen de (fra) *plus* ou (fra) *le plus*, et d'un autre adjectif, tel (fra) *plus important*. Il en est de même dans les autres langues romanes.

Nous ne ferons que signaler l'existence de rares cas où dans l'une des langues, c'est le pluriel qui est plus fréquent que le singulier anglais, à l'instar de (eng) *care* et de (fra) *soins*.

Une approche par lemmes fréquents gomme la présence de noms dans des locutions conjonctives telles que (eng) *despite the fact that*, nous pouvons souligner cette limite.

## 5.2.6 Convergence sémantique interlexicale

Un mot polysémique peut être traduit de la même façon pour ses deux ou trois acceptions, éventuellement dans les quatre langues. C'est ce qui explique qu'on ait dans le corpus 1944 mots à une traduction (cf. Analogie des 1944 mots anglais à une traduction, p.223). Voici deux exemples de ce type de mots.

### 5.2.6.1 *Le même rapport sémantique : l'exemple d'early*

*Early* peut avoir pour signification :

Acception 1) *Early* : '(medicine) Having begun to occur ; in its early stages'.

C'est le sens dans les textes du corpus ScienText lorsque cet adjectif a pour cooccurrents des noms qui expriment des affections ou des actions qui font partie de l'acte médical (*arthritis, cancer*), ou bien qui servent à décrire le caractère évolutif d'une affection (*period, phase, detection*).

Pour d'autres occurrences, il s'agit de :

Acception 2) *Early* : 'At a time in advance of the usual or expected event'.

Dans ce cas, on perçoit le jugement du spécialiste qui tente de tirer des enseignements de l'observation qu'il fait de quelque chose de statistiquement aberrant : *death, pregnancy, evolution, onset*.

Les deux acceptions de (eng) *early* se traduisent par le même mot dans les langues romanes : (spa) *precoz* / (fra) *précoce* / (ita) *precoce* / (por) *precoce*.

Dans les dictionnaires unilingues, l'acception 1) se retrouve dans les définitions suivantes :

(spa) *precoz* : « Med. Pertenciente o relativo a las etapas tempranas de una enfermedad o proceso orgánico ».

(fra) *précoce* : « [En parlant d'un signe clinique] Dont les manifestations se produisent au stade initial d'une maladie. »

En outre, on trouve en italien et en portugais la collocation *fase precoce*<sup>203 204</sup>.

---

203 Au pluriel en italien ; [https://frama.link/N\\_0j\\_mW](https://frama.link/N_0j_mW) consulté le 12/08/15

204 Au singulier en portugais ; <https://frama.link/o9-DHckD> consulté le 12/08/15

L'acception 2 de *early* est présente également dans les équivalents de traductions *precoz*, etc., comme en témoignent les définitions suivantes :

(spa) *precoz* : « Dicho de un proceso: Que aparece antes de lo habitual ».

(fra) *précoce* : « « [Appliqué à une transformation ou à une altération physique visible] Qui apparaît prématurément. »<sup>205</sup>

(ita) « Che si sviluppa, che fiorisce, che raggiunge la maturità prima dell'epoca normale. »<sup>206</sup>

(por) « Que se produz antes do tempo normal. »<sup>207</sup>

### 5.2.6.2 Convergence des séquences collocatives spécifiques

Certains mots courants ont posé question parce qu'ils peuvent être combinés à des collocatifs spécifiques. La séquence collocative qui s'ensuit se traduit-elle alors en combinant les équivalents de traductions de chacun des deux mots ?

En guise de premier exemple, prenons le nom (eng) *pair*, avec l'acception courante de :

*'Two similar or identical things taken together (...)'*. Il se rencontre essentiellement dans le domaine de la génétique, avec des collocatifs comme (eng) *base*, *primer*, *gene*, *transcript*, *alignment*.

(eng) *pair* est un mot de la langue courante, mais sa traduction courante dans les langues romanes est-elle pertinente ? Nous savions que le français utilise *paires de gènes*, mais ignorions ce qu'il en est dans les trois autres langues romanes, aussi avons-nous cherché si l'usage a consacré (ita) « *paia di geni* », (spa) « *pares de genos* », et (por) « *par de genes* ».

Voici les résultats<sup>208</sup> :

« *paia di geni* » : 47 200

« *pares de genos* » : 317 000

« *par de genes* » : 156 000

---

205 Pour les deux citations précédentes en français, voir le *TLFi*.

206 <http://www.treccani.it/vocabolario/precoce> consulté le 12/08/15

207 <http://www.dicio.com.br/precoce> consulté le 12/08/15

208 Les résultats obtenus dans le moteur de recherche google datent du 09/10/16.

Ces nombres élevés confirment l'existence de ces séquences collocatives calquées, nous en déduisons donc que (ita) *paio*, et (spa) / (por) *par* sont des traductions de l'anglais *pair* dans ce contexte spécifique.

Voici un deuxième exemple. La page du *Wiktionary* consacrée à l'adjectif (eng) *complex* ne donnait pas (spa) *complejo* comme traduction pour l'acception définie comme :

*'Intricate ; composite ; not simple'*<sup>209</sup>.

Or nous étions dubitatif face à l'adjectif (spa) *complicado* qui était proposé pour cette acception car il relève de la langue courante dans laquelle il équivaut souvent à (spa) *complejo*. Nous avons vérifié les significations possibles de *complejo* dans le *Diccionario de la Lengua Española*. *'Que se compone de elementos diversos'*<sup>210</sup> semblait confirmer cette possibilité. Mais cette définition manque de précision et d'exemples pour illustrer ce que seraient ces éléments.

En complément, nous avons donc associé l'adjectif (eng) *complex* à son collocatif (eng) *disease*. La séquence collocative (eng) *complex disease* nous intéresse particulièrement en ce qu'elle réfère très spécifiquement au champ de la santé. Nous avons cherché<sup>211</sup> dans google les groupes nominaux qui en seraient les traductions hypothétiques en espagnol : « *enfermedad compleja* » donne 38 900 résultats, et le pluriel « *enfermedades complejas* » 44 600 résultats. Ce nombre élevé a confirmé que l'équivalent de (eng) *complex* est (spa) *complejo* pour le collocatif (eng) *disease* de cet adjectif anglais dans le corpus de ScienText.

Voici un dernier exemple. Nous nous sommes appuyé sur plusieurs collocatifs pour traduire (eng) *island*. Mais si on parle de (ita) *isola CpG*<sup>212</sup> pour traduire (eng) « *CpG islands* », le terme (ita) *isola* est-il utilisé pour (eng) « *pathogenicity island* » ? Les banques de données terminologiques que nous avons consultées ne donnant aucun résultat pour l'italien, nous avons entré la séquence collocative hypothétique (ita) \* « *isole di patogenicità* » dans google et nous avons trouvé des pages à contenu académique où cette expression était attestée, si bien que nous avons traduit (eng) *island* par (ita) *isola* pour cette acception.

---

209 Nous pourrions du reste critiquer cette définition, intéressante en ce qu'elle englobe le sens, quoique manquant de précision. La reformulation de toute définition jugée imprécise dépasse cependant amplement le cadre de ce travail.

210 <http://lema.rae.es/drae/?val=complejo>

211 Le 01/11/14.

212 [https://it.wikipedia.org/wiki/Isola\\_CpG](https://it.wikipedia.org/wiki/Isola_CpG) consulté le 13/04/15.

## 5.2.7 Conséquences de la variation intralinguistique

### 5.2.7.1 Variation diatopique des traductions des mots courants de la santé

Une variation diatopique existe dans les traductions en langues romanes des mots anglais fréquents, en raison d'aires géographiques distinctes, particulièrement en espagnol (entre l'espagnol d'Espagne et les espagnols d'Amérique latine) et en portugais (entre le portugais du Portugal et le portugais du Brésil ; les éventuelles divergences dans les autres pays lusophones ne transparaissent pas dans les sources que nous avons utilisées). Il nous semblerait délicat de gommer l'un ou l'autre des mots dans une perspective didactique dans la mesure où l'interlexique établi peut servir à un enseignement d'une autre langue romane que le français, en Europe ou dans des pays romanophones extraeuropéens, ce qui ne nous autorise pas à choisir entre par exemple (por) *despistagem*, terme qui vient de (fra) *dépistage* et semble s'utiliser plutôt dans l'Union Européenne, et (por) *rastreio*, tous deux traduisant (eng) *screening*.

Le choix de l'un ou l'autre des équivalents possible se complique d'indications prescriptives parfois contradictoires. Pour ce qui est de ces deux derniers termes, que nous allons prendre comme exemple de la complexité occasionnelle à choisir un équivalent de traduction, nous avons consulté deux dictionnaires unilingues de spécialité, le *Dicionário de termos médicos* (Freitas e Costa : 2005) et le *Dicionário Médico* (Manuila et al. : 2011).

Dans le *Dicionário de termos médicos*, on trouve comme mot-vedette (por) *despiste ou despitagem*, défini comme suit :

*Procura ativa, numa população, de sinais de uma doença que ainda não foi determinada. Designação incorreta do termo rastreio (forma preferível)*<sup>213</sup>

Comme *rastreio* n'est pas défini dans ce dictionnaire, on peut s'étonner que la définition de ce procédé apparaisse avec le mot-vedette dont il est dit qu'il faut le considérer comme une forme incorrecte.

Dans le *Dicionário Médico*, une distinction est établie entre les deux termes. *Despistagem* est défini comme :

*Procura activa numa população de sinais de doença latente (ainda não detectada), em regra com o auxílio de métodos simple e pouco onerosos. [...] Sin de rastreio*<sup>214</sup>

---

213 Freitas e Costas (2015 : 333-334).

214 Manuila et al. (2011 : 237).

Suivent des définitions de types de *despistagens* : *despistagem por exames sucessivos*, *despistagem precoce*, *despistagem selectiva*, *despistagem sistemática*, et pour ce dernier, trois formes sont définies.

*Rastreio* est donné comme :

*Actividade para identificar presuntivamente por meio de provas, de exames ou de outras técnicas os indivíduos portadores de uma doença ou anomalia até então desconhecida. O rastreio deve permitir fazer a distinção entre as pessoas aparentemente sãs, mas provavelmente afectadas, e as que provavelmente estão isentas dessa doença. A finalidade não é o diagnóstico. As pessoas com resultados positivos ou duvidosos devem seguir para diagnóstico e tratamento, se necessário.*<sup>215</sup>

Comme ce dernier dictionnaire donne les deux équivalents de traduction *despistagem* et *rastreio* comme synonymes, indépendamment de considérations prescriptives, nous avons préféré les conserver tous deux.

La prise en compte du portugais du Portugal nous a posé le plus question quant à la légitimité de faire apparaître dans nos données les termes spécifiquement brésiliens / portugais : le plus grand nombre de lusophones se trouvant au Brésil, il semblerait légitime d'adopter dans notre lexique les termes en vigueur dans ce pays, par exemple (por) *arquivo* pour (eng) *file*. Toutefois, comme nous avons prévu de rechercher des énoncés dans des corpus de textes européens tels qu'Europarl et EMEA, qui adoptent la graphie du Portugal, nous avons finalement opté pour les deux termes de manière à garder la possibilité de recourir au terme portugais, par exemple *ficheiro* pour (eng) *file*, lors des requêtes dans les corpus européens. Pour la même raison, nous conservons en espagnol les paires (spa) *fichero* / *archivo*, etc.

### **5.2.7.2 Variation diastratique des traductions des mots courants de la santé**

La variation diastratique se perçoit dans les articles de ScienText où il est fait mention de questionnaires posés à des patients. La description des symptômes par le patient / le clinicien, pour ne citer qu'un exemple, diffère grandement, et on sait qu'elle provoque une difficulté à prendre en charge efficacement l'affection.

Nous avons naturellement rencontré dans les ressources consultées des termes qui témoignent de la variation diastratique dans le domaine de la santé, à l'instar de (eng) *doctor* traduit par

---

215 Manuila *et al.* (2011 : 662).



(fra) *médecin* et *docteur*; et par (ita) *medico* et *dottore*; (por) *registo* au Portugal, au Mozambique, etc., et *registro* au Brésil. Comme le corpus reflète les usages attestés dans le langage écrit du spécialiste dans ses articles scientifiques, nous avons essayé de mettre en face des équivalents de traduction qui ressortissent à ces usages génériques. Mais il ne nous est pas possible ici de traiter de cette question plus avant.

Nous sommes conscient de l'existence d'un fossé entre l'usage du savant et l'usage qui prévaut dans la relation thérapeutique entre le soignant et le patient. Certains outils sont actuellement développés afin de rendre plus facile l'accès du patient aux connaissances et à l'information qui ont trait à la description de la maladie qui l'affecte ou qui affecte ses proches – par exemple le Large Rated Lexicon (Grabar & Hamon, 2016 : 2643-2648), qui repose sur des méthodes de segmentation de mots composés et sur l'évaluation par l'humain de leur intelligibilité.

### **5.2.7.3 Variation diachronique et processus de néologisation en cours**

Nous faisons figurer plusieurs équivalents de traduction dans une langue romane pour certains mots anglais, monosémiques ou polysémiques. Plusieurs types de variations intralinguistiques motivent ce choix.

En effet, pour quelques termes déjà ancrés dans l'usage, quelques variantes graphiques coexistent, du type (spa) *microscopia* / *microscopía*, (ita) *alcol* / *alcole*, (por) *registo* / *registro*. Nous avons conservé ces formes concurrentes sans toutefois approfondir ce point situé à la marge de notre étude. Serait-il utile de rechercher lesquelles sont véritablement adoptées dans l'usage (écrit et éventuellement oral) afin de trier les mots obtenus et gagner en discernement? Ces doublons peuvent n'être en réalité que le reflet d'une variation diachronique dont se font le témoin les dictionnaires qui se posent comme les garants des *bons usages*.

Pour des mots assez récents, l'hésitation dans les usages se perçoit par exemple entre les équivalents de (eng) *monitoring*, avec (fra) *surveillance* / *monitoring* / *monitorage*; l'usage semblant aller actuellement en français vers l'emploi des deux derniers, sur la base du mot anglais.

Parfois, c'est un peu comme si nous assistions à la naissance de mots : les auteurs hésitent à utiliser en français *microréseau* / *micromatrice* / *biopuce* pour (eng) *microarray*, lui-même

souvent abrégé en (eng) *array*. Ce qui pourrait refléter une insécurité linguistique traduit en même temps un formidable processus de négociation en cours parmi les usagers d'une part, entre les usagers et les instances prescriptives, d'autre part, que ce soit à l'oral ou à l'écrit.

Nous trouvons ainsi pour (eng) *genomic library* (ita) *biblioteca genomica*, qui est la forme prescrite, mais également (ita) *libreria genomica*, forme critiquée et pourtant plus fréquente si l'on en fait la requête dans un moteur de recherche.

Faut-il considérer sur un même plan les équivalents de traduction de ce type de mots ? Certains devraient-ils être privilégiés ? Nous ne nous positionnons pas sur ce sujet, les traducteurs choisiront. Nous nous contentons de faire figurer dans nos données ceux qui semblent ne pas avoir (encore) donné lieu à un consensus.

#### **5.2.7.4 Variation stylistique de l'anglais**

La traduction que nous proposons ne prétend pas être adaptée à tous les contextes : un mot comme (eng) *great* exprime avant tout l'intensité, mais *great* a été employé dans certains passages quand (eng) *major*, *high*, ou *large* auraient probablement pu l'être pour certaines des occurrences de *great*.

Le choix de certains mots anglais peut refléter un choix stylistique de l'auteur entre plusieurs mots qui ont un signifié très proche. La variation intralinguistique est particulièrement sensible dans le domaine du changement, avec des verbes intransitifs tels que (eng) *increase*, *rise*, *grow*, ou encore *insert*, *introduce*, lesquels coexistent par ailleurs avec *incorporate* et *integrate*. On retrouve cette possibilité d'utiliser l'un ou l'autre mot avec des verbes tels que *link* et *relate*, *implicate*, et *imply* (et dans une moindre mesure *involve*), ou encore avec des adjectifs tels que *same* et *identical*.

Nous trouvons particulièrement intéressante la perspective d'enseigner le lexique d'une langue (sur objectifs spécifiques, mais pas seulement) en partant de ces variantes validées par l'usage par une communauté de chercheurs qui l'utilisent dans un même genre textuel. De la même façon, et quelle que soit la langue, on peut se demander quels mots sont usités dans les textes scientifiques par rapport à ceux de la langue courante, par exemple (por) *similar* / *parecido* / *semelhante*.

Le tri des 2208 acceptions a permis de regrouper les mots qui ont un sens très proche, voire identique. Dans le fichier lexiqueTRANSDISCIPLINAIRE.ods, qui regroupe les mots du

LST, on peut trouver dans quels sous-domaines sémantiques la variation est forte, comparer le nombre d'occurrences de chacun des mots concernés, éventuellement s'en inspirer pour faire écrire des écrits scientifiques, ou faire trouver à ses étudiants dans un corpus des marques de la variation sémantique et ainsi affiner la perception des usages stylistiques.

## **5.2.8 Forme et ordre des mots traduits dans les tableaux de données**

### **5.2.8.1 Forme des verbes**

Nous avons conservé le verbe dans sa forme infinitive, mais avec la profusion possible de *phrasal verbs*, on regrettera que certains mots courants ne figurent pas dans la liste obtenue. La flexion du verbe anglais est restreinte à peu de formes, nous l'avons dit plus haut, mais ce n'est pas le cas dans les langues romanes. Nous avons conservé la forme infinitive, celle des dictionnaires, mais nous savons que des formes similaires pourraient être identifiées sur la base des autres formes qu'un lecteur rencontre en lisant les textes de spécialité : troisième personne du singulier du présent de l'indicatif, etc. Toujours au sujet des verbes anglais, on observe que certains figurant dans notre liste peuvent être à la fois transitifs et intransitifs mais que leurs équivalents de traduction n'ont pas la même forme selon qu'ils sont eux-mêmes transitifs ou intransitifs.

### **5.2.8.2 Ordre des traductions multiples dans les tableaux**

Lorsque la traduction que nous proposons d'un mot anglais fréquent du corpus ScienText est composée de plusieurs équivalents de traduction, nous écrivons en premier le mot qui a la forme la plus proche du mot anglais, ceci pour la simple raison que le script `compareCols.csv` (voir plus bas *Mesure de similarité par le script CompareCols.csv.pl*, p. 189) prend en compte l'ordre dans lequel figurent les caractères des mots comparés. Par conséquent, (eng) *monitoring* ne serait pas considéré comme similaire à un mot français si les trois équivalents possibles apparaissaient dans l'ordre (fra) *surveillance / monitoring / monitorage*, au lieu qu'ils apparaissent dans l'ordre *monitoring / monitorage / surveillance*, celui que nous avons appliqué. Les deux, trois, voire exceptionnellement quatre mots que nous indiquons comme possibles équivalents sémantiques sont séparés dans les tableaux par une barre oblique (/).

## **5.3 Mesures de similarité : vers un interlexique et quatre lexiques bilingues**

### **5.3.1 Un interlexique au niveau des acceptions et non des mots**

Mesurer l'analogie au niveau du mot n'a pas paru pertinent à cause de la polysémie de certains mots fréquents. En effet, comme on peut s'y attendre, certains mots polysémiques anglais admettent des traductions distinctes en fonction de l'acception. Il en résulte éventuellement un risque de confusion pour un apprenant qui peut croire que tous les sèmes des mots de l'anglais sont superposables aux sèmes des mots des langues romanes. En conséquence, comment conviendrait-il de traiter les mots dont l'équivalent de traduction dans une langue romane serait similaire au mot anglais pour l'une des acceptions quand il ne le serait pas pour une autre acception ?

L'acception est donc la base des calculs. La granularité des données que nous avons ne nous permet pas de distinguer parmi les occurrences d'un mot polysémique le nombre de celles qui réfèrent à l'une ou l'autre de ses acceptions. Nous avons donc basé les mesures de similarité sur un calcul tout théorique du nombre d'occurrences : le nombre d'occurrences des mots polysémiques à deux acceptions est divisé par 2, le nombre d'occurrences des mots polysémiques à trois acceptions est divisé par 3.

Accessoirement, si les deux ou trois acceptions d'un mot peuvent être traduites par un seul et même mot, l'analogie est exprimée sur la base de toutes ses acceptions du mot (cf. Analogie des 1944 mots anglais à une traduction, p.223).

L'interlexique anglais-espagnol-français-italien-portugais est établi en partant de l'anglais. Il est constitué des séries d'adjectifs, de noms et de verbes analogues, c'est-à-dire des séries de cinq mots qui présentent une identité sémantique (ou quasi-identité sémantique, cf. Conséquences des différences lexicosémantiques interlinguistiques, p.173) ainsi qu'une similitude formelle.

La similarité formelle des mots a été mesurée au moyen d'un programme qui compare la forme des mots anglais avec celle de leurs équivalents de traduction dans chacune des quatre langues romanes pour une même acception. Deux moyens complémentaires ont été mis en œuvre en fonction de la longueur des mots comparés. D'une part automatiquement au moyen

d'un script quand le mot comporte au minimum 4 caractères (dans les faits, quatre lettres), d'autre part manuellement lorsque le mot comprend moins de 4 caractères.

### 5.3.2 Neutralisation des diacritiques des langues romanes

Dans les quatre langues romanes de notre étude, certaines lettres comportent des signes diacritiques : des accents sur des voyelles, le tréma, le tilde et la cédille.

Nous avons supposé qu'un lecteur voit la même chose entre une lettre diacritisée et la même lettre non diacritisée, qu'il ne fait pas de différence par exemple entre les graphèmes *é* et *e*, ou encore entre *n* et *ñ*. Cette supposition induit que la présence d'un diacritique n'empêcherait pas la reconnaissance d'une lettre dans sa forme dénuée de diacritique. Certes, il peut sembler abusif de considérer que *n* et *ñ* ne sont pas deux lettres différentes. Elles transcrivent assurément des phonèmes différents, mais qu'on les imagine avec le point de vue d'un lecteur ignorant la correspondance graphophonémique de l'espagnol, ce sont alors deux graphèmes que peu de choses distinguent visuellement.

Pour tracer un parallèle, un lecteur qui ignore le vietnamien et se trouve face à un texte écrit dans cette langue peut observer que presque toutes les voyelles comprennent des diacritiques. En vietnamien, certains diacritiques indiquent l'aperture de la voyelle et d'autres le ton de la syllabe. Cependant, un lecteur non initié à cette langue lira la voyelle sans pouvoir lui donner la réalisation phonique attendue, alors que *bá* (le roi), *bã* (le marc de café), *ba* (trois), *bả* (la pâture), *bà* (la grand-mère) et *bạ* (n'importe quel) ne sont pas prononcés de la même façon (Trần, 2007, cité par Đỗ, 2011 : 50). L'utilisation de diacritiques, moins discriminante dans les langues romanes qu'en vietnamien, permet la plupart du temps d'indiquer l'accent tonique ou l'ouverture de la voyelle.

La mesure de similarité entre les mots de deux langues part donc de formes artificiellement simplifiées, sans diacritiques. Ces formes ne sont pas attestées dans les langues de notre étude, ce ne sont que des formes de travail.

### 5.3.3 Mesure de similarité par le script CompareCols.csv.pl

#### 5.3.3.1 Les trois critères appliqués pour mesurer la similarité

Le script CompareCols.csv.pl, conçu par Olivier Kraif, permet de comparer les sous-chaines de caractères des mots de plusieurs langues. Il identifie la sous-chaine commune maximale (SCM) en fonction de trois critères réglables :

- 1) la longueur minimale des mots à comparer ;
- 2) le décalage maximal existant entre deux caractères. Par exemple, dans la comparaison de *surgery* avec *chirurgie*, les deux premiers 'u' alignés sont décalés de 3 caractères (le premier en position 2 et le second en position 5).
- 3) le rapport de la plus longue sous-chaine commune vis-à-vis de la plus courte des deux sous-chaines (respectivement 11, 8 et 7 dans le graphique).

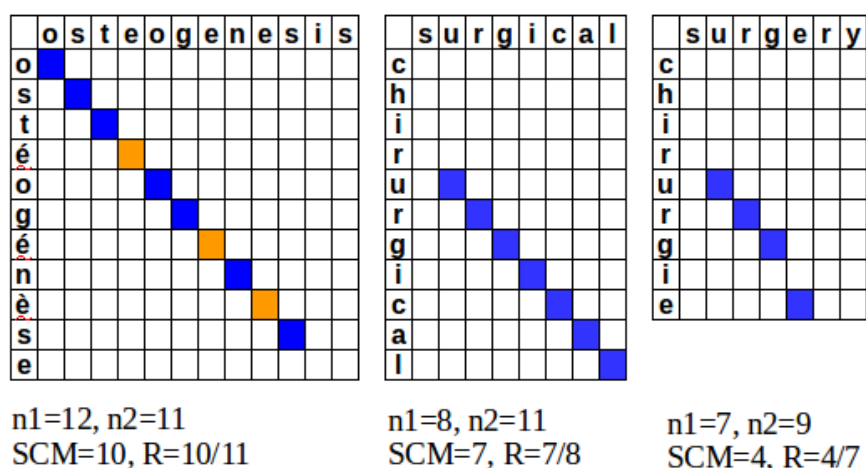


Image 11 : Trois exemples Sous-chaine commune maximale (SCM)

#### 5.3.3.2 Réglages successifs des critères

Initialement, nous avons appliqué une longueur minimale de 4 caractères, un décalage maximal de 3 et un rapport minimal de 60 %. De ce réglage, il résultait que les mots anglais étaient donnés comme similaires à un équivalent de traduction dans les quatre langues romanes pour 1261 de leurs 2208 acceptions, c'est-à-dire pour 57,11 %.

Cependant, certains couples de mots, comme (eng) *write* et (fra) *écrire*, faisaient partie des mots similaires alors qu'un lecteur ne les voit manifestement pas comme tels. En conséquence, les résultats reposent sur une comparaison faite avec un réglage plus restrictif :

- 1) longueur minimale des mots à comparer : 4 caractères ;
- 2) décalage maximal existant entre deux caractères : 2 ;
- 3) rapport minimal de la sous-chaine commune vis-à-vis de la plus courte des deux sous-chainnes : 0,7 (70%).

En complément, nous avons neutralisé les finales équivalentes.

### 5.3.3.3 *Neutralisation des finales équivalentes*

Dans le domaine de la traduction automatique, Lavallée et Langlais (2010 : 617-624) ont utilisé la segmentation morphologique en découpant les mots par analogie. Langlais et al. (2008 : 284-295) ont suggéré que soient forgées des équivalences terminologiques dans le domaine médical. Dans cette idée et afin d'obtenir des résultats plus consistants, nous avons cherché à faire prendre en compte au script les suffixes. Au mot suffixe, nous n'attribuons pas une définition linguistique mais informatique. Il s'agit là simplement d'une terminaison d'au moins deux caractères, indépendamment du radical du mot et du domaine de spécialité.<sup>216</sup>

Le script `sortReverse.pl` a permis d'effectuer un tri automatique par ordre inverse. Puis la finale des mots des quatre langues romanes a été observée manuellement afin de chercher si à un suffixe d'un mot traduit dans une langue romane donnée correspond préférentiellement un suffixe des 2000 mots anglais. Enfin, une liste des suffixes présents dans les traductions en espagnol, français, italien et portugais a été dressée. En leur ajoutant les suffixes auxquels ils correspondent en anglais, on obtient un tableau des allosuffixes, autrement dit un tableau des suffixes équivalents dans les cinq langues (cf. Annexe 4 : Tableau des allosuffixes, p.394). Afin de simplifier les formes, seules les paires de suffixes qui ont plus d'une lettre de différence ont été ajoutés au script `compareCols.csv.pl` : *-clude*, *-cluir*, *-clure*, *-cludere*, *-cluir*, mais pas *-al*, *-all-el*, *-al*, *-ale*, *-al*.

Les calculs de similarité obtenus suite aux nouveaux réglages et en prenant en compte les allosuffixes sont plus satisfaisants. Ainsi, (eng) *write* et (fra) *écrire*, ou (eng) *step* et (fra)

---

<sup>216</sup> Une liste des affixes utilisés en terminologie médicale se trouve dans le *Dorland dictionnaire médical bilingue* (Maniez, 2009 : ix-xxxiii).

*étape*, ne sont plus donnés comme similaires, ni (eng) *binding* et (spa) *unión*, (eng) *draft* et (spa) *redactar*, (eng) *mouse* et (fra) *souris*, (eng) *right* et (ita) *diritto*, (eng) *smoker* et (fra) *fumeur*, (eng) *light* et (spa) *ligero*, pour citer des exemples de couples de mots qui ne seraient certainement pas considérés comme similaires s'ils étaient comparés manuellement.

Il demeure néanmoins des cas contestables. Ainsi, selon ces nouveaux réglages, (ita) *ragione* et (eng) *ratio* seraient analogues, (ita) *diluire* serait analogue à (eng) *dilute*, mais pas (spa) et (por) *diluir*. Il serait intéressant de prendre en compte dans ces mesures les allographèmes afin que soient considérés comme analogues des mots comme (eng) *enzyme* et (spa), (ita) et (por) *enzima*, ou encore (eng) *unique*, (spa) et (por) *único*, et (ita) *unico*.

À l'issue de ces réglages, l'interlexique se compose de 1038 séries de cinq mots, autrement dit 47,01 % des 2208 séries sont analogues dans les cinq langues.

### 5.3.4 Étapes suivies pour mesurer la similarité formelle

Voici comment comment la similarité des mots des langues de cette étude a été mesurée au moyen du script CompareCols.csv. La feuille 2208ac.ods contient les 2208 acceptions fréquentes des 2000 mots anglais fréquents du corpus ScienText et les traductions en espagnol, français, italien et portugais pour ces acceptions. Nous avons copié ces données dans un nouveau fichier calc et avons suivi la procédure suivante :

- 1) suppression des éventuels espaces inutiles au moyen de la fonction =SUPPRESPEACE(A1), étendue aux 2208 X 5 cellules contenant les mots des 5 langues<sup>217</sup> ;
- 2) suppression des marques de formatage (sélection des données par CTR + A, puis clic droit, et effacer le formatage direct) ;
- 3) suppression des colonnes contenant les acceptions et la catégorie sémantique (sinon, la ponctuation présente dans les définitions des acceptions entraîne des problèmes de mise en page) ;

---

<sup>217</sup> Cette étape est très utile car il n'est pas rare qu'on ajoute un espace en recopiant les données depuis les ressources consultées. Or un espace est considéré comme un caractère, et le script mesure la similarité en s'appuyant sur la chaîne de caractères, ce qui entraîne alors des erreurs de calcul. Un tutoriel des principales fonctions utilisables dans un fichier de calcul se trouve ici : <https://frama.link/1SfDe-1L>



- 4) enregistrement des données sous le nom ANA2208acceptions.csv (le script traite les fichiers dans le format informatique ouvert CSV<sup>218</sup>) dans le répertoire où se trouve le script compareCols.csv.pl ; le codage de caractères UTF8 permet de conserver les diacritiques<sup>219</sup> ;
- 5) édition du script contenu dans le fichier compareCols.csv.pl (nous avons utilisé SciTE<sup>220</sup>) ;
- 6) copie à la ligne 6 (du script) du nom du fichier ANA2208acceptions sans l'extension csv ;
- 7) exécution du script (par la commande f5 du clavier) ; un fichier portant le même nom, mais l'extension .comp.csv, est généré, dans le cas présent, il s'appelle donc ANA2208acceptions.comp.csv ; ce fichier contient les résultats de la comparaison des langues : 1 signifie qu'un couple de mots anglais – langue romane est similaire selon les trois critères énoncés ci-dessus (cf. p.189), 0 signifie que le couple de mots n'est pas similaire.
- 8) ouverture du fichier ANA2208acceptions.comp.csv et copie des données de ce fichier dans une feuille intitulée Brut2208EN\_ROM.ods ;
- 9) copie dans cette feuille, colonnes M et N, des deux colonnes que nous avons supprimées lors de l'étape 3 (nous avons besoin de la colonne qui précise la définition de l'acceptation de façon à pouvoir exprimer pour quelle acceptation un mot anglais est similaire à un mot d'une langue romane) ;
- 10) application de la fonction =SOMME(H1:K1) dans la colonne L et extension de cette fonction à toutes les cellules de cette colonne, puis tri de la colonne L par ordre décroissant : les 1013 lignes qui contiennent le chiffre 4 sont les lignes qui contiennent les mots anglais fréquents de la santé d'au moins 4 caractères qui sont similaires avec les quatre langues romanes ;
- 11) dans une nouvelle feuille, intitulée INTERLEXIQUE1038EN\_ROM.ods, copie des 1013 lignes qui contiennent les mots similaires d'au moins 4 caractères, ajout à partir de la feuille ANAmoinsde4caract.ods des lignes des 25 séries de mots de moins de 4 caractères dont nous avons estimé l'analogie manuellement comme nous le détaillons dans la section qui suit, et enfin tri dans l'ordre alphabétique des 1038 séries de mots similaires.

---

218 CSV : comma-separated values

219 Les lettres diacritées figurent dans les fichiers, elles sont néanmoins neutralisées par le script.

220 <http://www.scintilla.org/SciTE.html>

### 5.3.5 Mesure de similarité des mots de moins de 4 lettres

#### 5.3.5.1 *Obtention des termes de moins de 4 lettres dans chaque langue*

Nous avons vu que le script ne traite pas les mots de deux ou de trois lettres (cf. 5.3.3.2, p.190). Leur comparaison donne 0 dans les fichiers en comp.csv quand bien même ils sont similaires, voire identiques. En deçà du seuil de 4 caractères (lettres), l'analogie a donc été estimée manuellement, sans l'aide de script. Le fichier ANAmoinsde4caract.ods contient ces mots. Les 25 acceptions renvoient toutes à des noms monosémiques, pour la plupart des symboles d'unités de mesure internationales (*cm, kg, etc.*), des sigles (*PC, DNA, etc.*) et des abréviations (*fig, lab, etc.*).

Tout d'abord, le nombre de caractères de chaque cellule de la feuille de calcul Brut2208EN\_ROM.ods a été calculé dans les colonnes O, P, Q, R et S : la fonction =NBCAR(C1) permet de dénombrer dans une feuille de calcul les caractères de la cellule C1. Cette fonction a été étendue à chacune des cellules de la colonne et aux colonnes des quatre langues romanes. Les données des colonnes O, P, Q, R et S ont été triées pour ne garder que les mots de deux et trois lettres, qui ont été copiés dans une feuille par langue.

Les mots de deux et trois caractères de ces feuilles ont été l'objet d'un examen systématique, langue par langue. Se mêlent dans notre démarche deux approches, une approche formelle, et une approche davantage liée à la reconnaissance visuelle que l'on peut attendre d'un lecteur.

#### 5.3.5.2 *Approche formelle*

Appliquant une approche formelle, assez proche somme toute de ce que fait automatiquement le script, nous avons considéré comme similaires les équivalents de traduction suivants :

- deux mots composés des trois mêmes premiers caractères, y compris s'ils apparaissent dans un ordre différent – (eng) *oil*, (ita) *olio*.
- deux mots dont les deux premiers caractères de la langue connue sont identiques aux deux premiers caractères de l'autre langue, et dont le troisième se trouve parmi les lettres présentes dans la suite du mot de l'autre langue, avec un maximum de 2 caractères intercalés : (eng) *use* et (fra) *usage* ;

- deux mots dont les lettres du mot de la langue connue se retrouvent dans le même ordre dans l'autre langue, avec au maximum un caractère intercalé entre chaque lettre : (fra) *bon* et (ita) *buono* ; (por) *ovo* et (ita) *uovo*.

- un mot de deux lettres de la langue connue et son équivalent si les deux lettres du premier sont au début du mot de l'autre langue : (fra) *os*, (ita) *osso* ; (por) *nó*, (ita) *nodo*.

### 5.3.5.3 *Approche complémentaire plus intuitive*

L'approche complémentaire que nous avons suivie compte sur la souplesse dont fait inconsciemment preuve un lecteur qui sait qu'en anglais, un phonème peut être rendu par plusieurs allographèmes.

Les voyelles *y* et *i* peuvent être utilisées en anglais pour écrire la diphtongue /aɪ/, comme dans (eng) *rely*, *rise*, ou bien le phonème /ɪ/, comme dans (eng) *body*, *lipid*, c'est pourquoi dans un même contexte, (eng) *day* et (por) *dia* sont peut-être perçus comme similaires par un anglophone.

De la même façon, en anglais comme en français, les consonnes *s* et *z* peuvent transcrire le son /z/ : (eng) / (fra) *cause*. Par conséquent, (eng) *gas* et (fra) *gaz* sont triés parmi les analogues. Autre exemple : les voyelles *y* et *i* sont utilisées en espagnol et en portugais pour transcrire le yod, en cela, (spa) *ley* et (por) *lei* peuvent être vus comme similaires par les lecteurs lusophones et hispanophones. Enfin, les lettres *c* et *z* sont utilisées en espagnol pour transcrire le son /θ/, considérant ce fait, (spa) *luz* et (ita) *luce* devraient paraître similaires aux yeux d'un lecteur qui connaît le castillan tel qu'il est parlé en Espagne, bien que *ce* en italien se prononce autrement (/tʃe/ ou /tʃɛ/).

D'autres phénomènes linguistiques propres à la graphie de chaque langue jouent également et parfois s'entrecroisent : pour un lusophone, (spa) *fin* est similaire parce que les mots de la famille de (por) *fim* contiennent le *n* : (por) *infinito*, *finalidade*, etc. Mais pour un hispanophone, (por) *fim*, hors contexte, pourrait à la limite être vu comme analogue à (spa) *fin* car le morphème de négation *in-* devant *p* donne lieu à des mots comme (spa) *imposible*.

Cette aptitude se situe dans la continuité de notre approche et préfigure ce que nous attendons des lecteurs qui ont comme L2 l'anglais. Bien entendu, la réalisation orale des allographèmes peut ne pas être la même dans l'une et l'autre langues, mais un lecteur ne le perçoit pas.

Dans cette approche complémentaire, nous n'avons en revanche pas classé comme analogues les sigles qui ne sont pas exactement identiques, même si ce sont les mêmes lettres qui sont employées : (fra) *ADN* n'est pas analogue à (eng) *DNA*.

En outre, si un mot de la langue connue est traduit par un mot qui existe dans la langue connue avec une autre signification, ces deux mots ne sont pas similaires : (fra) *nom* et (spa) *nombre* ne sont pas similaires parce que la forme (fra) *nombre* est un mot qui existe en français avec une signification distincte.

### 5.3.6 Similarité : trois catégories de mots traduits

Afin d'affiner le portrait des monosémiques / polysémiques au regard de leur traduction, nous avons distingué trois catégories.

- Mots monosémiques dans le corpus

Ce type de mot admet une seule acception dans le corpus si l'on examine ses collocatifs grâce au *Lexicoscope*. Ils sont 1807 dans le corpus. Il s'agit par exemple de (eng) *chronic*, *confirm*, ou encore de *neuron*.

Prenons l'exemple de *confirm*. Selon la définition de cette acception, *confirm* signifie :

*To assure the accuracy of previous statements*<sup>221</sup>

Il se traduit par les mots (spa) *confirmar*<sup>222</sup>, (fra) *confirmer*<sup>223</sup>, (ita) *confermare*<sup>224</sup>, et (por) *confirmar*<sup>225</sup>.

- Mots polysémiques à une traduction

Ce type de mot admet plusieurs acceptions mais chacune des acceptions présentes dans le corpus se traduit par le même mot dans une langue cible.

137 des 193 polysémiques sont dans ce cas, par exemple *central*, *code*, *error*, *public* ou encore le verbe *recognize*.

---

221 <https://en.wiktionary.org/wiki/confirm#Verb>

222 « Corroborar la verdad, certeza o el grado de probabilidad de algo. », <http://lema.rae.es/drae/?val=confirmar>

223 « Présenter comme certaine une chose considérée jusque là comme quelque peu douteuse ou seulement possible ou probable, établir avec plus de certitude la réalité ou la vérité d'un fait, d'une nouvelle, la valeur d'une idée, d'une théorie, le bien-fondé d'un sentiment, par de nouveaux indices, de nouveaux arguments, de nouvelles preuves ou simplement de nouvelles assurances données par une personne autorisée. », *TLFi*.

224 « Rendere sicura una cosa, dimostrarla vera », <http://www.treccani.it/vocabolario/confermare>

225 « Tornar mais certo, mais firme, mais seguro », <http://www.dicio.com.br/confirmar/>

- Mots polysémiques à plus d'une traduction

Ils représentent une très faible proportion des lemmes du corpus ScienText : 2,75 % des lemmes et tout juste 2 % des acceptions.

C'est par exemple *grow*, traduit dans son acception « *to become bigger* » par (es) *crecer*, (fra) *grandir / croître*, (ita) *crescere*, (por) *crescer*, et dans son acception « *To cause or allow something to become bigger, especially to cultivate (...)* » par (es) *cultivar*, (fra) *cultiver*, (ita) *coltivare* et (por) *cultivar*.

Au delà de ces trois catégories, nous avons ajouté les mots anglais monosémiques aux mots anglais polysémiques à une traduction parce que les mots de ces deux catégories ont une seule traduction, ce qui peut gommer la polysémie des seconds. Nous rendons compte des résultats dans la section Convergence des monosémiques et des polysémiques à une traduction, p.220.

### 5.3.7 Étapes pour obtenir quatre lexiques bilingues anglais-langue romane des mots analogues

L'association par paires des mots anglais fréquents et de leur traduction dans les langues romanes permet d'établir quatre lexiques bilingues anglais – langue romane des mots analogues dans les deux langues :

- ◆ le lexique anglais – espagnol des mots analogues pour les acceptions fréquentes des écrits scientifiques de la santé ;
- ◆ le lexique anglais – français des mots analogues etc.
- ◆ le lexique anglais – italien des mots analogues etc.
- ◆ le lexique anglais – portugais des mots analogues etc.

Ces quatre lexiques ont pour entrée les lemmes anglais fréquents dans chacune des acceptions fréquentes identifiées dans le corpus ScienText. Ils ont été élaborés en suivant les étapes suivantes :

1) copie de la la feuille Brut2208EN\_ROM.ods (cf.5.3.4, p. 192) dans quatre nouvelles feuilles, une par couple de langue, ANA1397EN-ES.ods, ANA1459EN-FR.ods, ANA1439EN-IT.ods et ANA1279EN-PT.ods ;

2) dans chacune de ces feuilles, copie du contenu de la feuille ANAmoinsde4caract ; les quatre feuilles obtenues comprennent alors les mesures de similarité de tous les mots, indépendamment de leur longueur ;

3) dans chacune de ces 4 feuilles, suppression des colonnes qui contiennent les trois langues qui ne nous intéressent pas, dit autrement, pour obtenir le lexique bilingue anglais-espagnol, nous supprimons les colonnes français, italien et portugais, etc. ;

4) tri par ordre alphabétique ; les mots de moins de 3 caractères, qui figurent en double puisqu'ils sont à la fois dans les données de la feuille Brut2208EN\_ROM.ods et dans celles de la feuille ANAmoinsde4caract.ods, sont éliminés par ce tri une ou deux fois, selon qu'ils sont ou non analogues à l'anglais.

5) suppression de la colonne H sur laquelle ne figurent logiquement que des 1 puisque 1 indique que les couples mot anglais - mot de la langue romane en question sont similaires.

L'Annexe 7 : l'interlexique : 1038 séries de mots analogues en anglais, espagnol, français, italien et portugais., p.408, contient les 1038 séries analogues. Les résultats tirés de ces feuilles sont analysés dans la section Analogie par couples de langues : quatre lexiques bilingues anglais-langue romane, p.238.

## 5.4 Conclusion du chapitre 5

Nous avons présenté dans ce chapitre la méthodologie suivie afin premièrement de traduire en espagnol, français, italien et portugais les 2000 adjectifs, noms et verbes fréquents dans ScienText pour les 2208 acceptions fréquentes que nous leur avons identifiées dans le corpus, deuxièmement de comparer les cinq formes dont le sens est équivalent. Il en a résulté un interlexique ainsi que quatre lexiques bilingues anglais – langue romane. L'interlexique est constitué de séries de cinq mots, à savoir un mot anglais fréquent et ses équivalents de traductions en espagnol, français, italien et portugais. Ces cinq mots ont une forme similaire et une acception identique, autrement dit ils sont analogues par paires mot anglais – mot d'une langue romane. Les lexiques bilingues anglais – langue romane donnent une image de l'analogie par couples de langues.

L'approche traductologique est essentiellement cibliste. Nous nous sommes appuyé sur nos connaissances dans les langues cibles et sur des descriptions. Pour trouver les équivalents de traduction que nous ne connaissions pas, nous avons suivi une démarche hypothético-déductive à partir de la morphologie latine ou grecque du mot anglais, dont nous avons supposé qu'elle suggérerait des emprunts parallèles dans les autres langues de notre étude.

Les différences lexicosémantiques et lexicosyntaxiques entre les mots anglais et leurs équivalents sont fréquentes, mais on observe aussi une convergence interlexicale des collocations de spécialité qui nous paraît intéressante dans la perspective d'une application didactique.

La concurrence de deux équivalents de traduction (ou plus) pour un même mot anglais est le résultat de la variation diatopique, diastratique, diachronique, ou encore stylistique. Partant du nombre d'acceptions et du nombre de traductions, nous avons distingué trois types de mots anglais : les mots monosémiques dans le corpus, les mots polysémiques à une traduction, les mots polysémiques à plus d'une traduction.

Enfin, les scripts qui utilisent la sous-chaine commune maximale (SCM) sont des outils dont le réglage est facilement adaptable en fonction des retours. Ces scripts gagneraient à l'avenir à prendre en compte les allographèmes par couples de langues. En effet, les résultats seraient probablement plus conformes à la capacité dont est doté l'esprit humain à reconnaître des

mots d'une langue à partir des allographèmes qui existent dans sa L1, sa L2 ou les éventuelles autres langues de son répertoire.



## Troisième partie - Résultats et perspectives

### Chapitre 6 - Analogies interlinguistiques des mots anglais fréquents de la santé du corpus ScienText et de leurs traductions en espagnol, français, italien et portugais

#### Sommaire du chapitre 6

6.1 Introduction du chapitre 6.....	201
6.2 Étude des caractéristiques du lexique anglais fréquent de la santé du corpus scientifique de ScienText.....	202
6.2.1 Représentativité des occurrences des 2 000 mots anglais les plus fréquents dans le corpus ScienText.....	202
6.2.2 Catégories grammaticales des 2000 mots anglais fréquents.....	203
6.2.3 Monosémie / polysémie des 2000 mots anglais les plus fréquents de ScienText .....	209
6.2.4 Caractérisation sémantique du lexique anglais de ScienText.....	212
6.3 Convergence des monosémiques et des polysémiques à une traduction.....	220
6.3.1 Traduction des mots polysémiques.....	220
6.3.2 Analogie des mots monosémiques et polysémiques à une traduction.....	221
6.4 Un interlexique anglais-espagnol-français-italien-portugais des mots de la santé.....	227
6.4.1 Analogie selon l'acception.....	227
6.4.2 Analogie dans les cinq langues (lexique spécifique de la santé).....	229
6.4.3 Analogie dans les cinq langues (LST).....	231
6.4.4 Comparaison de l'analogie dans les deux catégories sémantiques.....	232
6.4.5 Analogie, rang et catégorie sémantique.....	233
6.5 Analogie par couples de langues : quatre lexiques bilingues anglais-langue romane .....	238
6.5.1 Résultats globaux.....	238
6.5.2 Analogie anglais – langue romane selon la catégorie grammaticale.....	239
6.5.3 Analogie anglais-langue romane selon la catégorie sémantique.....	241
6.5.4 Améliorer les mesures de similarité par la prise en compte des allographèmes .....	242
6.6 Conclusion du chapitre 6.....	245

## 6.1 Introduction du chapitre 6

Ce chapitre rend compte de l'analyse des adjectifs, noms et verbes anglais fréquents dans le corpus de spécialité de ScienText, et de leurs équivalents de traduction possibles dans quatre langues romanes en fonction de la catégorie grammaticale des mots, leur mono- ou polysémie et leur catégorie sémantique.

Les 2208 acceptions fréquentes identifiées permettent de quantifier l'analogie au niveau sémantique, et les 1944 mots à une traduction au niveau du mot. Par conséquent, s'interroger sur l'analogie revient à se demander quelle proportion des équivalents de traduction espagnols, français, italiens et portugais des mots fréquents anglais sont potentiellement compréhensibles par un locuteur d'anglais L2 (ou L1) qui n'a pas appris ces langues, à l'échelle des acceptions fréquentes des mots fréquents ou à l'échelle des mots fréquents.

Dans la section 6.2, p.202, le lexique anglais fréquent de la santé de ScienText est étudié.

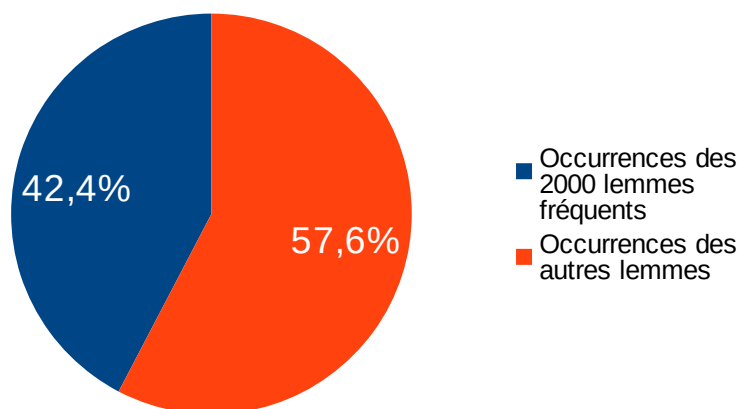
La section 6.3, p.220, envisage le nombre de traductions qu'admettent les mots : une seule traduction ou plus d'une traduction.

La section 6.4, p.227, analyse l'interlexique anglais-espagnol-français-italien-portugais et la section 6.5, p.238, les lexiques bilingues anglais – langue romane.

## 6.2 Étude des caractéristiques du lexique anglais fréquent de la santé du corpus scientifique de ScienText

### 6.2.1 Représentativité des occurrences des 2 000 mots anglais les plus fréquents dans le corpus ScienText

La version 8 du corpus d'écrits scientifiques en anglais comporte 35 244 378 tokens. Les 2 000 mots (adjectifs, noms et verbes) les plus fréquents, qui apparaissent chacun entre 1 556 fois (*internet* et *permit*) et 510 052 fois (*be*), sont représentés par 14 532 344 occurrences (cf. la feuille 2000lemm.ods). Le graphique 1 figure la proportion des occurrences des 2000 lemmes fréquents dans le corpus.



Graphique 1 : Part des occurrences des 2000 mots anglais les plus fréquents dans le corpus ScienText

Plus de deux occurrences sur cinq de mots du corpus sont des occurrences de mots anglais fréquents. Mais dans les faits, nombreuses sont les occurrences du corpus qui réfèrent à des *mots* qui sont des signes de ponctuation (1,274 millions d'occurrences) et des noms propres (3,569 millions), autrement dit à des unités qui ne font pas partie du lexique de l'anglais.

Nous avons dit dans la section Première vérification de l'étiquetage des mots à 1 000 occurrences, p.135, l'immense tâche qui consisterait à rectifier l'étiquetage de tout ce que l'étiqueteur syntaxique TreeTagger a identifié comme des mots dans le corpus. TreeTagger (Schmid, 1995) est un outil gratuit qui a été développé par l'Institut für Maschinelle Sprachverarbeitung de l'Universität Stuttgart. Les adjectifs, noms et verbes fréquents dans ce corpus de spécialité représentent au minimum 42,4 % de l'ensemble des tokens du corpus. En

conséquence, un enseignement de l'écrit en anglais de la santé peut légitimement se baser sur ces 2000 mots fréquents, à plus forte raison sur les mots les plus fréquents d'entre eux. En effet, un lecteur régulier de ce type de discours rencontre fréquemment les plus fréquents de ces mots, en fonction notamment de sa sous-spécialité - thérapeutique, biochimie clinique, etc.

## 6.2.2 Catégories grammaticales des 2000 mots anglais fréquents

### 6.2.2.1 Résultats globaux : les catégories grammaticales des 2000 mots anglais fréquents

La répartition des mots relevant de chaque catégories grammaticale donne une indication sur le fonctionnement de la langue dans le discours. Les données de la feuille 2000lemm.ods permettent de représenter la part de chacune des trois catégories grammaticales au sein de l'ensemble des 2000 mots, la part de leur nombre d'occurrences, leur nombre minimum, maximum et moyen d'occurrences. Le tableau 4 le récapitule<sup>226</sup>.

SCIENTEXT	ADJECTIFS	NOMS	VERBES	Ensemble
Nombre de mots fréquents	476	1150	374	2000
Part des mots fréquents de chaque catégorie grammaticale	23,8 %	57,5 %	18,7 %	100 %
Nombre d'occurrences des mots fréquents	2 503 603	8 806 665	3 222 076	14 532 344
Part des occ. des mots fréquents par catégories grammaticales	17,23 %	60,60 %	22,17 %	100 %
Nombre min. d'occurrences	1 559	1 556	1 556	1 556
Nombre max. d'occurrences	56 715	201 521	510 052	510 052
Nombre moyen d'occurrences	5 260	7 658	8 615	7 266

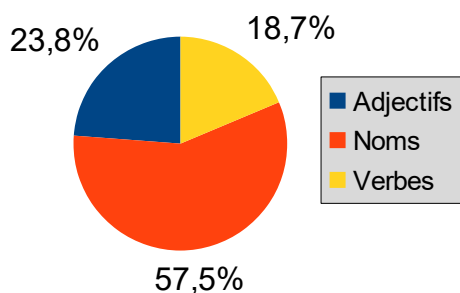
Tableau 4 : Les 2000 mots fréquents de ScienText par catégories grammaticales

<sup>226</sup> En textométrie, on utilise parfois les termes *types* et *occurrences* pour distinguer ce que nous appelons *mots* et *occurrences*.

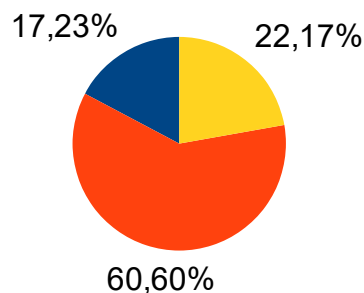
Les erreurs d'étiquetage que nous avons signalées ne nous permettent pas de mesurer la proportion des 476 adjectifs (et de leurs occurrences) au regard du nombre total de tokens du corpus étiquetés comme adjectifs (et de leurs occurrences), idem des noms et des verbes. Les chiffres que nous donnerons ne porteront donc plus désormais que sur les lemmes et les occurrences de lemmes que nous avons clairement identifiés.

Si le nombre minimal d'occurrences est quasiment identique (1 556 à 1 559), puisque nous avons sélectionné les 2000 mots qui ont le plus grand nombre d'occurrences, la disparité est considérable entre le nombre maximal d'occurrences d'adjectifs et le nombre maximal d'occurrences de verbes : 56 715 pour l'adjectif *high* et 510 052 pour le verbe *be*, 3,5 % des occurrences des mots fréquents du corpus à lui tout seul. Ce pourcentage atteste-t-il du rôle central de la description dans les écrits scientifiques de la santé ? S'explique-t-il par le recours à *be* dans des formes passives qui seraient mal étiquetées, ou encore par d'autres raisons ? Est-il dû en partie à certaines erreurs d'étiquetage ? Nous ne pouvons qu'émettre des hypothèses à ce sujet.

#### 6.2.2.2 *Part de chaque catégorie grammaticale parmi les mots anglais fréquents*



*Graphique 2 : Part des 2000 mots fréquents par catégories grammaticales*



*Graphique 3 : Part des occurrences des 2000 mots fréquents par catégories grammaticales*

Du tableau 2 sont issus les deux graphiques 2 et 3, qui permettent de comparer la part de chaque catégorie grammaticale, en nombre de mots et d'occurrences de mots. Les noms sont la catégorie la plus représentée en nombre de mots et d'occurrences, la proportion de ces

derniers atteignant les deux cinquièmes. On compte un cinquième d'adjectifs, mais leur nombre d'occurrences est moindre en proportion, ce qui signifie qu'ils occupent une place moins haute dans le rang. Au contraire, la proportion d'occurrences des verbes est plus élevée que la part occupée par cette catégorie grammaticale, ce que traduit aussi le nombre moyen d'occurrences des verbes.

### 6.2.2.3 *Part de chaque catégorie grammaticale : comparaison de ScienText avec l'Oxford English Dictionary*

L'ordre noms, adjectifs, et verbes dans le corpus ScienText est identique à celui des entrées de l'Oxford English Dictionary (OED)<sup>227</sup>, indiqué par le tableau 5 :

<b>Oxford English Dictionary</b>	ADJECTIFS	NOMS	VERBES	Ensemble
Nombre d'entrées	83 252	163 079	31 968	278 299
Part de chaque catég. grammat.	29,91 %	58,6 %	11,49 %	100 %

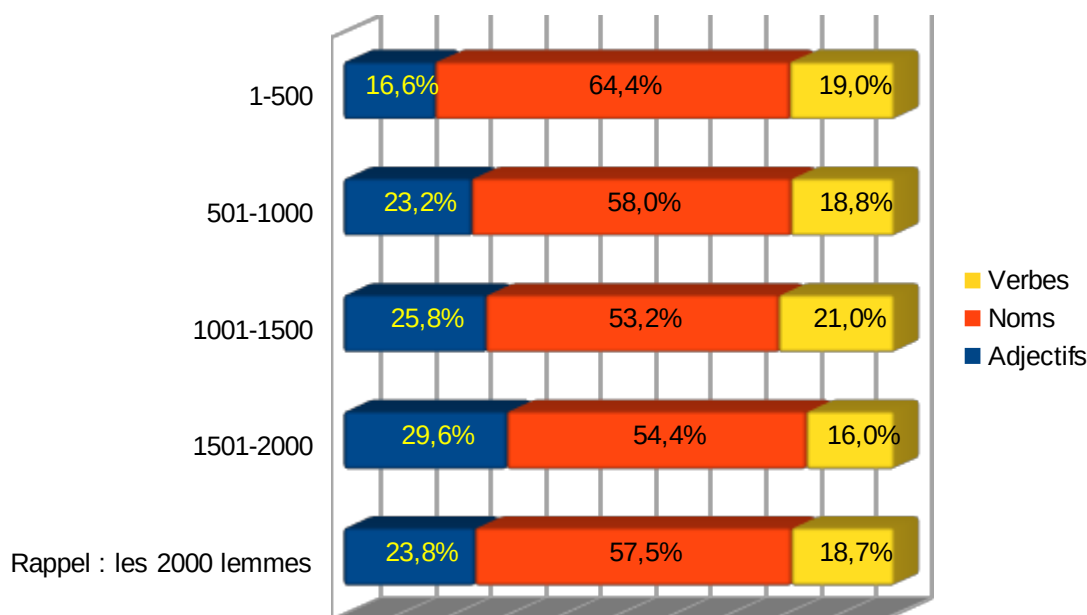
Tableau 5 : Les entrées de l'OED par catégories grammaticales

La proportion des noms dans l'OED et parmi les 2000 mots est très semblable – 58,6 % / 60,6 %. Cependant, dans le discours scientifique de la santé, la proportion d'adjectifs est moindre que celle de l'OED. Bien entendu, ces deux ensembles de mots se distinguent. D'un côté, l'OED entend définir tous les mots de l'anglais, dont certains sont sortis de l'usage. D'un autre côté, les 2000 mots de ScienText sur lesquels reposent ces tableaux ont été sélectionnés sur le critère de la fréquence, ils relèvent d'un type de discours et d'un domaine de spécialité. On remarque toutefois que la proportion de verbes fréquents est beaucoup plus élevée parmi les 2000 que la proportion d'entrée de verbes dans l'OED : 18,7 / 11,5.

### 6.2.2.4 *Catégories grammaticales en fonction du rang des 2000 mots*

Affinons ces résultats en examinant la part de chacune des trois catégories grammaticales parmi les 500 premiers mots, les 500 suivants, etc., ce qui permet d'établir le graphique 4.

227 <http://www.oed.com/>



Graphique 4 : Catégorie grammaticale des mots selon le rang des 2000 mots

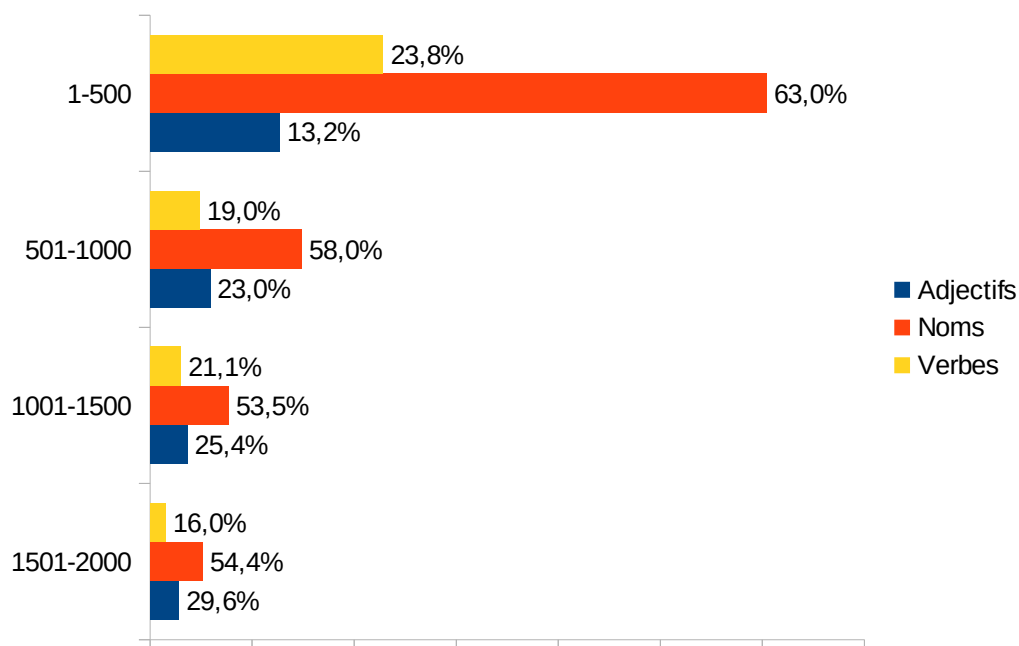
Les noms sont la catégorie grammaticale la plus représentée dans chaque quartile. On note la forte proportion des noms, environ deux mots sur trois, dans le premier quartile. Parmi ces mots les plus fréquents, les verbes viennent en deuxième position, contrairement à ce qu'on observe dans les 1500 mots suivants, où l'ordre est : noms, adjectifs, verbes. Parmi les 500 plus fréquents, la différence est cependant peu significative entre la proportion de verbes (19%), et celle d'adjectifs (16,6%). Elle est accrue dans les 1500 suivants.

La comparaison des occurrences de chaque catégorie, menée dans les sections suivantes, est probablement plus significative que la proportion de chaque catégorie de mots.

#### 6.2.2.5 *Catégories grammaticales des occurrences en fonction du rang des 2000 lemmes*

Le graphique 5 confirme la place prépondérante des 500 premiers mots : un quart des mots fréquents regroupe 9,6 millions des 14,5 millions d'occurrences, soit les 2/3 des occurrences des 2000 mots fréquents. Cette proportion indique qu'ils sont incontournables pour accéder au sens de ces textes. Pour cette raison, on peut dire que ces 500 mots les plus fréquents devraient faire l'objet de l'enseignement de l'anglais de spécialité. Les noms occupent la première place en termes d'occurrences : dans chacun des segments, ils dépassent 50 %, avec

une part accrue parmi les 500 premiers. La part des occurrences d'adjectifs croît à mesure que le rang diminue, autrement dit, moins un mot est fréquent, plus il a de chance d'être un adjectif. C'est la tendance inverse pour les verbes, mais nous avons vu la place du verbe *be* dans le premier quartile. Affinons ces résultats en examinant la distribution des 2000 mots et de leurs 14,5 millions d'occurrences.

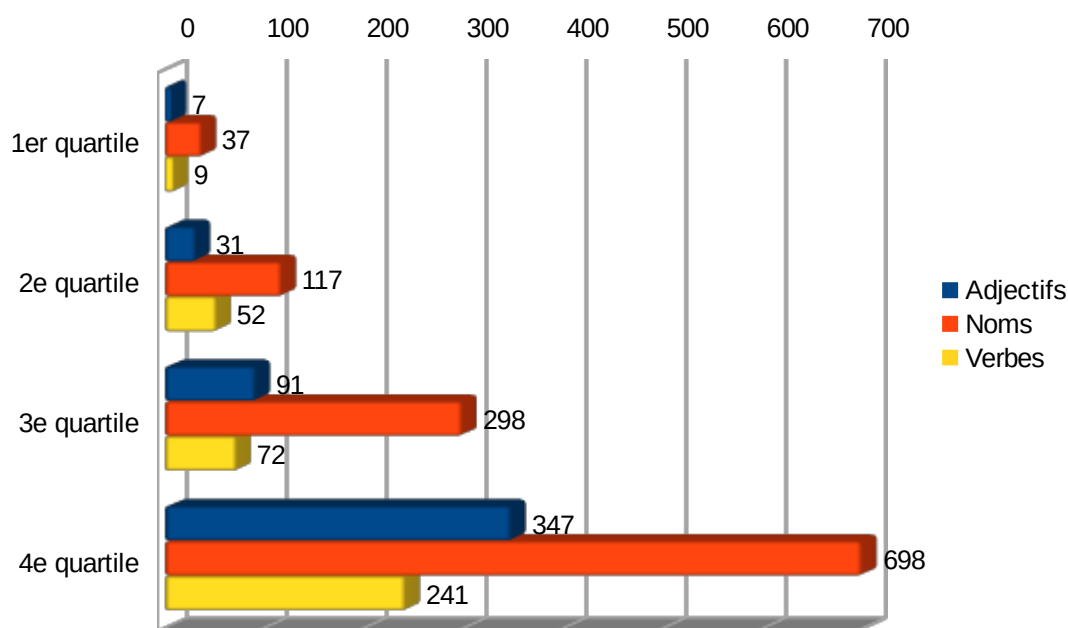


Graphique 5 : Catégorie grammaticale des occurrences selon le rang des 2000 mots

#### 6.2.2.6 Les mots les plus fréquents : catégories grammaticales des 2000 mots en fonction des quartiles d'occurrences

Le graphique 6 représente la distribution de chaque catégorie grammaticale parmi les quartiles d'occurrences. Les occurrences des deux premiers quartiles se répartissent ainsi : 15 % sont des occurrences d'adjectifs, 60,9 % de noms, 24,1 % de verbes. Les occurrences de 253 mots représentent 50 % des occurrences des 2000 mots ; on peut donc dire que ces 253 mots sont incontournables.





Graphique 6 : Catégories grammaticales par quartiles d'occurrences

Plus encore, les occurrences des 53 mots les plus fréquents du corpus ScienText, c'est-à-dire de 2,65 % des 2000 mots, représentent un quart des 14,5 millions d'occurrences. Ce très faible nombre regroupe les mots que nous avons repris dans le tableau 6. 13,2 % sont des adjectifs, 69,8 % des noms et 17 % des verbes. Un mot, *analysis*, est polysémique, il figure donc dans les deux catégories sémantiques. Environ 6/10 des mots relevant du LST.

	ADJECTIFS	NOMS	VERBES
<b>LST</b>	different high low other specific	analysis case control datum difference factor figure group level method model number percent rate result study table time type value	be base compare find have report show suggest use
<b>Lexique de spécialité de la santé</b>	clinical human	activity analysis cancer cell disease effect expression gene health mouse patient protein receptor response sample sequence treatment tumor	

Tableau 6 : Les 53 mots du premier quartile d'occurrences

### 6.2.3 Monosémie / polysémie des 2000 mots anglais les plus fréquents de ScienText

#### 6.2.3.1 Données globales sur les monosémiques / polysémiques

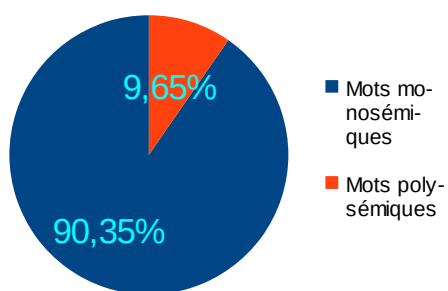
Nous avons expliqué dans la section Détermination de la monosémie / polysémie d'un mot du corpus ScienText, p. 147, comment nous avons déterminé les acceptions fréquentes valables dans le corpus. Voici dans le tableau 7 la répartition des 2000 mots selon qu'ils admettent une ou plusieurs acceptions dans le corpus ScienText. Par commodité, nous les nommons monosémiques et polysémiques.

Nous étudions la répartition des mots monosémiques / polysémiques parce que la polysémie d'un mot peut être un obstacle pour sa compréhension. Ce tableau n'est pas commenté, nous tirons de ces données les graphiques de la sous-section suivante.

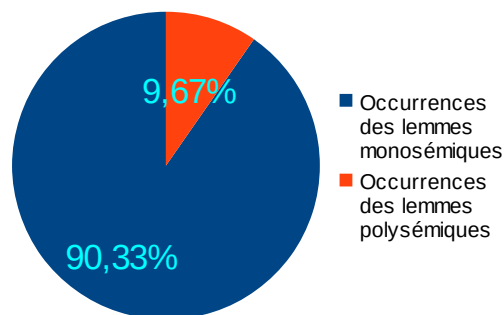
	ADJECTIFS	NOMS	VERBES	Ensemble
Nombre de mots fréquents	476	1150	374	2000
Part des mots fréq. de chaque cat. gr.	23,8 %	57,5 %	18,7 %	100 %
Nombre de monosémiques	437	1024	346	1807
Nombre de polysémiques	39	126	28	193
Part des monosém. / 2000 mots	21,85 %	51,2 %	17,3 %	90,35 %
Part des polysém. / 2000 mots	1,95 %	6,3 %	1,4 %	9,65 %
Occurrences des monosémiques	2 253 733	7 874 865	2 998 054	13 126 652
Occurrences des polysémiques	249 870	931 800	224 022	1 405 692
Part des occ. des monosém. / 2000mots	15,51 %	54,19 %	20,63 %	90,33 %
Part des occ. des polysém. / 2000mots	1,72 %	6,41 %	1,54 %	9,67 %

Tableau 7 : Monosémie / polysémie des 2000 mots fréquents de ScienText

### 6.2.3.2 Répartition globale des monosémiques / polysémiques



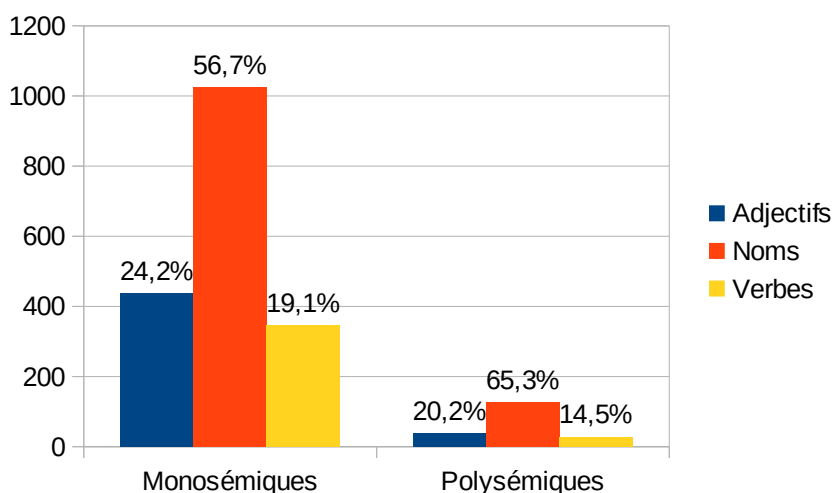
Graphique 7 : Répartition globale des monosémiques / polysémiques



Graphique 8 : Répartition globale des occurrences des monosémiques / polysémiques

Pour la très grande majorité des 2000 mots, une seule acception s'applique : 9/10 des mots sont monosémiques dans le corpus ScienText. Les occurrences respectent presque exactement les mêmes proportions. L'étude des trois catégories grammaticales donne une autre image de la répartition entre mots monosémiques et mots polysémiques.

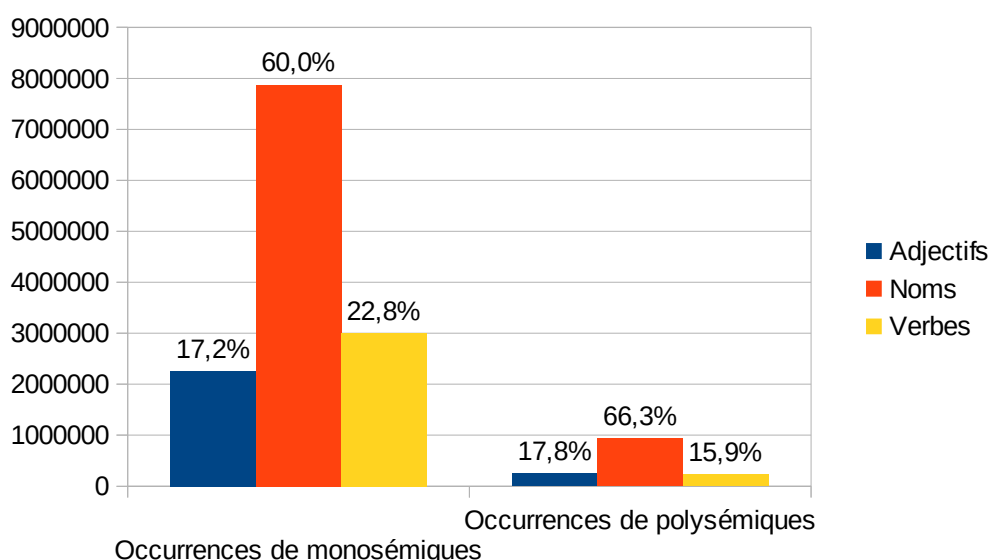
### 6.2.3.3 Répartition monosémiques / polysémiques par catégories grammaticales



Graphique 9 : Répartition des mots monosémiques / polysémiques par catégories grammaticales

Parmi les polysémiques et les monosémiques, les noms sont la catégorie grammaticale la plus représentée, puis viennent les adjectifs et les verbes. Il n'y a en cela aucune différence avec la répartition par catégories grammaticales telle que nous l'avons vue dans la section Résultats globaux : les catégories grammaticales des 2000 mots anglais fréquents, p.203. En revanche, on observe dans le graphique 9 une prédominance des noms encore plus marquée parmi les polysémiques : 2/3 des mots polysémiques sont des noms.

#### 6.2.3.4 Répartition des occurrences des mots monosémiques / polysémiques par catégories grammaticales



Graphique 10 : Occurrences des monosémiques / polysémiques par catégories grammaticales

Là encore, les noms sont plus fréquents parmi les polysémiques – ils représentent 2/3 des occurrences. La proportion d'occurrences d'adjectifs ne dépend pas de leur monosémie / polysémie. Les verbes en revanche sont moins fréquents parmi les polysémiques.

L'examen de la différence entre le nombre de mots monosémiques / polysémiques anglais et leur nombre d'occurrences ne nous permet pas de tirer d'enseignement dans la perspective d'une didactique de L3, mais nous réexaminons la portée de ces chiffres dans la section 6.3, p.220, lorsque nous additionnons les monosémiques aux polysémiques à une traduction.

## 6.2.4 Caractérisation sémantique du lexique anglais de ScienText

### 6.2.4.1 Récapitulatif des données sur les catégories sémantiques

Dans la section Deux catégories sémantiques pour les 2208 acceptions fréquentes, p.152, nous avons expliqué comment nous avons trié les 2208 acceptions fréquentes des 2000 mots anglais fréquents du corpus selon que le sens du mot anglais relevait du LST ou du lexique de spécialité de la santé. Ce tri permet de comparer la part du lexique commun aux autres textes scientifiques à la part du lexique spécifique de la santé.

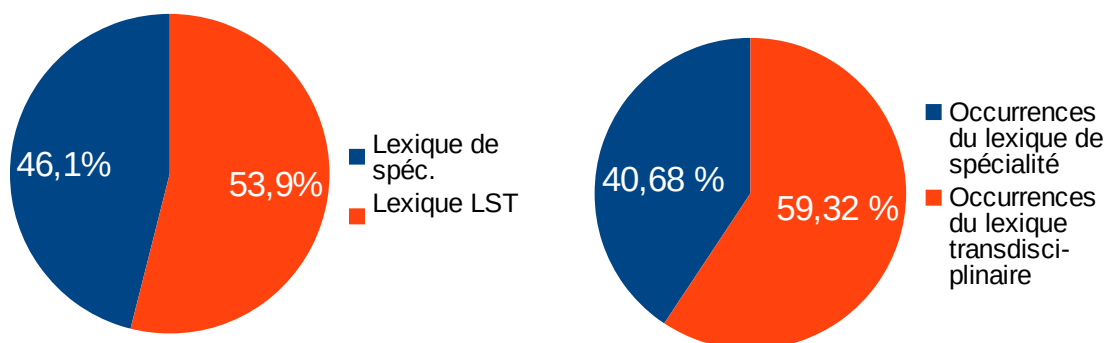
Les données du tableau récapitulatif 8 montrent que la somme des occurrences présente une légère différence en raison du mode d'estimation du nombre d'occurrences des polysémiques dont nous avons parlé dans la section 5.3.1., p.187.

	ADJECTIFS	NOMS	VERBES	Tous
Nombre d'acc. de chaque cat. gr.	522	1284	402	2208
Part d'acceptions de chaque catégorie	23,64 %	58,15 %	18,21 %	100 %
Nombre d'acc. du lexique de la santé	230	686	101	1017
Nombre d'acceptions du LST	292	598	301	1191
Part d'acceptions du lexique de la santé	22,62 %	67,45 %	9,9 %	100 %
Part d'acceptions du LST	24,52 %	50,21 %	25,27 %	100 %
Nombre d'occ. du lexique de spécialité	939467	4539284	433275	5912026
Nombre d'occ. relevant du LST	1564117	4267309	2788787	8620213
Proportion d'occ. du lexique de spéc.	15,89 %	76,78 %	7,33 %	100 %
Proportion d'occ. du LST	18,14 %	49,50 %	32,35 %	100 %

Tableau 8 : Tableau récapitulatif des catégories sémantiques

Ce tableau n'est pas commenté, nous tirons de ces données les graphiques et tableaux qui suivent.

#### 6.2.4.2 Représentation globale des catégories sémantiques des 2208 acceptions

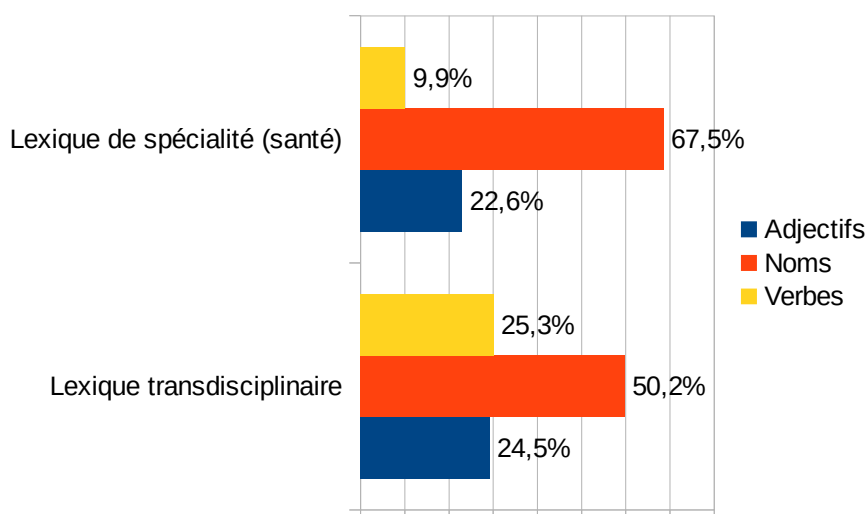


Graphique 12 : Nombre d'acceptions de chaque catégorie sémantique

Graphique 11 : Nombre d'occurrences de mots de chaque catégorie sémantique

Les deux graphiques ci-dessus permettent de comparer le nombre d'acceptions et le nombre d'occurrences selon la catégorie sémantique dont relèvent les acceptions. Dans le corpus ScienText, la part du LST est plus importante que celle du lexique de spécialité. La part d'occurrences des mots du LST est encore plus élevée, presque 3/5, ce qui peut signifier que le métalangage sur le discours de spécialité occupe davantage de place que le lexique de spécialité lui-même.

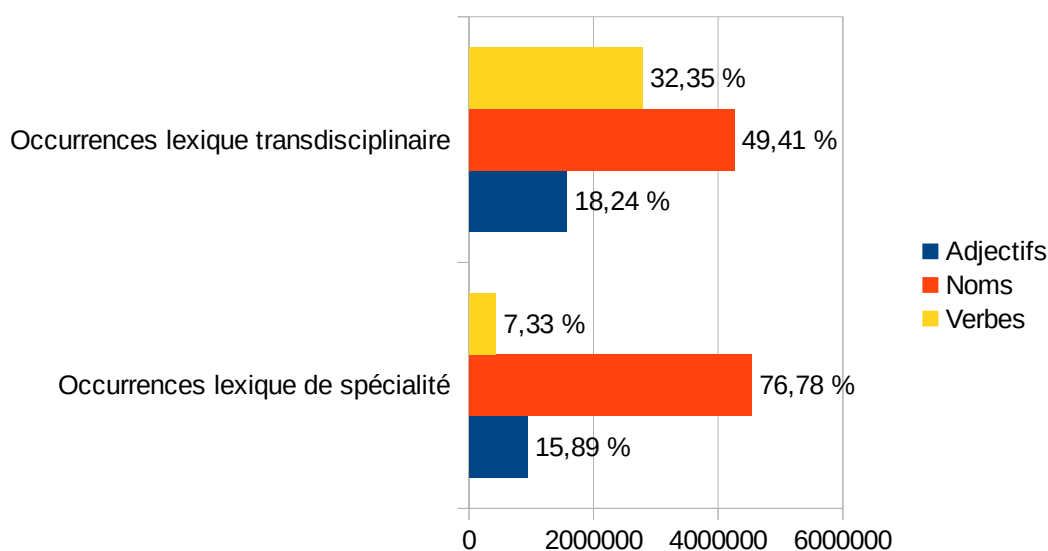
#### 6.2.4.3 Catégories grammaticales selon la catégorie sémantique des acceptions fréquentes



Graphique 13 : Catégorie grammaticale d'un mot selon la catégorie sémantique

Le graphique 13 fait apparaître que la différence dans la proportion des adjectifs dont il est fait usage n'est pas significative, au contraire de ce qu'on observe pour les verbes : ces derniers représentent 1/4 des acceptions fréquentes du LST, mais seulement 1/10 des acceptions du lexique de spécialité. Le lexique de spécialité de la santé est constitué majoritairement de noms : 2/3 des acceptions fréquentes de la santé sont exprimées par des noms.

Si l'on examine maintenant dans le graphique 14 la répartition des occurrences des termes dont l'acception se rattache au LST ou au lexique de spécialité, la compréhension des données diffère légèrement. Le nombre d'occurrences de noms dans le lexique de spécialité confirme la place des noms dans cette catégorie sémantique : plus des 3/4 des occurrences y réfèrent.



Graphique 14 : Catégorisation grammaticale de l'occurrence selon la catégorie sémantique

Les occurrences d'adjectifs n'offrent pas de contraste digne d'être étudié, à la différence des verbes, dont le nombre d'occurrences, 1/3 du LST, ne représente qu'une occurrence sur 14 dans le lexique de spécialité. Les verbes de spécialité, un quart du lexique de spécialité, sont donc peu utilisés chacun.

#### 6.2.4.4 *Catégorie sémantique et rang par acceptions*

Le tableau 9 récapitule les données que nous avons tirées par des tris successifs de la feuille 2208ac.ods. La répartition des catégories sémantiques en fonction du rang par acceptions

indique si l'on trouve plus d'acceptions de mots du lexique de spécialité ou d'acceptions du LST. En cela, elle permet d'avoir un aperçu des mots incontournables de chaque catégorie.

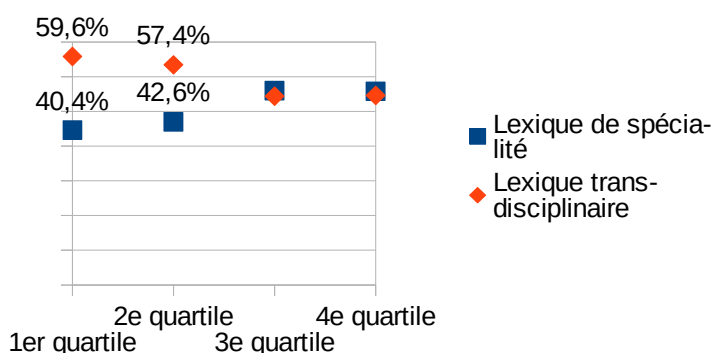
	Acceptions du lexique de spécialité	Acceptions du LST	Occurrences du lexique de spécialité	Occurrences du LST
<b>1<sup>er</sup> quartile</b>	<b>223</b>	<b>329</b>	<b>3635302</b>	<b>6007749</b>
Adjectifs	33	55	374811	872746
Noms	170	178	3058405	2968199
Verbe	20	96	202086	2166804
<b>2<sup>e</sup> quartile</b>	<b>235</b>	<b>317</b>	<b>1082916</b>	<b>1458186</b>
Adjectifs	60	73	271116	340889
Noms	154	175	709546	804519
Verbe	21	69	102254	312778
<b>3<sup>e</sup> quartile</b>	<b>280</b>	<b>272</b>	<b>737089</b>	<b>709710</b>
Adjectifs	65	86	173549	222198
Noms	182	106	477689	273192
Verbe	33	80	85851	214320
<b>4<sup>e</sup> quartile</b>	<b>279</b>	<b>273</b>	<b>456719</b>	<b>444568</b>
Adjectifs	72	78	119991	128284
Noms	180	139	293644	221399
Verbe	27	56	43084	94885

*Tableau 9 : Récapitulatif catégorie sémantique et rang*

Nous pouvons tirer de ce tableau la distribution des acceptions relevant de chaque catégorie sémantique en fonction du quartile. Les quartiles sont calculés selon le nombre d'acceptions : le premier regroupe le premier quart des 2208 acceptions fréquentes, le second les 552 suivantes, etc. D'après le graphique 15, les acceptions les plus fréquentes, celles des premier

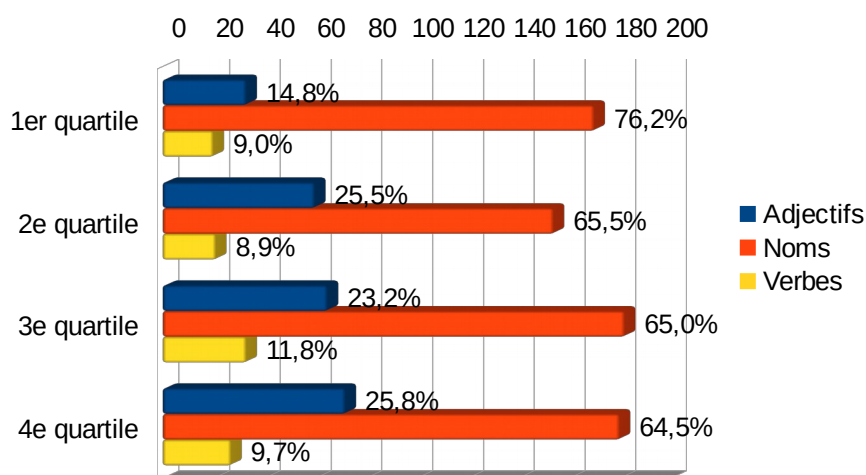


et deuxième quartiles, relèvent davantage de la catégorie du LST. On peut expliquer cela par la nécessité d'exposer les enjeux et les hypothèses, d'expliquer la méthode et de décrire les phénomènes, de présenter et d'interpréter les résultats, autant d'étapes communes à toutes les disciplines scientifiques au cours desquelles c'est le même lexique qui est utilisé. Dans les deux derniers quartiles, où les pourcentages ne sont pas affichés afin d'éviter leur superposition, on observe une répartition presque identique. Ce tableau montre l'importance du LST dans l'enseignement d'une langue étrangère : c'est ce lexique qui est le plus fréquent, et qui permet de donner du sens aux phénomènes observés et aux données recueillies.



Graphique 15 : Distribution des acceptations du lexique de spécialité / du LST

Poursuivons l'analyse avec la distribution des acceptations en fonction de la catégorie grammaticale. Nous pouvons comparer les graphiques 16 (lexique de spécialité) et 17 (LST).

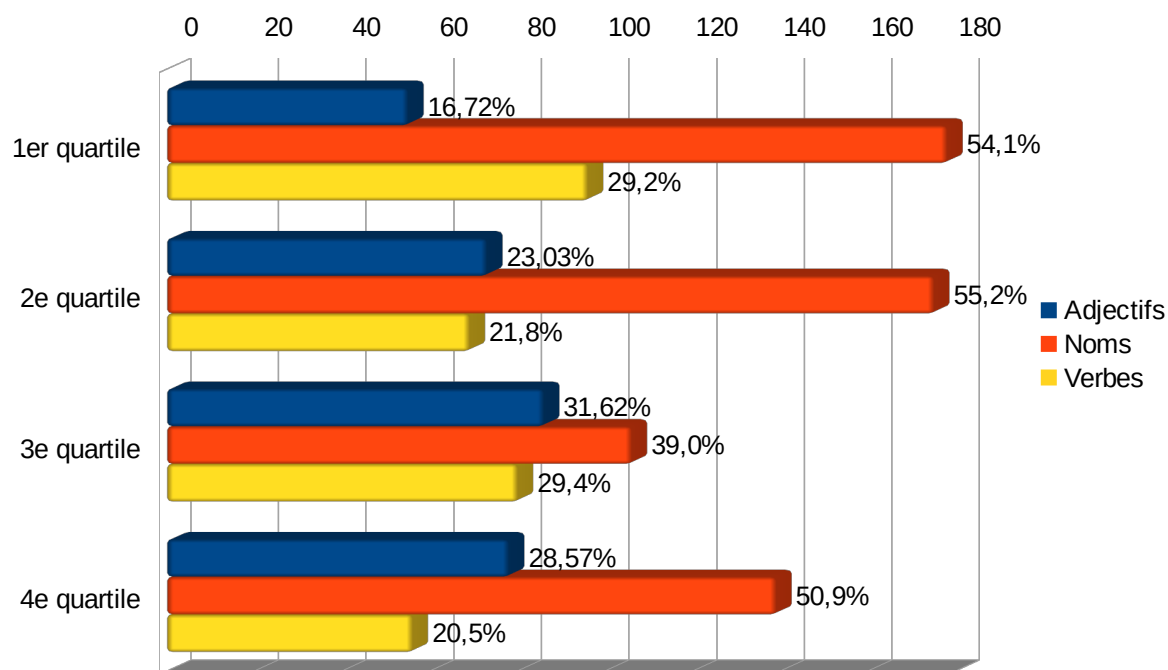


Graphique 16 : Catégories grammaticales des acceptations du lexique de spécialité par quartiles

Pour le lexique de spécialité de la santé, les noms sont la catégorie grammaticale la plus représentée dans chaque quartile, suivent les adjectifs, en enfin les verbes ; il en est de même dans le LST<sup>228</sup>, si bien que nous ne reviendrons pas sur la répartition relative de chaque catégorie pour examiner plutôt la progression de chacune.

La proportion des noms du lexique de spécialité de la santé est la plus élevée dans le premier quartile, où ils comptent pour 3/4 des mots. Dans les trois autres quartiles, elle diminue, mais les noms représentent encore presque les 2/3. Le nom est donc très clairement la catégorie grammaticale dominante dans le lexique de spécialité.

Le nombre d'adjectifs dans chaque quartile ne manque pas de surprendre, surtout dans les trois derniers quartiles, où ils avoisinent le quart des acceptions. Les verbes sont uniformément peu représentés et l'évolution de leur proportion n'est pas significative.



Graphique 17 : Catégorie grammaticale des acceptions du LST par quartiles

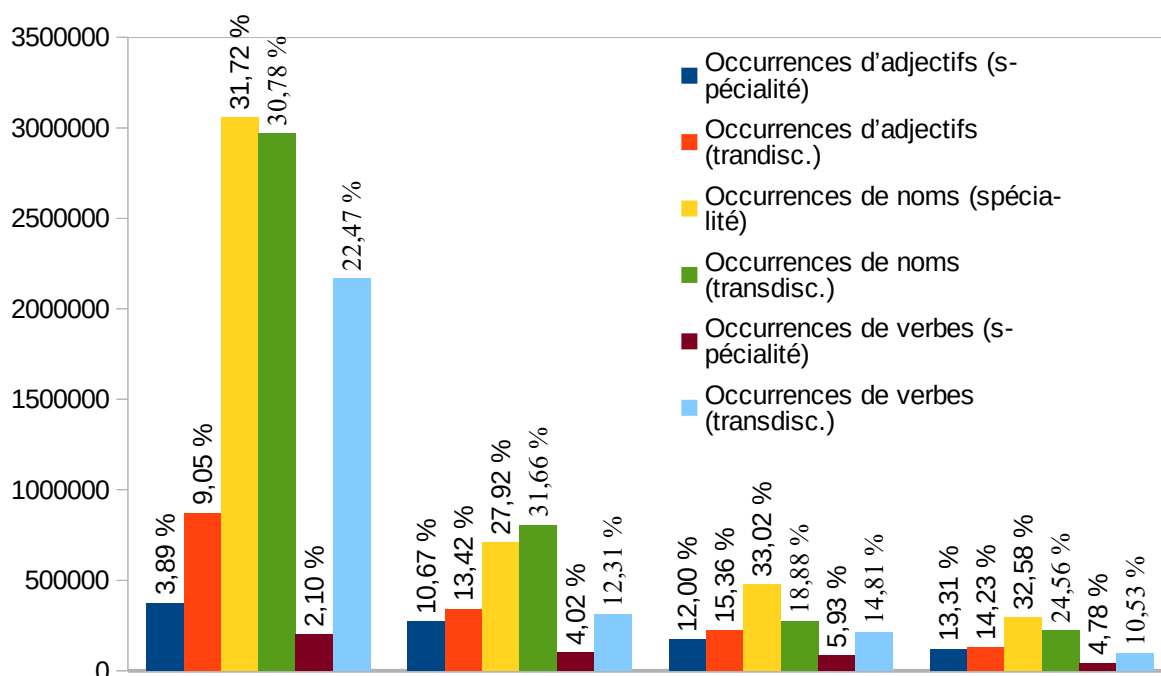
La proportion des noms du LST est la plus élevée dans les premier, deuxième et quatrième quartiles, un peu plus de la moitié des mots exprimant des acceptions fréquentes. Mais cette catégorie grammaticale est moins utilisée qu'elle ne l'est dans le lexique de spécialité de la

<sup>228</sup> Si l'on excepte le premier quartile, mais nous avons dit que les occurrences du verbe *be* représentent à elles seules 3,5 % des occurrences du corpus.

santé, ce qui confirme la place particulière des noms dans le lexique de spécialité utilisé dans le type de discours qui fait l'objet de notre étude.

Hormis dans le premier quartile, les adjectifs avoisinent ou dépassent le 1/4 des acceptions. Leur part est plus élevée parmi les acceptions moins fréquentes. La proportion de verbes, plus d'1/5 des acceptions, est deux fois plus élevée dans le LST que dans le lexique de spécialité. Le troisième quartile est à part, où la différence entre chacune des catégories grammaticales est la plus faible, sans que nous puissions l'expliquer.

La distribution des occurrences par quartile donne une image plus fidèle de l'utilisation de chaque catégorie grammaticale et de chaque catégorie sémantique. Dans la section 6.2.2.5, p.206, nous examinons la catégorie grammaticale des occurrences selon le rang des 2000 lemmes anglais. Chaque quartile correspondait donc à 500 lemmes. Nous avons vu dans les lignes et les tableaux qui précèdent les catégories grammaticales et sémantique des mots qui expriment les 2208 acceptions des 2000 lemmes. Chaque quartile correspondait donc à 552 acceptions. Le graphique 18 donne une image de la répartition des occurrences de ces mots pour les 2208 acceptions, pour chaque catégorie grammaticale et sémantique.



Graphique 18 : Distribution des occurrences de mots anglais du corpus ScienText en fonction de la catégorie grammaticale et de la catégorie sémantique

Les noms sont la catégorie grammaticale la plus représentée dans chaque quartile : si la proportion est moindre dans la deuxième moitié des données, presque la moitié des occurrences tout de même, elle atteint 63 % des occurrences du premier quartile. La proportion de noms du lexique de spécialité de la santé est à peu près constante dans chaque quartile, elle avoisine le tiers des occurrences. Dans les deux premiers quartiles, la proportion des noms du LST atteint également presque un tiers, puis elle oscille entre 1/4 et 1/5. Nous pouvons en conclure qu'en matière d'occurrences, le nom domine de la même manière le lexique de la santé et le LST, le nombre d'occurrences de noms de la santé étant de 6 % inférieur à celui des occurrences de noms du LST.

La proportion d'occurrences d'adjectifs du lexique de spécialité croît à mesure que la fréquence des adjectifs diminue : elle triple du premier quartile au quatrième. Pour ce qui est des occurrences des adjectifs du LST, on peut noter dans le premier quartile qu'elles sont deux fois plus nombreuses que celles des adjectifs de la santé. Dans les autres quartiles, la proportion d'adjectifs est semblable, à 1 à 3 points près. Que les adjectifs de sens spécifiques comptent pour moins d'un tiers des adjectifs très fréquents souligne le rôle des adjectifs très fréquents du LST pour organiser le discours scientifique.

Enfin, les occurrences de verbes du lexique de spécialité comptent pour une part très faible des occurrences, entre 2 % et 6 %, contrairement aux occurrences de verbes du LST, qui représentent en moyenne six fois plus. Cependant, la proportion d'occurrences de verbes du LST dans le premier quartile, plus d'1/5, signifie la place déterminante des mots communs à d'autres discours scientifiques.

## 6.3 Convergence des monosémiques et des polysémiques à une traduction

### 6.3.1 Traduction des mots polysémiques

#### 6.3.1.1 *Des mots polysémiques à une traduction*

Dans une perspective didactique, nous pouvons approcher les résultats sous l'angle de leur nombre de traductions. En effet, nous avons voulu déterminer les acceptions dans lesquelles les mots anglais sont analogues avec leurs traductions dans les quatre langues romanes. Ce travail devait mener à des séries de cinq mots qui ont une forme similaire et une signification identique. On pourrait estimer *a priori* qu'il est préférable d'éviter l'ambiguïté en sélectionnant des mots anglais dont le sens est univoque dans ce type de discours. Selon cette logique, seuls les mots anglais monosémiques analogues avec leurs traductions dans les quatre langues romanes auraient leur place dans ce dispositif.

Nous avons vu dans le Répartition globale des monosémiques / polysémiques, p.210, que 9/10 des 2000 mots anglais fréquents sont monosémiques. En outre, nous avons constaté que certains mots polysémiques anglais ont une seule traduction en français, espagnol, etc., indépendamment du nombre d'acceptions fréquentes dans lesquelles on les rencontre dans le corpus ScienText. Ne peut-on pas alors considérer que ces mots fonctionnent comme des mots monosémiques ? En effet, pour autant qu'un mot anglais et son équivalent dans une langue romane soient analogues, l'apprenant qui connaît le mot anglais devrait pouvoir, dans un contexte donné, reconnaître sa traduction dans une langue romane. Nous distinguons par conséquent ce que nous allons appeler les polysémiques à une traduction dans les quatre langues romanes des polysémiques à plus d'une (>1) traduction dans les quatre langues romanes.

Précisons ce que nous entendons par *à une traduction*. Dans de rares cas, pour des raisons liées à l'usage - qui n'est pas encore fixé, qui varie selon l'aire géographique etc., nous avons proposé pour certains mots anglais deux ou trois mots qui peuvent les traduire dans une langue romane. Nous avons examiné ce point dans la partie méthodologie (cf. Ordre des traductions multiples dans les tableaux, p.186). S'il est possible de traduire un mot anglais pour chaque acception par ces deux ou trois mots d'une langue romane, nous considérons qu'il

s'agit là d'une même traduction. Ainsi, *head* peut se traduire en italien par *capo* et *testa* dans les deux acceptions suivantes :

1) *head* : *The part of the body of an animal or human which contains the brain, mouth and main sense organs* ;

2) *head* : *The rounded part of a bone fitting into a depression in another bone (...)*.

Par conséquent, nous estimons qu'il n'y a qu'une traduction à *head* en italien, *capo* et *testa*, quand bien même cette traduction se compose de deux mots.

### 6.3.1.2 Mots polysémiques selon le nombre de traductions

	Ensemble	Polysémiques à 1 traduction	Polysémiques à >1 traductions
Nombre de mots	193	137	56
Proportion des mots	100 %	71 %	29 %
Nombre d'occurrences	1405692	1111956	293736
Proportion des occurrences	100 %	79,1 %	20,9 %

Tableau 10 : Répartition des deux catégories de polysémiques

Le tableau 10 rend compte de la répartition des deux catégories polysémiques.

Les mots polysémiques à une seule traduction dans les quatre langues romanes, quelle que soit leur acception, sont largement majoritaires, et encore plus majoritaires, presque 8 sur 10, sont leurs occurrences. On peut par conséquent considérer que la majorité des mots polysémiques fonctionnent à la manière de mots monosémiques. Cette convergence nous incite à examiner la proportion des 2000 mots fréquents du corpus qui n'ont qu'une traduction, qu'ils soient monosémiques ou polysémiques, afin d'estimer dans quelle mesure la polysémie des mots anglais risque d'être facteur de confusion dans un dispositif didactique qui favorise des transferts de l'anglais vers les langues romanes.

## 6.3.2 Analogie des mots monosémiques et polysémiques à une traduction

### 6.3.2.1 Un premier niveau : monosémiques / polysémiques

	Monosémiques	Polysémiques
Nombre de mots	1807	193
Nombre de mots analogues	865	87
Proportion d'analogues	47,87 %	45,08 %

Tableau 11 : Analogie des monosémiques / polysémiques

Raisonnons au niveau du mot. La proportion de séries analogues est la même dans les deux catégories, un peu moins de la moitié, avec une proportion des monosémiques supérieure de moins de 3 points. Notre hypothèse 2 est donc confirmée, mais de peu : nous croyions en effet que l'existence de plusieurs acceptions pour un mot aurait pour effet qu'il admettrait plusieurs équivalents de traduction différent. D'après le tableau 10, 56 lemmes sur 193 polysémiques admettent plus d'une traduction. La faible différence entre les deux catégories tient de plus à ce que certains équivalents de traduction ont une forme très proche, si bien que chacun d'entre eux est similaire au mot anglais : (eng) *language* et (fr) *langage / langue*. Nous pouvons affiner ces résultats à l'échelle non plus des mots mais des acceptions.

### 6.3.2.2 Analogie des monosémiques / polysémiques dans les cinq langues

	Adjectifs	Noms	Verbes	Ensemble
Rappel : les 2208 acceptions	522	1284	402	2208
Nombre de séries d'analogues	299	585	154	1038
<b>Répartition des séries d'ana.</b>	<b>57,28 %</b>	<b>45,56 %</b>	<b>38,31 %</b>	<b>47,01 %</b>
Acc. de mots polysémiques	85	260	56	401
Acc. de mots polys. ana.	51	102	17	170
<b>Proportion de polys. ana.</b>	<b>60 %</b>	<b>39,23 %</b>	<b>30,36 %</b>	<b>42,39 %</b>
Acceptions de mots monosémiques	437	1024	346	1807
Acceptions de mots monos. ana.	248	483	137	868
<b>Proportion de monos. ana.</b>	<b>56,75 %</b>	<b>47,16 %</b>	<b>39,59 %</b>	<b>48,04 %</b>

Tableau 12 : Analogie des polysémiques et des monosémiques selon la catégorie grammaticale

Les équivalents de traduction dans les quatre langues romanes des mots anglais qui n'ont qu'une acception fréquente dans le corpus ScienText sont analogues à l'anglais pour presque la moitié. Ils le sont davantage que les mots polysémiques, mais la différence n'est pas si importante au regard de la complexité que représenterait le lexique polysémique pour des apprenants : 48 % / 42 %, soit 13 % de plus (tableau 12).

Parmi les monosémiques et les polysémiques, ce sont les adjectifs, puis les noms et les verbes qui présentent la plus grande analogie, mais on note que les adjectifs polysémiques sont plus analogues que les adjectifs monosémiques, contrairement aux deux autres catégories grammaticales, ce qui signifie que les sens fréquents des mots des langues romanes recouvrent davantage ceux des mots anglais dans cette catégorie que dans les autres, pour des raisons que nous ne pouvons pas déduire sans une étude plus approfondie.

### 6.3.2.3 Analogie des 1944 mots anglais à une traduction

Rapprochons les 1807 mots monosémiques et les 137 mots polysémiques dont les deux ou trois acceptions se traduisent d'une seule manière. Nous obtenons par le tableau 13 une représentation des 1944 mots du corpus qui n'ont qu'une seule traduction en espagnol, français, italien et portugais et ceux qui ont plus d'une traduction dans au moins l'une de ces quatre langues.

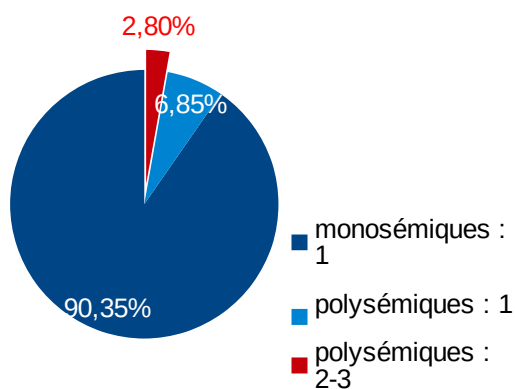
	Ensemble	mots à 1 traduction		mots à >1 traduction
		Monosémiques	Polysémiques	
Nombre de mots	2000	1807	137	56
Proportion des mots	100 %	90,35 %	6,85 %	2,8 %
Nombre d'occurrences	14532344	13126652	1111956	293736
Proportion des occurrences	100 %	90,33 %	7,65 %	2,02 %

Tableau 13 : Mots à une ou plus d'une traduction

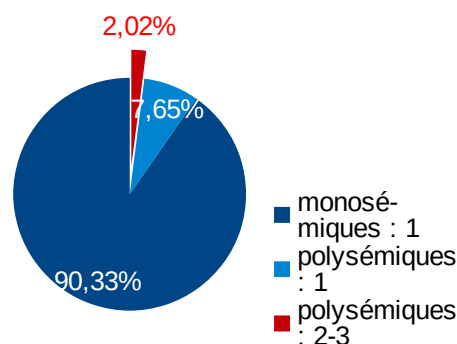
Dans les deux graphiques suivants, les monosémiques et les polysémiques à une traduction apparaissent dans deux nuances de la même couleur afin de signifier qu'ils peuvent être



considérés comme une même catégorie. Se détache en rouge la petite part constituée par les mots polysémiques à deux ou trois traductions.



*Graphique 19 : Mots anglais fréquents selon le nombre de traductions dans les cinq langues romanes*



*Graphique 20 : Occurrences des mots anglais fréquents selon le nombre de traductions dans les cinq langues romanes*

Plusieurs traductions distinctes existent pour à peine 3 % des 2000 mots, autrement dit la quasi totalité des mots n'a qu'une seule traduction, que le mot accepte une, deux ou trois acceptions.

C'est encore plus net pour les occurrences des 2000 mots : tout juste 2 % des occurrences sont celles de mots qui admettent plus d'une traduction. Nous pouvons donc estimer que la polysémie des mots anglais les plus fréquents du domaine de la santé n'est pas tant un obstacle pour les locuteurs d'espagnol, de français, d'italien et de portugais en ce que les mots anglais polysémiques sont traduits dans leur immense majorité dans les quatre langues romanes par des mots dont les acceptions recouvrent celles de l'anglais dans le corpus.

#### **6.3.2.4 Catégories grammaticales des 1944 mots à une traduction**

Considérer les adjectifs, noms et verbes des 1944 mots du corpus qui ont une traduction dans les quatre langues romanes peut permettre d'éclairer si l'une des catégories grammaticales se détache et présenterait une difficulté sémantique particulière dans les écrits scientifiques de la santé. Le tableau 14 confronte des données du tableau 7, p. 209 aux données concernant les mots à une traduction.

	ADJECTIFS	NOMS	VERBES	Ensemble
Nombre de monosémiques	437	1024	346	1807
Nombre de polysém. à une trad.	34	88	15	137
Nombre de mots fréquents	476	1150	374	2000
Part de mots à une traduction	98,95 %	96,7 %	96,52 %	97,2 %
Occurrences des monosémiques	2253733	7874865	2998054	13126652
Occ. des polysémiques à une trad.	227152	734152	150652	1111956
Nombre d'occ. des mots fréquents	2503603	8806665	3222076	14532344
Part des occ. des mots à une trad.	99,09 %	97,76 %	97,72 %	97,98 %

*Tableau 14 : Répartition des monosémiques et des polysémiques à une traduction selon la catégorie grammaticale*

Si on additionne le nombre d'adjectifs monosémiques à celui des adjectifs à une traduction et qu'on le rapporte au nombre d'adjectifs fréquents, on obtient le pourcentage de 98,95 %, le plus élevé des trois catégories grammaticales. Les chiffres équivalents obtenus pour les noms et pour les verbes accusent une faible différence, autrement dit il n'y a pas d'écart significatif entre les trois catégories grammaticales. Il en est de même des occurrences, où l'écart est encore moindre.

### **6.3.2.5 Analogie des 1944 mots à une traduction**

Nous avons expliqué dans la section Analogie des mots monosémiques et polysémiques à une traduction, p.222, pourquoi et comment nous avons additionné les mots monosémiques et les polysémiques à une traduction. 946 des 1944 mots anglais fréquents à une traduction sont analogues à leurs équivalents de traduction dans les quatre langues romanes, soit 47,3 % des 2000 mots. Par conséquent, on peut espérer qu'un apprenant d'anglais L2 (mais aussi un locuteur d'anglais L1) soit en mesure de reconnaître les traductions en espagnol, français, italien et portugais de presque la moitié des mots anglais fréquents dans le corpus ScienText, une proportion très encourageante dans la perspective d'un enseignement du FLE tel que nous l'entendons.

Le tableau 15 montre dans quelle mesure chacune des trois catégories est analogue dans les cinq langues.

	Mots fréquents	Mots analogues	Proportion de mots analogues
Adjectifs	477	271	56,81 %
Noms	1150	530	46,09 %
Verbes	373	145	38,87 %
Ensemble	2000	946	47,3 %

*Tableau 15 : Analogie des 2000 mots selon la catégorie grammaticale*

Les adjectifs sont la catégorie grammaticale dans laquelle les mots de langues romanes sont le plus analogues avec l'anglais. Plus de la moitié le sont. Les noms s'approchent les 50 %, et on observe là aussi l'originalité du rapport entre les verbes anglais et leurs équivalents de sens dans les langues romanes : moins de deux verbes anglais sur cinq sont analogues avec leurs quatre équivalents, autrement dit le verbe est la catégorie de mots dont la forme est la plus opaque.

Nous avons dans cette section 6.3 raisonné sur la base des mots. Il convient maintenant de revenir aux acceptions afin d'étudier les séries de mots qui constituent l'interlexique et les lexiques bilingues que forment l'anglais et les quatre langues romanes de notre étude (cf. Analogie par couples de langues : quatre lexiques bilingues anglais-langue romane, p.238).

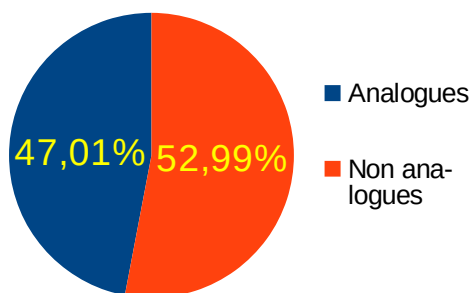
## 6.4 Un interlexique anglais-espagnol-français-italien-portugais des mots de la santé

### 6.4.1 Analogie selon l'acception

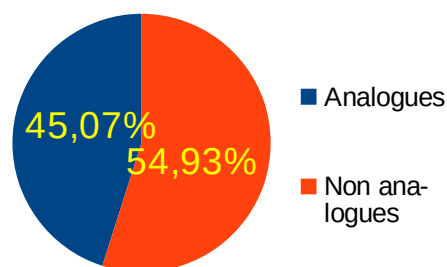
#### 6.4.1.1 Résultats globaux : pour 1038 acceptions sur 2208, les équivalents sémantiques sont analogues dans les 5 langues

Rappelons ce que nous appelons *interlexique*. L'interlexique anglais-français-espagnol-italien-portugais est un lexique en cinq langues composé des mots anglais fréquents de la santé et de ceux de leurs équivalents de traduction qui leur sont analogues dans les quatre langues romanes pour les acceptions fréquentes des mots anglais.

Les graphiques 21 et 22 synthétisent l'analogie au niveau des acceptions et des occurrences de mots qui expriment ces acceptions.



Graphique 21 : Analogie dans les 5 langues des mots exprimant les 2208 acceptions

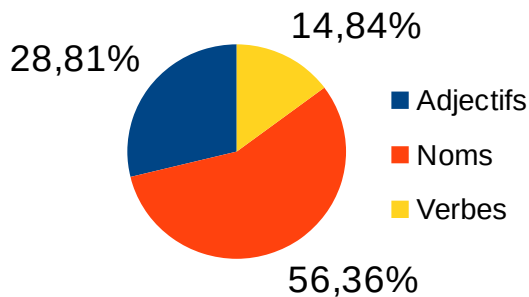


Graphique 22 : Analogie dans les 5 langues des occurrences des mots exprimant les 2208 acceptions

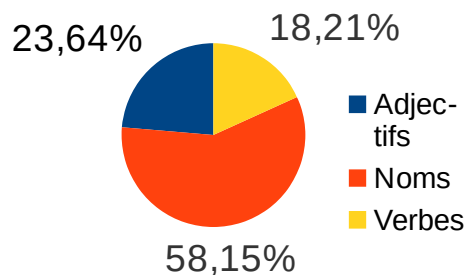
Presque la moitié des mots anglais fréquents qui expriment l'une des 2208 acceptions fréquentes du corpus sont analogues avec leurs équivalents de traduction en espagnol, français, italien et portugais. Calculée sur la base des occurrences de ces mots, la proportion de mots analogues est un peu moindre, 45 %, mais reste remarquable.

#### 6.4.1.2 Catégories grammaticales des mots analogues dans les 5 langues exprimant les 1038 acceptions

Pour le graphique 24, nous nous appuyons sur le nombre d'acceptions relevant de chaque catégorie grammaticale, et le graphique 23 repose sur les analogues par catégories grammaticales.

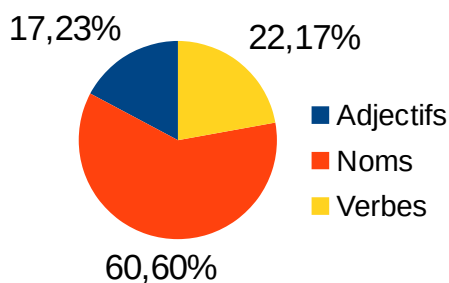


Graphique 23 : Catégories grammaticales des mots pour les 1038 acceptions où ils sont analogues dans les 5 langues

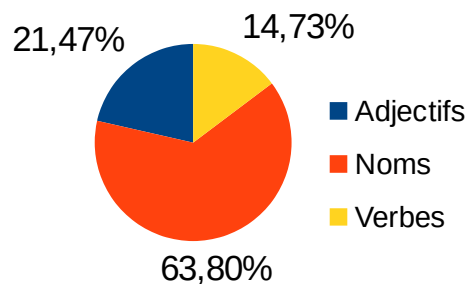


Graphique 24 : Catégories grammaticales des mots exprimant les 2208 acceptions

Si l'on compare la proportion de mots de chaque catégorie grammaticale dans le corpus (graphique 24) à celle des mots analogues dans les cinq langues (graphique 23), on remarque que les analogues sont surtout constitués de noms, puis d'adjectifs et de verbes, ce qui est identique à l'importance respective de ces trois catégories dans le corpus. Cependant, les adjectifs sont la catégorie qui contient le plus de séries de cinq mots analogues. Pour les noms, il y a peu de différence. Enfin, une plus faible proportion des verbes sont analogues.

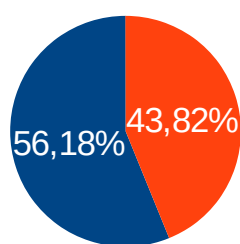


Graphique 26 : Catégories grammaticales des occurrences des mots exprimant les 2208 acceptions



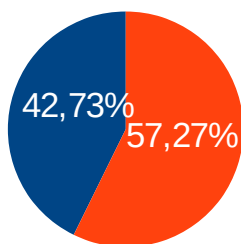
Graphique 25 : Catégorie grammaticale des occurrences des mots pour les 1038 acceptions où ils sont analogues dans les 5 langues

L'examen des occurrences (graphiques 25 et 26) confirme que les adjectifs sont analogues dans une plus grande proportion, de même que les noms. Les verbes se démarquent, comme le montrent les graphiques suivants (27, 28 et 29).



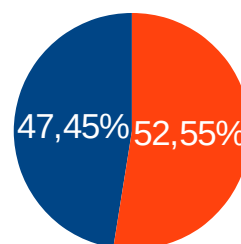
■ Occurrences d'adjectifs analogues  
 ■ Occurrences d'adjectifs non analogues

*Graphique 27 : Occurrences des adjectifs analogues / non analogues*



■ Occurrences de verbes analogues  
 ■ Occurrences de verbes non analogues

*Graphique 28 : Occurrences des verbes analogues / non analogues*



■ Occurrences de noms analogues  
 ■ Occurrences de noms non analogues

*Graphique 29 : Occurrences des noms analogues / non analogues*

Les verbes anglais n'ont pas le morphème grammatical de suffixation (par exemple *-er* en français) qui rend les mots des langues romanes plus longs et peut ainsi diminuer le rapport entre les deux chaînes de caractères comparées.

#### 6.4.2 Analogie dans les cinq langues (lexique spécifique de la santé)

Selon le tableau 16, pour plus de la moitié des acceptions de spécialité, les équivalents de traduction en espagnol, français, italien et portugais sont analogues à l'anglais. En revanche, si on distingue les catégories grammaticales, seuls les adjectifs sont analogues à plus de 50 % dans les cinq langues, un peu moins de la moitié des noms, 47 % et tout de même 44 % des verbes. C'est déjà une proportion très intéressante dans la perspective d'un enseignement plurilingue.

	Adjectifs	Noms	Verbes	Ensemble
<b>Rappel : analogie en 5 langues par catégories gram.</b>	<b>57,28 %</b>	<b>45,56 %</b>	<b>38,31 %</b>	<b>47,01 %</b>
Acceptions de spécialité	230	686	101	1017
Séries de mots analogues en 5 langues du lexique de spécialité	151	323	45	519
<b>Analogie en 5 langues du lexique de spécialité</b>	<b>65,65 %</b>	<b>47,08 %</b>	<b>44,55 %</b>	<b>51,03 %</b>

*Tableau 16 : Analogie en cinq langues du lexique anglais de spécialité de la santé*

La mesure de l’analogie des mots anglais avec leurs équivalents de traduction nous permet de comparer les systèmes, mais elle ne nous renseigne pas autant sur le fonctionnement du discours scientifique que le fait le nombre d’occurrences (tableau 17).

	Adjectifs	Noms	Verbes	Ensemble
Nombre d’occurrences de mots du lexique de spécialité	939467	4 539 284	433 275	5 912 026
Proportion d’occurrences du lexique de spécialité	15,89 %	76,78 %	7,33 %	100 %
Occurrences de mots analogues en 5 langues du lexique de spécialité	655 163	2 293 748	205 365	3 154 276
<b>Analogie en 5 langues des occurrences du lexique de spécialité</b>	<b>69,74 %</b>	<b>50,53 %</b>	<b>47,4 %</b>	<b>53,35 %</b>

*Tableau 17 : Analogie en cinq langues des occurrences du lexique anglais de spécialité de la santé*

Calculée sur le nombre d’occurrences, la similarité majoritaire du lexique anglais de spécialité avec ses équivalents de traduction se confirme : plus de la moitié des occurrences des mots anglais fréquents du corpus ScienText peuvent se traduire en espagnol, en français, en italien et en portugais avec des mots dont la forme est similaire. L’analogie dans les cinq langues est majoritaire parmi les adjectifs, où elle atteint presque 7/10, ainsi que dans les noms, un peu plus de 50 %, or nous savons que les noms sont la catégorie grammaticale la mieux représentée dans le lexique de spécialité, avec plus des 3/4 des occurrences. Pour les verbes, la proportion n’est pas loin d’atteindre 50 %. Comme seules 7 % des occurrences de cette catégorie sémantique appartiennent à cette catégorie grammaticale, une proportion analogie a peut-être moins de poids.

Il convient néanmoins de prendre avec recul cette interprétation des résultats, car les données ne sauraient refléter le fonctionnement réel des équivalents de traduction dans des textes qui appartiennent au même discours. Pour en avoir une image, un corpus comparable serait nécessaire, qui contiendrait des textes de spécialité dans chacune des langues. Nous ne serions alors plus dans une étude interlinguistique intersystémique basée sur une comparaison formelle d’équivalents de traduction, mais dans une comparaison interlinguistique intradiscursive.

### 6.4.3 Analogie dans les cinq langues (LST)

	Adjectifs	Noms	Verbes	Ensemble
<b>Rappel : analogie en 5 langues par catégories gram.</b>	<b>57,28 %</b>	<b>45,56 %</b>	<b>38,31 %</b>	<b>47,01 %</b>
Acceptions du LST	292	598	301	1191
Séries de mots analogues en 5 langues du LST	148	262	109	519
<b>Analogie en 5 langues du LST</b>	<b>50,68 %</b>	<b>43,81 %</b>	<b>36,21 %</b>	<b>43,58 %</b>

Tableau 18 : Analogie en cinq langues du lexique anglais transdisciplinaire

Le tableau 18 synthétise les chiffres clés pour le LST. Dans les trois catégories grammaticales, l'analogie dans les cinq langues est moindre dans le LST que pour l'ensemble des 2208 acceptions. La moitié des adjectifs du LST sont néanmoins analogues dans les cinq langues.

	Adjectifs	Noms	Verbes	Ensemble
Nombre d'occurrences de mots du LST	1 564 117	4 267 309	2 788 787	8620213
Proportion d'occurrences du LST	18,14 %	49,50 %	32,35 %	100 %
Occurrences de mots analogues en 5 langues du LST	751 421	1 885 309	759 218	3395948
<b>Analogie en 5 langues des occurrences du LST</b>	<b>48,04 %</b>	<b>44,18 %</b>	<b>27,22 %</b>	<b>39,39 %</b>

Tableau 19 : Analogie en cinq langues des occurrences du lexique anglais transdisciplinaire

À l'examen des occurrences des mots du LST (tableau 19), on note que les adjectifs sont là encore la catégorie qui présente la plus forte analogie dans les cinq langues, 48 %, tout en restant inférieure à la moitié, mais les adjectifs ne constituent que 18 % des occurrences. Avec une analogie de quatre points inférieure, la catégorie la plus représentée, les noms, devrait se prêter à des recherches de contextes favorables à l'enseignement des langues romanes par une entrée qui serait le LST. Les occurrences de verbes - mais nous avons vu la place dominante de *be* - sont analogues pour 1/4 seulement dans les cinq langues.



#### 6.4.4 Comparaison de l'analogie dans les deux catégories sémantiques

Nous voulions évaluer les analogies interlexicales dans un domaine de spécialité, ici, la santé, entre l'anglais et les quatre langues romanes de notre étude. Ces analogies proviennent d'origines étymologiques communes, nous l'avons vu dans la section 2.2, p.72, et probablement davantage de contacts séculaires entre savants et spécialistes des disciplines de la santé. Les données obtenues sur l'analogie interlexicale dans ce domaine peuvent orienter le travail des didacticiens des langues romanes, mais aussi des didacticiens de l'anglais qui voudraient concevoir des dispositifs destinés à des étudiants locuteurs de langues romanes. Nous pouvons tirer des chiffres obtenus de premiers enseignements sur la faisabilité d'une intercompréhension de spécialité dans le discours scientifique de la santé.

	Adjectifs	Noms	Verbes	Ensemble
Proportion d'occ. du lexique de spéc.	15,89 %	76,78 %	7,33 %	100 %
<b>Ana. en 5 langues des occurrences du lexique de spécialité</b>	<b>69,74 %</b>	<b>50,53 %</b>	<b>47,4 %</b>	<b>53,35 %</b>
Proportion d'occ. du lexique transd.	18,14 %	49,50 %	32,35 %	100 %
<b>Ana. en 5 langues des occ. du LST</b>	<b>48,04 %</b>	<b>44,18 %</b>	<b>27,22 %</b>	<b>39,39 %</b>

Tableau 20 : Analogie par catégories grammaticales selon la catégorie sémantique

L'hypothèse 3 se vérifie. En effet, on constate que dans le discours scientifique de la santé écrit en anglais, le lexique de spécialité présente plus de similarité avec l'espagnol, le français, l'italien et le portugais que le LST. Si on compare la proportion totale de mots analogues dans l'une et l'autre catégorie sémantique, on note que plus de la moitié des mots du lexique de spécialité sont analogues contre 2/5 des mots du LST. L'hypothèse se vérifie pour les trois catégories grammaticales, particulièrement les adjectifs. On note que les verbes du LST sont la catégorie dans laquelle on compte le moins de mots analogues dans les cinq langues. Le verbe des langues romanes, qui présente déjà de nombreuses formes fléchies, sont donc probablement plus difficile à comprendre sur la base d'acquis en anglais et requièrent un traitement particulier.

### 6.4.5 Analogie, rang et catégorie sémantique

Le tableau 21 permet de voir la distribution des analogues selon les quartiles des acceptions.

	Nombre de séries analogues	Part des séries ana. par quartile	Séries ana. (lexique de spécialité)	Séries ana. (LST)	SPÉ/ANA	LST/ANA
1 <sup>e</sup> quartile	267	48,37 %	117	150	43,82 %	56,18 %
2 <sup>e</sup> quartile	244	44,2 %	120	124	49,18 %	50,82 %
3 <sup>e</sup> quartile	271	49,09 %	144	127	53,14 %	46,86 %
4 <sup>e</sup> quartile	256	46,38 %	138	118	53,91 %	46,09 %
Ensemble	1038	47,01 %	519	519	50 %	50 %

Tableau 21 : Séries analogues par quartiles

Si la part d'analogues dans chaque quartile reste dans une fourchette limitée, entre 44 % et 49 %, on remarque une analogie un peu plus marquée dans les deux derniers quartiles, ce qui signifierait que les mots moins fréquents sont un peu plus analogues. La comparaison par catégories sémantiques montre que les deux catégories sont représentées par un nombre identique de séries analogues. Toutefois, on observe qu'il y a une plus forte proportion de LST que de lexique de spécialité dans les deux premiers quartiles. C'est exactement l'inverse pour le lexique de spécialité de la santé : on le trouve en plus forte proportion que le LST dans les deux quartiles de basse fréquence. Conformément aux usages concernant les probabilités conditionnelles, on notera LST/ANA (prononcé LST sachant ANA) la part du LST parmi les analogues et SPÉ/ANA (prononcé SPÉ sachant ANA) la part du lexique spécifique parmi les analogues.

Du tableau 9, il ressortait que les 552 premières acceptions, c'est-à-dire le premier quartile, sont exprimées par 9,6 millions des 14,5 millions d'occurrences, soit les deux tiers des occurrences des 2000 mots fréquents. Par conséquent, les mots anglais qui expriment ces acceptions peuvent être considérés comme incontournables pour accéder au sens des articles scientifiques anglais de la santé. Ces mots très fréquents sont-ils analogues avec leurs équivalents de traduction dans les quatre langues romanes ? Dans quelle mesure les mots analogues dépendent-ils de la catégorie sémantique de laquelle relève l'acception ? Le tableau 22 répertorie les occurrences du lexique de spécialité et celle du LST selon le rang des 2208 acceptions.

	Nombre d' occ.	Nombre d' occ. des ana.	Part des occ. des ana.	Occ. Du lexique spécialisé (analogues)	SPÉ/ANA	Occ. Du LST (analogues)	LST/ANA
1 <sup>e</sup> quart.	9643051	4288548	44,47 %	2000412	46,65 %	2288136	53,35 %
2 <sup>e</sup> quart.	2541102	1124792	44,26 %	544781	48,43 %	580011	51,57 %
3 <sup>e</sup> quart.	1446799	710968	49,14 %	378456	53,23 %	332512	46,77 %
4 <sup>e</sup> quart.	901287	425916	47,26 %	230627	54,15 %	195289	45,85 %
Ensemble	14532239	6550224	45,07 %	3154276	48,15 %	3395948	51,84 %

*Tableau 22 : Analogie du lexique de spécialité et du LST selon le rang de l'acception*

Au niveau des occurrences de mots anglais, l'analogie est légèrement moins forte qu'au niveau des mots : 47 % contre 45 %, ce qui signifie que si on mesure l'analogie en texte, elle risque d'être moindre que si on la mesure en langue. Tout cela reste bien entendu théorique, car nous raisonnons sur la seule base d'équivalents de traduction. Il convient donc d'être prudent sur l'interprétation de ces données.

Nous notons une analogie plus marquée dans les deux derniers quartiles, toutes catégories confondues, mais l'évolution des chiffres d'un quartile sur l'autre n'est pas significative. Sur la base des occurrences comme sur la base des mots, il y a une plus forte proportion de LST dans les deux premiers quartiles. Nous pouvons en déduire une tendance générale de la langue anglaise : plus une acception du LST est fréquente dans le discours scientifique, plus les équivalents de traduction de l'anglais en espagnol, français, italien et portugais semblent être similaires ; moins une acception du lexique de spécialité de la santé est fréquente dans le discours scientifique, plus les équivalents de traduction de l'anglais dans les quatre langues romanes semblent être similaires.

Nous pouvons examiner si la catégorie grammaticale influe sur la présence de séries analogues. Si nous reprenons le tableau 9 et lui ajoutons le nombre de séries analogues par catégories grammaticales et par quartiles ; nous obtenons le tableau 23 grâce auquel une comparaison de l'analogie dans le lexique de spécialité de la santé et dans le LST est possible.

Dans les trois premiers quartiles, les séries analogues dans les cinq langues dépassent la moitié des séries qui relèvent du lexique de spécialité, et dans le dernier, elles s'en approchent.

L'analogie des cinq formes est moins marquée pour le LST, surtout parmi les mots les moins fréquents. Parmi les trois catégories grammaticales, nous noterons que l'analogie des verbes très fréquents de spécialité dépasse très nettement celle des verbes exprimant des acceptions du LST, ce qui signifie que les verbes très courants de la santé pourraient servir d'entrée dans le lexique des langues romanes, et plus difficilement les verbes du LST. Nous ne commenterons ni les adjectifs, ni les noms.

	Acceptions du lex. de spécialité	Analogues du lex. de spécialité	ANA/SPÉ	Acceptions du LST	Analogues du LST	ANA/LST
<b>1<sup>er</sup> quartile</b>	<b>223</b>	<b>117</b>	<b>52,47 %</b>	<b>329</b>	<b>150</b>	<b>45,59 %</b>
Adjectifs	33	25	75,76 %	55	28	50,91 %
Noms	170	81	47,65 %	178	85	47,75 %
Verbe	20	11	55 %	96	37	38,54 %
<b>2<sup>e</sup> quartile</b>	<b>235</b>	<b>125</b>	<b>53,19 %</b>	<b>317</b>	<b>124</b>	<b>39,12 %</b>
Adjectifs	60	35	58,33 %	73	32	43,84 %
Noms	154	77	50 %	175	72	41,14 %
Verbe	21	13	61,9 %	69	20	28,99 %
<b>3<sup>e</sup> quartile</b>	<b>280</b>	<b>144</b>	<b>51,43 %</b>	<b>272</b>	<b>127</b>	<b>46,69 %</b>
Adjectifs	65	44	67,69 %	86	47	54,65 %
Noms	182	86	47,25 %	106	45	42,45 %
Verbe	33	14	42,42 %	80	35	43,75 %
<b>4<sup>e</sup> quartile</b>	<b>279</b>	<b>138</b>	<b>49,46 %</b>	<b>273</b>	<b>118</b>	<b>43,22 %</b>
Adjectifs	72	46	63,89 %	78	41	52,56 %
Noms	180	85	47,22 %	139	60	43,17 %
Verbe	27	7	25,93	56	17	30,36 %

Tableau 23 : Récapitulatif catégorie sémantique et rang

Du tableau 23, nous tirons d'abord le tableau 24 qui représente la part des séries analogues parmi les mots qui expriment les 552 acceptions de chaque quartile.

Rapportées à l'ensemble des 552 acceptions de chaque quartile, les séries analogues de chaque catégorie sémantique ne présentent pas de variation remarquable. Nous noterons toutefois moins de séries analogues qui expriment des acceptions appartenant au LST dans le dernier quartile et la tendance inverse pour ce qui concerne le lexique de spécialité de la santé, avec un quart des séries.

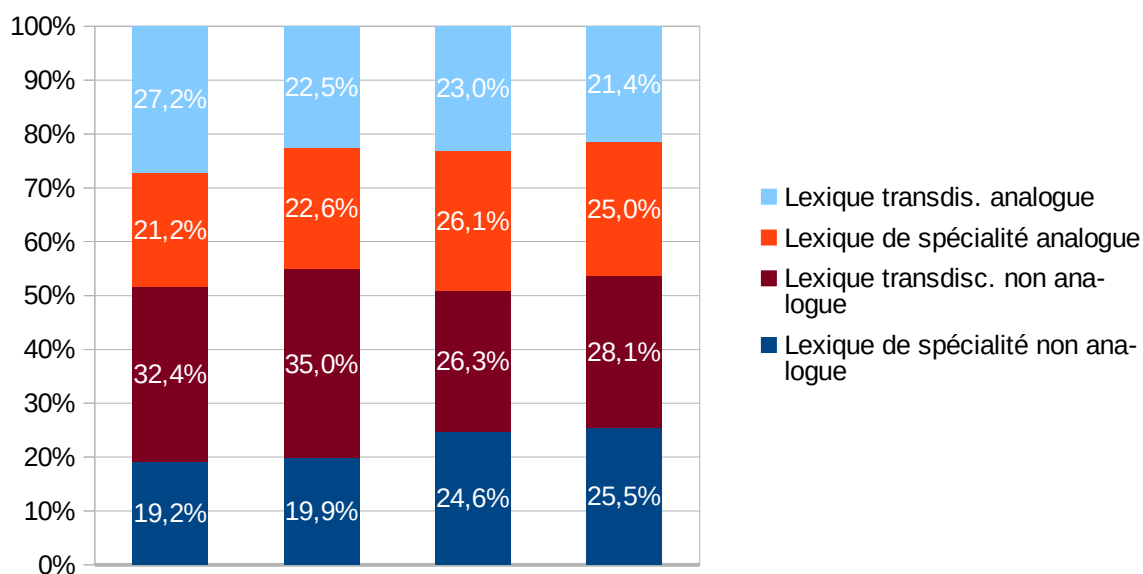


Tableau 24 : Distribution des mots analogues et non analogues du lexique de spécialité de la santé et du LST en fonction des quartiles d'acceptions

À partir du graphique 14, nous avons étudié la distribution des acceptions relevant de chaque catégorie sémantique en fonction du quartile.

Nous pouvons réexaminer ces chiffres pour comparer dans le tableau 25 les séries analogues correspondantes. Si l'on se fonde sur les occurrences du lexique de la santé, on voit que les termes équivalents seraient analogues à plus de 50 % dans chacun des quartiles. En revanche, pour les occurrences du LST, l'analogie n'atteindrait 50 % dans aucun quartile. Le lexique de la santé serait plus analogue dans le premier quartile alors que le LST le serait davantage dans les deux derniers. Les noms présenteraient quartile par quartile la moins grande différence

quant à l’analogie des deux catégories sémantiques. Pour les adjectifs et les verbes, la différence serait plus marquée et plus contrastée aussi. Nous utilisons le conditionnel pour commenter ce tableau parce qu’une telle comparaison repose sur de la pure spéculation : le contenu lexical des textes similaires dans les langues romanes est probablement bien différent ; encore une fois, des corpus comparables en donneraient une idée exacte.

	Occ. du lexique de spécialité	Occ. d'ana (lexique de spécialité)	ANA/SPÉ	Occ. du LST	Occurrences d'analogues (LST)	ANA/LST
<b>1<sup>er</sup> quartile</b>	<b>3635302</b>	<b>2000412</b>	<b>55,03 %</b>	<b>6007749</b>	<b>2288136</b>	<b>38,09 %</b>
Adjectifs	374811	299798	79,99 %	872746	406949	46,63 %
Noms	3058405	1600997	52,35 %	2968199	1334398	44,96 %
Verbe	202086	99617	49,29 %	2166804	546789	25,23 %
<b>2<sup>e</sup> quartile</b>	<b>1082916</b>	<b>544781</b>	<b>50,31 %</b>	<b>1458186</b>	<b>580011</b>	<b>39,78 %</b>
Adjectifs	271116	160036	59,03 %	340889	156902	46,03 %
Noms	709546	326879	46,07 %	804519	334760	41,61 %
Verbe	102254	57866	56,59 %	312778	88349	28,25 %
<b>3<sup>e</sup> quartile</b>	<b>737089</b>	<b>378456</b>	<b>51,34 %</b>	<b>709710</b>	<b>332512</b>	<b>46,85 %</b>
Adjectifs	173549	120037	61,17 %	222198	119542	53,8 %
Noms	477689	222323	46,54 %	273192	118562	43,4 %
Verbe	85851	36096	42,04 %	214320	94408	44,05 %
<b>4<sup>e</sup> quartile</b>	<b>456719</b>	<b>230627</b>	<b>50,5 %</b>	<b>444568</b>	<b>195289</b>	<b>43,93 %</b>
Adjectifs	119991	75292	62,75 %	128284	68028	53,0 %
Noms	293644	143549	48,89 %	221399	97589	44,08 %
Verbe	43084	11786	27,36 %	94885	29672	31,27 %

Tableau 25 : Analogie du lexique des deux catégories selon le rang des acceptions

## 6.5 Analogie par couples de langues : quatre lexiques bilingues anglais-langue romane

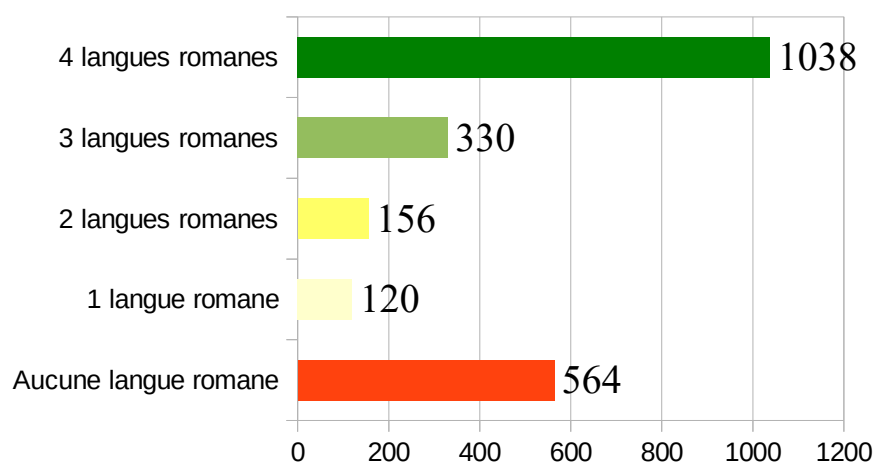
### 6.5.1 Résultats globaux

Calculé sur la base du nombre d'acceptions fréquentes du corpus, le tableau 26 récapitule combien de fois on a 1 équivalent de traduction analogue avec l'anglais, 2 équivalents analogues, etc.

Acceptions analogues avec aucune langue romane	564
Acceptions analogues avec 1 langue romane	120
Acceptions analogues avec 2 langues romanes	156
Acceptions analogues avec 3 langues romanes	330
Acceptions analogues avec 4 langues romanes	1038

Tableau 26 : Nombre d'acceptions pour lesquelles le mot anglais est analogue avec 1, 2, 3 ou 4 langues romanes.

Le graphique 30 le représente visuellement : un quart des acceptions dans le corpus ScienText ne sont pas exprimées par des mots similaires à l'anglais, mais les 3/4 le sont dans au moins une langue, et environ les deux tiers le sont dans 3 ou 4 langues romanes, ce qui est considérable et pourrait inciter les enseignants d'une langue romane à rechercher l'analogie de l'anglais avec d'autres langues romanes dans les cas d'opacité de la langue qu'ils enseignent.



Graphique 30 : Nombre de langues avec lesquelles l'anglais est analogue

Des données obtenues suite à la procédure décrite dans la section Étapes pour obtenir quatre lexiques bilingues anglais-langue romane des mots analogues, p.196, nous pouvons dresser le tableau 27 qui montre l'analogie du lexique anglais avec chacune des langues romanes de notre étude.

	Séries analogues	Pourcentage de séries analogues
anglais-espagnol	1397	63,27 %
anglais-français	1459	66,08 %
anglais-italien	1439	65,17 %
anglais-portugais	1279	57,93 %
Moyenne	1393,5	63,11 %

*Tableau 27 : Récapitulatif de l'analogie de l'anglais avec chacune des langues romanes*

Le graphique 21, p.227, montrait que pour les 2208 acceptions, 47 % des mots anglais sont similaires à leurs équivalents de traduction dans les cinq langues. Pour chaque couple de langue anglais-langue romane, l'analogie, toutes catégories grammaticales et toutes catégories sémantiques confondues, dépasse la moitié, et atteint même les 2/3 pour la paire anglais-français. Nous allons affiner ces résultats remarquables.

### **6.5.2 Analogie anglais – langue romane selon la catégorie grammaticale**

	Adjectifs	Noms	Verbes
Nombre d'acceptions dans le corpus	522	1284	402
Analogues anglais-espagnol	358	817	222
Proportion parmi les ana. anglais-espagnol	68,58 %	63,63 %	55,22 %
Analogues anglais-français	363	883	213
Proportion parmi les ana. anglais-français	69,54 %	68,77 %	52,99 %
Analogues anglais-italien	365	859	215
Proportion parmi les ana. anglais-italien	69,92 %	66,9 %	53,48 %
Analogues anglais-portugais	350	703	226
Proportion parmi les ana. anglais-portugais	67,05 %	54,75 %	56,22 %
Moyenne d'ana. anglais-langues romanes	68,77 %	63,51 %	54,48 %

*Tableau 28 : Analogie par catégories grammaticales et par langues*



Le tableau 8 ainsi que les données des analogues par paires de langues nous permettent d'obtenir le tableau 28 dans lequel est calculée la proportion d'analogues anglais-langue romane parmi les paires de mots de chaque lexique bilingue. L'examen de ces données permet de savoir si l'une des classes grammaticales semble se détacher dans l'une ou l'autre des langues romanes.

Les adjectifs et les noms présentent une analogie très semblable, autour des 2/3 dans les quatre langues, le portugais étant à part pour les noms, qui sont pour moitié analogues avec l'anglais.

Les verbes sont la catégorie dans laquelle l'analogie est le moins nette, mais la différence n'atteint que 3 points. On note toutefois que les verbes du portugais et de l'espagnol (dans leur forme infinitive) présentent une analogie plus marquée que l'italien, et surtout que le français. Cette catégorie grammaticale ne semble donc pas à privilégier comme entrée dans les langues romanes.

Le tableau 29 montre la répartition de chaque catégorie grammaticale dans les quatre lexiques bilingues.

	Adjectifs	Noms	Verbes
Nombre d'acceptions dans le corpus	522	1284	402
Taux d'acceptions dans le corpus	23,64 %	58,15 %	18,21 %
Analogues anglais-espagnol	358	817	222
Taux parmi les 1397 ana. anglais-espagnol	25,63 %	58,48 %	15,89 %
Analogues anglais-français	363	883	213
Taux parmi les 1459 ana. anglais-français	24,8 %	60,52 %	14,6 %
Analogues anglais-italien	365	859	215
Taux parmi les 1439 ana. anglais-italien	25,36 %	59,69 %	14,94 %
Analogues anglais-portugais	350	703	226
Taux parmi les 1279 ana. anglais-portugais	27,37 %	54,96 %	17,67 %
Moyenne des analogues parmi chaque catégorie gr.	25,76 %	58,52 %	15,72 %

*Tableau 29 : Répartition de chaque catégorie grammaticale dans les quatre lexiques bilingues*

L'ordre d'importance des analogues est identique dans les quatre lexiques bilingues à l'ordre d'importance de l'ensemble des acceptions : la plus grande part des analogues est constituée de noms, plus de la moitié, puis viennent les adjectifs analogues, environ un quart, enfin, les verbes sont la catégorie la moins représentée, un peu moins d'1/6. La variation des taux est faible : entre 3,4 % et 12,4 %, si bien qu'on peut estimer comparables l'analogie de l'anglais et de chacune des langues romanes de notre étude pour chaque catégorie grammaticale.

### 6.5.3 Analogie anglais-langue romane selon la catégorie sémantique

Le tableau 30 rappelle les données des tableaux 16 p.229 et 18, p.231, et s'appuie sur les quatre lexiques bilingues. Il permet de comparer la proportion d'analogues dans chaque couple de langues par catégories sémantiques.

	Acceptions du lexique de spécialité	Analogie du lex. de spécialité	Acceptions du LST	Analogie du LST	Analogie des deux catégories confondues
Paires anglais-espagnol	685	67,35 %	712	59,78 %	63,27 %
Paires anglais-français	722	70,99 %	737	61,88 %	66,08 %
Paires anglais-italien	702	69,03 %	737	61,88 %	65,17 %
Paires anglais-portugais	625	61,46 %	654	54,91 %	57,93 %
Moyenne pour les 4 paires	683,5	67,21 %	710	59,61 %	63,11 %

*Tableau 30 : Acceptions pour lesquelles les paires de langues sont analogues, par catégories sémantiques*

Le lexique de spécialité de la santé est dans chaque langue romane plus analogue à l'anglais que le LST. La moyenne de l'écart entre l'analogie dans chacune des deux catégories sémantiques se situe à 7,59 % ; dit autrement, dans les langues romanes étudiées, la probabilité qu'un mot anglais soit analogue à son équivalent de sens dépend de la catégorie sémantique à laquelle appartient son acception. Si l'on compare les lexiques bilingues, on note une plus grande différence pour les paires de mots anglais-français, avec 9,11 % d'écart.

D'autres points de comparaison des lexiques bilingues pourront faire l'objet d'études ultérieures : l'analogie par catégories grammaticales et catégories sémantiques, le rang des mots anglais analogues avec le français, les mots absents des dictionnaires généralistes, etc.

#### **6.5.4 Améliorer les mesures de similarité par la prise en compte des allographèmes**

Parmi les résultats obtenus au moyen du script, certains couples qui semblent pourtant des allolexèmes, ne sont pas identifiés comme analogues. Il en est ainsi de (eng) *ethical* et de (fra) *éthique* : pourtant, non seulement ces deux mots ont en commun les quatre premières lettres, mais au graphème *c* correspond le *qu-* en français, allographème qui existe aussi en anglais pour transcrire le phonème /k/. Le classement parmi les non analogues s'explique par la sous-chaine commune de 4 caractères, *e t h i*<sup>229</sup>, à rapporter au nombre de caractères de la sous-chaine la plus courte, 7 ici puisque les deux mots comparés comprennent 7 lettres.  $4/7$  s'élève à 0,57, soit un rapport inférieur à 0,7 qui est le minimum pour que deux mots soient rangés parmi les mots analogues (cf. Mesure de similarité par le script CompareCols.csv.pl, p.189).

D'autres couples de langues paraîtraient néanmoins formellement moins proches : les mots (eng) *ethical* et (spa) / (ita) / (por) *etico* sont analogues parce qu'ils ont comme sous-chaine commune *e t i c*, soit 4 caractères. Or la sous-chaine la plus courte est prise en compte pour mesurer l'analogie. Ici, (ita) *etico* contient 5 lettres, on a donc  $4/5$ , soit 0,8.

Nous en avons déduit de tels cas qu'il était nécessaire de modifier les réglages du script afin d'incorporer ces couples d'allolexèmes dont on peut supposer qu'ils peuvent être vus comme similaires par un lecteur, et analogues de surcroît puisque les deux termes ont le même sens. Le phénomène en jeu ici est la correspondance graphophonémique. Dans certaines langues, dont l'anglais et le français, la correspondance graphophonémique est faible : certains phonèmes peuvent être transcrits au moyen de plusieurs graphèmes, inversement, certains graphèmes transcrivent plusieurs phonèmes. Comme nous nous intéressons ici à la lecture de documents en français (espagnol, etc.), il est utile de se demander comment un lecteur réalise mentalement la prononciation de mots écrits qu'il n'a jamais entendus.

Un lecteur qui connaît l'anglais sait, consciemment ou non, selon l'enseignement plus ou moins formel qu'il a reçu, que le graphème anglais *ph-* transcrit en général le phonème /f/

---

229 Le diacritique du *é* de *éthique* est neutralisé, cf. Neutralisation des diacritiques des langues romanes, p.188.

(Treiman, 1993 : 74). Il a éventuellement aussi à l'esprit que ce graphème se prononce parfois /v/, comme dans le prénom *Stephen*. D'autre part, il sait que le phonème /f/ peut être transcrit par f (*first*), ff (*difficult*), gh (*enough*), lf (*half*) ou encore ft (*often*)<sup>230</sup>.

<b>Allographèmes</b>	<b>Exemples</b>
<b> et <bb>	(eng) <i>public</i> , (ita) <i>pubblico</i>
<cc> et <c>	(eng) <i>accumulation</i> , (spa) <i>acumulación</i>
<ch> et <qu>	(eng) <i>biochemical</i> , (por) <i>bioquímico</i>
<ck> et <c>	(eng) <i>block</i> , (fra) <i>bloc</i>
<ck> et <qu>	(eng) <i>bloc</i> , (spa) <i>bloque</i>
<c> et <qu>	(eng) <i>ethical</i> , (fra) <i>éthique</i>
<dd> et <d>	(eng) <i>addiction</i> , (spa) <i>adición</i>
<ff> et <f>	(eng) <i>affect</i> , (spa) <i>afectar</i>
<k> et <c>	(eng) <i>eukaryote</i> , (spa) et (ita) <i>eucariota</i>
<ll> et <l>	(eng) <i>allele</i> , (por) <i>alelo</i>
<mm> et <m>	(eng) <i>gamma</i> , (spa) et (por) <i>gama</i>
<nn> et <n>	(eng) <i>annotate</i> , (spa) et (por) <i>anotar</i>
<pp> et <p>	(eng) <i>apparent</i> , (spa) et (por) <i>aparente</i>
<ph> et <f>	(eng) <i>phase</i> , (spa), (ita) et (por) <i>fase</i>
<qu> et <c>	(eng) <i>qualitative</i> , (spa) <i>cualitativo</i>
<ss> et <s>	(eng) <i>depression</i> , (spa) <i>depresión</i>
<s> et <ss>	(eng) <i>parasite</i> ; (ita) <i>parassita</i>
<tt> et <t>	(eng) <i>attitude</i> , (por) <i>atitude</i>
<y> et <i>	(eng) <i>catalytic</i> , (ita) <i>catalitico</i>

Tableau 31 : Allographèmes anglais - langues romanes

230 Allographèmes définis sur la page <https://frama.link/NJwNmxxq> comme des « graphèmes ayant les mêmes fonctions d'Unité minimale distinctive »

Ces connaissances sont plus ou moins conscientisées, plus ou moins exhaustives, mais on peut imaginer qu'elles peuvent aider à faire coïncider une forme lue et une forme intégrée dans la mémoire. Nous pouvons estimer à l'instar de Caure (2009 : 103) que « vu la place qu'occupe le système graphique dans notre société, un lecteur a *forcément* le phonème correspondant à tel graphème à l'esprit pendant sa lecture », et qu'un lecteur qui s'initie aux langues romanes est prêt à voir comme allographèmes les lettres et digrammes du tableau 31, parce qu'ils existent en anglais.

La prise en compte des allographèmes bénéficierait probablement davantage aux couples de langues anglais-espagnol, anglais-portugais et anglais-italien. En effet, le français et l'anglais sont plus proches par leur graphie.

## 6.6 Conclusion du chapitre 6

L'analyse des données permet de caractériser d'une part le lexique anglais, d'autre part ses équivalents de traduction en espagnol, français, italien et portugais, aussi les résultats présentent-ils un triple intérêt : pour la connaissance de la langue anglaise, pour l'enseignement de cette langue et pour l'enseignement des langues romanes sur la base d'analogies lexicales. Nous allons récapituler les principaux enseignements en croisant ces trois niveaux.

Les 500 mots les plus fréquents du corpus ScienText représentent les 2/3 des occurrences des 2000 mots fréquents. Plus importants encore, 253 mots rassemblent la moitié des occurrences, ce qui plaide en faveur d'un enseignement de ces mots très fréquents en classe d'anglais de la santé. Enfin, la majorité des acceptions et des occurrences des 2000 mots fréquents se rattachent au lexique scientifique transdisciplinaire.

Selon les critères appliqués au moyen du profil combinatoire, les 2000 mots anglais fréquents de ScienText sont monosémiques à 90 %, et la part de mots qui n'admettent qu'une traduction en espagnol, français, italien et portugais est encore plus élevée, 97 %, ce qui signifie que les deux ou trois acceptions fréquentes des mots anglais polysémiques recouvrent celles de leurs équivalents de traduction. On peut alors raisonner en nombre de mots anglais dont les traductions sont analogues dans les quatre langues romanes : 47 % sont dans ce cas, particulièrement les adjectifs et les noms. La forme des verbes est moins similaire à celle de leurs équivalents de traduction.

Le nom est la catégorie grammaticale qui domine le lexique de la santé et le LST, mais c'est parmi les adjectifs que l'analogie est la plus forte. Au total, la proportion de mots anglais analogues dans les quatre langues romanes pour une même acception, également 47 %, témoigne de bonnes bases de transfert entre l'anglais et les langues romanes. C'est au sein du lexique de la santé que l'on trouve le plus de séries analogues dans les cinq langues, plus de la moitié, particulièrement les adjectifs et les noms.

La comparaison de la forme des mots anglais fréquents avec celle de leurs équivalents dans les quatre langues romanes a abouti à des lexiques bilingues : le lexique anglais-espagnol des mots espagnols analogues à l'anglais, le lexique anglais-français des mots français analogues à l'anglais, etc. L'analogie va de 58 % (anglais-portugais) à 66 % (anglais-français). En conséquence, un enseignant qui enseigne une langue romane dans le domaine de la santé peut

puiser dans l'un ou l'autre de ces lexiques bilingues des mots qui devraient être facilement mémorisés par les apprenants (de L1 anglais et, nous l'espérons, de L2 anglais) du fait de leur analogie avec l'anglais. De plus, on peut en comparant ces lexiques bilingues repérer quels mots anglais ne sont pas analogues avec l'une des langues romanes mais le sont avec une autre. Nous y voyons une potentialité didactique : faire comprendre des mots opaques du français avec des mots d'une autre langue romane qui sont analogues avec l'anglais.

# Chapitre 7 - Perspectives : vers un dispositif didactique plurilingue d'enseignement du FOS sur la base d'analogies interlinguistiques

## Sommaire du chapitre 7

7.1 Introduction du chapitre 7.....	247
7.2 Articulation de l'étude lexicologique et d'activités de FLE construites sur des analogies avec l'anglais.....	249
7.2.1 Des analogies lexicales anglais – langue romane pour l'apprentissage d'une L3 romane sur objectifs spécifiques.....	249
7.2.2 Configuration : L1 langue distante, L2 anglais L3 français.....	250
7.2.3 Quels mots de l'interlexique et des lexiques bilingues incorporer ?.....	250
7.3 Créer des conditions favorables à la perception d'analogies interlinguistiques.....	252
7.3.1 Lire une langue sans l'avoir apprise.....	252
7.3.2 Application de principes de la didactique des langues tierces.....	253
7.3.3 S'appuyer sur le lexique analogue.....	253
7.3.4 Valorisation d'analogies et de dissemblances syntaxiques et morphosyntaxiques.....	255
7.4 Requêtes dans des corpus multialignés et exemples d'activités.....	257
7.4.1 Des textes authentiques contenant des mots de l'interlexique.....	257
7.4.2 Présentation des corpus multialignés.....	257
7.4.3 Extractions de chaînes de caractères.....	259
7.4.4 Résultats partiels d'extractions dans les deux corpus multialignés.....	260
7.5 Conclusion du chapitre 7.....	267

## 7.1 Introduction du chapitre 7

L'étude lexicologique dont nous avons livré des résultats dans le chapitre précédent répond à un objectif premier : faciliter l'accès au français de la santé à des apprenants qui ont au moins un niveau B2 en anglais (cf. l'introduction, p.19). La plus grande partie de cette recherche a été consacrée à la délimitation du lexique anglais fréquent et ses acceptions fréquentes dans le corpus ScienText, et à sa traduction en espagnol, français, italien et portugais. Les mesures de similarité interlinguistique montrent qu'il existe de bonnes bases de transferts lexicaux entre l'anglais et ces langues, et plus encore entre l'anglais et le français. Une didactique du



français sur objectifs spécifiques inspirée de l'intercompréhension devrait donc pouvoir s'appuyer sur des éléments de ce lexique analogue et sur des analogies syntaxiques que nous avons vues section Analogies et dissemblances syntaxiques de l'anglais et des langues romanes, p.54.

Nous présentons dans ce chapitre comment nous pourrions passer d'une étude prospective du lexique à un dispositif didactique plurilingue d'apprentissage du FLE appliqué au domaine de la santé et inspiré des approches de l'intercompréhension. Nous entamons une réflexion sur l'élaboration d'activités où la comparaison interlinguistique est un ressort et l'extraction de concordances dans des corpus multilingues alignés un moyen.

Dans la section 7.2, p.249, nous explicitons le lien que nous voyons entre l'interlexique et les lexiques bilingues élaborés pendant cette thèse et la conception à venir d'activités destinées à un public de professionnels de la santé qui souhaitent apprendre le FLE en L3.

La section 7.3, p.252 explique comment nous comptons créer des conditions favorables à la découverte du FLE L3 après l'anglais L2.

La section 7.4, p.257, rend compte des requêtes que nous avons menées dans deux corpus alignés : EMEA et Europarl. Les résultats partiels indiquent des concordances qui pourraient être didactisées et quelques pistes possibles.

## **7.2 Articulation de l'étude lexicologique et d'activités de FLE construites sur des analogies avec l'anglais**

### **7.2.1 Des analogies lexicales anglais – langue romane pour l'apprentissage d'une L3 romane sur objectifs spécifiques**

Nous avons déterminé les 2000 adjectifs, noms et verbes anglais fréquents dans le corpus scientifique anglais de ScienText (cf. Constitution d'une liste des adjectifs, noms et verbes anglais fréquents du corpus ScienText, p.135) et 2208 acceptions dans lesquelles on les rencontre fréquemment dans le corpus (cf. Détermination du sens des 2000 mots anglais fréquents, p.143) ; on peut considérer que les plus fréquents d'entre eux sont essentiels pour accéder au sens de ce type d'écrits. Nous supposons qu'un lecteur qui a appris l'anglais en L2 et dont la santé est le domaine de spécialité connaît beaucoup d'entre eux. Toutefois, nous ne nous hasarderons pas à en proposer une estimation dans la mesure où les pratiques d'enseignement de l'anglais de spécialité varient d'un pays à l'autre, et où les acquis dans une langue étrangère varient évidemment d'un individu à l'autre.

Nous avons comparé la forme des 2000 mots anglais à celle de leurs équivalents de traduction en espagnol, français, italien et portugais pour les 2208 acceptions (cf. Mesures de similarité : vers un interlexique et quatre lexiques bilingues, p.187), et en avons tiré un interlexique de 1038 séries de mots analogues à l'anglais (cf. Résultats globaux : pour 1038 acceptions sur 2208, les équivalents sémantiques sont analogues dans les 5 langues, p.227) ainsi que quatre lexiques bilingues anglais – langue romane des mots analogues (cf. Analogie par couples de langues : quatre lexiques bilingues anglais-langue romane, p.238). L'interlexique et les lexiques bilingues sont une base de travail en vue d'élaborer des ressources didactiques qui permettent à des apprenants de L3 romane de construire leur apprentissage sur les analogies entre l'anglais L2 et une L3 romane. La lisibilité du français devrait être possible grâce aux analogies interlinguistiques qui forment des bases de transfert (Meißner, 1998 et 2000) vers des langues romanes. Dans les activités que nous envisageons, la L3 romane est le français, et nous nous situons dans une configuration particulière, où la L1 est une langue distante.

## **7.2.2 Configuration : L1 langue distante, L2 anglais L3 français**

Un anglophone de naissance qui apprend une langue romane constate immanquablement des analogies à l'écrit entre sa L1 et la L2 (ou L3) romane qu'il entreprend d'apprendre. Nous nous plaçons dans une autre perspective, celle d'un locuteur d'une langue distante, c'est-à-dire typologiquement éloignée des langues romanes et des langues germaniques, et qui a été initié à l'anglais, comme c'est le cas de centaines de millions d'apprenants qui vivent en Asie. Considérant le potentiel du eTandem pour l'enseignement-apprentissage des langues et cultures françaises et chinoise, Wang Szilas (2016 : 12, note 12) souligne la difficulté que représente la distance linguistique : «la langue chinoise, comparée avec le français, est considérée comme une langue particulièrement distante géographiquement, culturellement, linguistiquement, graphiquement (Bellassen, 2011) ».

Le public cible se compose de professionnels ou futurs professionnels des métiers de la santé asiatiques ou africains qui ont un projet de mobilité professionnelle dans un pays francophone, et qui ont pour cette raison besoin d'apprendre le français. Il ne sait pas nécessairement que les locuteurs d'anglais et de français sont en contact depuis plus de mille ans, et que ces deux langues présentent pour cette raison de nombreuses analogies linguistiques. Cependant, cela importe peu dans la mesure où les mots et structures syntaxiques sur lesquels reposeront les activités d'apprentissage du FOS sont analogues en synchronie : ses connaissances dans le domaine de spécialité et ses compétences en anglais devraient donc permettre à l'apprenant de comprendre des textes sélectionnés. On ne peut pas non plus espérer que l'apprenant soit capable de décrire en anglais la langue anglaise. Aussi faut-il prévoir un questionnement métalinguistique sans trop de jargon. L'enjeu est de créer des conditions dans lesquelles un apprenant de L3 français puisse reconnaître grâce à un étayage les mots fréquents de sa L2 anglais dans des mots analogues en français.

## **7.2.3 Quels mots de l'interlexique et des lexiques bilingues incorporer ?**

### **7.2.3.1 LST vs. lexique de spécialité de la santé**

Comme nous l'avons expliqué (cf. Deux catégories sémantiques pour les 2208 acceptions fréquentes, p.152), nous avons trié les acceptions dans deux catégories sémantiques. Distinguer des catégories avait du sens pour éprouver notre hypothèse 3 sur une plus grande analogie du lexique de spécialité que du LST. Comparer l'analogie au sein du lexique

spécifique de la santé et du LST permet de faire avancer la connaissance et prépare un terrain pour de futures études synchroniques et diachroniques.

Pour ce qui est d'activités de FLE destinées à un public spécialiste de la santé, devrait-on privilégier des textes qui contiennent en priorité des éléments du lexique de spécialité ? Il n'est peut-être pas pertinent dans un premier temps de distinguer les deux catégories, étant donné que l'objectif est d'abord de mesurer si une approche du FOS inspirée de l'intercompréhension porterait des fruits. Des études ultérieures pourront approfondir ce point. Toutefois, tous les apprenants ne connaissent probablement pas certains termes de la santé très spécialisés - nous pensons à des termes de la génomique et de la biologie interventionnelle. On peut donc juger préférable de ne pas incorporer dans un dispositif d'enseignement de contextes qui les contiennent : inconnus en anglais, ils ne seraient pas reconnus en français. Le LST, en revanche, rendrait certaines activités accessibles à des apprenants d'autres spécialités que la santé.

### ***7.2.3.2 Incorporer le lexique polysémique de la langue passerelle ?***

On peut se demander si les mots qui sont polysémiques en anglais risquent de poser problème aux apprenants si on utilise l'anglais comme langue passerelle vers le français (les langues romanes en général). Dans les chapitres 4, p.130) et 5, p.160), consacrés à la méthodologie, nous avons abordé la question de la polysémie et nous avons distingué trois catégories de mots traduits : les mots anglais monosémiques dans le corpus, les mots anglais polysémiques à une traduction et les mots anglais polysémiques à plus d'une traduction (cf. Similarité : trois catégories de mots traduits, p.195).

Hors contexte, un apprenant qui rencontre un mot polysémique peut être troublé parce qu'il ne sait pas en face de quelle acception il se trouve. Mais l'apprentissage est contextualisé dans le domaine d'activité de l'apprenant et de surcroît, 97,2 % des lemmes anglais fréquents ont une seule et même traduction dans les quatre langues romanes, qu'ils admettent une, deux ou bien trois acceptions fréquentes dans le corpus ; par conséquent, nous en avons déduit que les risques de confusion des apprenants de français L3 liés à la polysémie de l'anglais ne semblent pas tellement à craindre.

## 7.3 Créer des conditions favorables à la perception d'analogies interlinguistiques

### 7.3.1 Lire une langue sans l'avoir apprise

Les activités pédagogiques d'enseignement du FLE que nous souhaiterions développer s'inspirent de l'idée que quelqu'un peut lire une langue sans l'avoir apprise, s'il est soutenu par un étayage métalinguistique. Dès 1992, Havran *et al.* (1992) ont montré comment un locuteur d'allemand décode un mot international du français grâce à son analogie avec un mot allemand. Nous avons la conviction qu'un apprenant de FLE qui connaît l'anglais peut comprendre certains passages écrits dès le début de son apprentissage si ces passages sont contextualisés dans son domaine de spécialité<sup>231</sup>.

Selon Meißner (2016 : 13), « moins il y a de bases de transfert (= formes et fonctions partagées) entre langue-pont et langues cibles, plus un guidage pédagogique en conséquence se révèle indispensable. ». Aussi cherchons-nous à mener l'apprenant du connu en anglais à l'inconnu en français par :

- la mise en pratique de principes de la didactique des langues tierces et de l'intercompréhension énoncés par Caddéo et Jamet (cf. section Des approches novatrices qui développent l'intercompréhension, p.113) ;
- des éléments de l'interlexique ;
- des analogies syntaxiques et morphosyntaxiques (cf. section Analogies et dissemblances syntaxiques de l'anglais et des langues romanes, p.54) ;

Certaines compétences de compréhension écrite décrites dans le *Référentiel de Compétences de communication plurilingue en intercompréhension (REFIC, De Carlo et al., 2015)* sont visées : s'appuyer sur le lexique aisément identifiable, savoir faire des hypothèses sur le sens des mots à partir des contextes sémantiques et discursifs, savoir exploiter la comparaison entre plusieurs langues à l'écrit. Nous allons reprendre les trois leviers dans les sections qui suivent.

---

231 Le contexte, combiné avec le remarquable outil qu'est l'*Éventail des langues* (programme multilatéral Grundtvig 2012 Take Care, <http://takecareproject.eu>, Dinu *et al.*, 2016), aide les praticiens à trouver avec les patients migrants allophones les mots dont ils ont besoin pour échanger en consultation ou en pharmacie.

### **7.3.2 Application de principes de la didactique des langues tierces**

Nous appliquons certains principes mis en avant par les approches de l'intercompréhension (Caddéo & Jamet, 2013) : stimuler la perception de la proximité, et encourager les inférences. Plus généralement, nous appliquons les principes de la didactique des langues tierces énoncés par Neuner (2003 : 27 - 33) :

- l'orientation des textes : les textes sont choisis en fonction de leur forte similarité avec l'anglais ;
- l'apprentissage cognitif : l'anglais L2, le français L3, et parfois d'autres L3 romanes sont mis délibérément en relation par la comparaison des langues (Bento, 2009) ; des questions contribuent à mettre en lumière des traits du français différents de ce qu'ils sont en anglais ;
- la compréhension comme base et point de départ de l'apprentissage : l'attention est attirée sur un trait morphosyntaxique du français par des questions qui permettent de comparer avec l'anglais et éventuellement d'autres langues romanes ;
- l'orientation du contenu : nous choisissons des textes ancrés dans le domaine de spécialité des apprenants.

### **7.3.3 S'appuyer sur le lexique analogue**

#### ***7.3.3.1 Rendre transparent le lexique analogue***

Dans les activités conçues, nous adoptons une approche contrastive synchrone du français (et d'autres langues romanes) sur la base d'analogies avec l'anglais. Afin de distinguer l'analogie de la transparence, faisons un bref rappel sur la définition que nous avons donnée de l'analogie. Nous avons dit dans la section 1.5.6 (cf. Notre définition de l'analogie lexicale et de l'interlexique anglais-espagnol-français-italien-portugais, p.66) que nous comprenons l'analogie lexicale comme un double rapport : un rapport d'identité de sens et une similarité formelle. Mais l'existence d'une analogie entre deux mots n'implique pas qu'ils soient perçus comme analogues. Nous dirons donc que deux mots sont transparents pour un individu s'ils les perçoit comme analogues. Cette perception dépend de l'expérience de chacun, et elle peut être développée.

La compréhension humaine d'un texte s'opère par projections sémiologiques, autrement dit, une signification est projetée sur une forme. Lorsqu'une forme est susceptible d'être reconnue, sa reconnaissance s'obtient au prix de la tolérance d'une part de flou, par un rapprochement entre la forme connue (un mot dans la langue passerelle qu'est l'anglais dans notre configuration), et la forme inconnue (un mot de la langue cible qu'est le français ou l'une des autres langues romanes). Le rapprochement qui est fait par le lecteur dépend de la convergence / divergence qu'il perçoit entre des formes écrites dans un contexte donné. Des choix didactiques peuvent encourager le rapprochement.

### **7.3.3.2 *Traitement de l'opacité lexicale : plusieurs langues en parallèle***

Il est répandu en didactique des langues de travailler les inférences afin de faire saisir grâce au contexte le sens d'un mot opaque. Les résultats que nous avons obtenus par des mesures de similarité ouvrent d'autres possibilités.

Le graphique 30 montre que pour les 3/4 des acceptions fréquentes de ScienText, un mot anglais se traduit par un mot qui a une forme similaire dans au moins l'une des quatre langues romanes. Nous pensons qu'il est possible de valoriser l'analogie entre l'anglais et les autres langues dans les cas d'opacité du français. Pour cela, nous présentons en parallèle des phrases de sens identique. La présence d'une phrase en français et d'une phrase dans une autre langue romane, devrait atténuer la tension de l'apprenant face à ce qu'il ne connaît pas et permettre la compréhension du mot français opaque. Peut-être facilite-t-elle aussi la mémorisation du mot français. La forme du mot (fra) *données*, par exemple, n'est pas similaire à celle de (eng) *data*. Dans un contexte, son sens pourrait être éclairé par le mot (ita) *dati*, par exemple, qui est analogue à l'anglais.

Une autre direction pourra être explorée. Nous avons mentionné dans la section Vers une compréhension plus fine du discours de spécialité, p.120, l'intérêt que revêt la linguistique de corpus pour mieux appréhender la dimension phraséologique d'un discours de spécialité. Nous avons également expliqué (cf. Seconde tentative : déterminer le sens par le profil combinatoire, p.144) comment les acceptions fréquentes des mots anglais fréquents ont été délimitées au moyen du profil combinatoire. Des outils comme le Lexicoscope permettent de repérer les séquences collocatives. On peut ensuite chercher si dans la langue cible leurs équivalents sont aussi des séquences collocatives. On sélectionne alors les séquences

collocatives analogues, on extrait dans des corpus multialignés des concordances qui en contiennent et on les insère dans des activités d'apprentissage plurilingues afin que le sens d'un mot opaque du français soit éclairé par les séquences collocatives qui lui correspondent sémantiquement dans une autre langue romane et qui sont analogues à l'anglais.

À des analogies lexicales entre l'anglais, le français et les autres langues romanes s'ajoutent des analogies syntaxiques et morphosyntaxiques à prendre en compte.

### **7.3.4 Valorisation d'analogies et de dissemblances syntaxiques et morphosyntaxiques**

Nous avons dans les sections 1.5.2 et 1.5.4, pp.54 et 57, abordé des analogies et des dissemblances syntaxiques et morphosyntaxiques entre l'anglais, le français et les langues romanes. Il est temps d'expliquer comment nous envisageons leur valorisation dans des activités d'apprentissage du FOS L3 qui les fassent observer.

Premièrement, on peut appliquer dans des corpus multialignés des chaînes de traitement permettant d'extraire des concordances multilingues qui contiennent des éléments du lexique analogue et qui correspondent à des patrons morphosyntaxiques. Nous partons pour cela des structures syntaxiques panromanes du cinquième tome d'EuroCom (Klein, 2000) qui existent aussi en anglais. Voici quelques uns de ces patrons syntaxiques, Adj, Det, N, plur et V signifiant respectivement adjectif, déterminant, nom, pluriel et verbe :

- 1) Syntaxe de l'attribut : Det N est Det N
- 2) Syntaxe de l'attribut : Det N est Adj
- 3) Syntaxe ordre Sujet Verbe Objet : Det N V Det N
- 4) Morphosyntaxe du pluriel : Det plur N sont adj
- 5) Morphosyntaxe du pluriel : Det plur N sont N adj

Deuxièmement, on peut sélectionner des traits (morpho)syntaxiques divergents, extraire des concordances qui en contiennent et en organiser la découverte :

- 1) Syntaxe nom adjectif : Det N Adj est Adj



- 2) Syntaxe nom adjectif : Det N est Det N Adj
- 3) Syntaxe du groupe nominal : Det N est Det N de N

Les analogies de surface et les divergences qui existent entre l'anglais et les quatre langues romanes peuvent être étudiées par un questionnement contrastif qui vise à acquérir ces traits (morpho)syntaxiques communs ou divergents.

#### ***7.3.4.1 Un questionnement métalinguistique contrastif***

L'aptitude à généraliser à partir de l'analogie repose sur la logique. L'apprenant la développe tout autant dans sa L1 que dans d'autres langues, mais l'effort qu'il doit fournir pour s'approprier des traits morphosyntaxiques dans la langue romane cible est moindre si le lexique choisi est similaire en anglais et dans la L3 romane.

Les activités pourront alterner la découverte de traits morphosyntaxiques tantôt communs, tantôt divergents entre l'anglais et le français. La conscientisation s'effectuera par un questionnement métalinguistique dont la langue d'interface est l'anglais : au moyen de questions qu'il faudrait simples, l'apprenant sera invité à observer le français et les autres langues romanes, et à faire des comparaisons avec l'anglais qu'il connaît. L'association du français avec des mots analogues de l'anglais et des autres langues romanes devrait permettre la mémorisation. Le questionnement métalinguistique portera sur les langues repérées par un code de couleurs.

Il nous paraîtrait judicieux de réfléchir aux modalités d'un apprentissage mobile du français de la santé au moyen d'activités en ligne ou à télécharger, par exemple sur un téléphone ou une tablette. Le dispositif en ligne, qui serait une introduction à la langue française de la santé, serait compatible avec le curriculum très chargé des formations de la santé : il permettrait d'acquérir le fonctionnement des principaux traits syntaxiques et morphologiques du français et préparerait à un apprentissage dans un groupe classe.

## 7.4 Requêtes dans des corpus multialignés et exemples d'activités

### 7.4.1 Des textes authentiques contenant des mots de l'interlexique

De la même manière qu'il faisait sens d'utiliser un corpus unilingue (ScienText) pour en extraire le lexique anglais fréquent, nous nous sommes tourné vers des corpus multialignés afin d'y trouver des concordances en plusieurs langues romanes. Nous avons défini les corpus multialignés section Distinction entre corpus comparables et corpus parallèles, p.118. Nous avons eu recours à deux corpus en raison de leurs thématiques respectives, qui devraient permettre de contextualiser les apprentissages langagiers dans un domaine de spécialité, mais pas seulement, tout en privilégiant l'autorité des textes authentiques.

Les requêtes menées dans ces corpus répondent à la volonté de fournir aux apprenants des phrases qui contiennent à la fois des éléments du lexique des 1038 séries de mots analogues présents dans l'interlexique ou celui des lexiques bilingues, et les structures syntaxiques identiques aux langues romanes et à l'anglais. Les concordances trouvées contiennent aussi des séquences collocatives.

### 7.4.2 Présentation des corpus multialignés

#### 7.4.2.1 *Le corpus multiparallèle EMEA*

Ces corpus font partie du projet OPUS<sup>232</sup>. Ils sont libres de droits et téléchargeables. Une interface<sup>233</sup> offre également la possibilité d'effectuer des requêtes dans le corpus EMEA.

Le corpus d'EMEA<sup>234</sup> (Tiedemann, 2009 : 6) est un corpus parallèle publié par l'European Language Resources Association. Il se compose de 41 957 fichiers en 22 langues, avec environ 1900 fichiers par langue. Le corpus entier provient de documents pdf en ligne qui ont été convertis en XML. L'alignement phrastique a été fait pour chaque couple de langues. Nous avons recherché dans la version 0.3 du corpus EMEA.

---

232 <http://opus.lingfil.uu.se/>

233 <http://opus.lingfil.uu.se/bin/opuscqp.pl?corpus=EMEA;lang=en;adv=1>

234 Téléchargeable sur la page <http://opus.lingfil.uu.se/EMEA.php>.

Les textes, essentiellement des rapports européens publics d'évaluation, proviennent de l'Agence européenne des médicaments<sup>235</sup> et contiennent 311,65 millions d'occurrences. Cet organisme public publie un rapport pour chaque médicament auquel la Commission Européenne attribue ou refuse une Autorisation de Mise sur le Marché (AMM) suite à une évaluation par le Comité des Médicaments à Usage humain<sup>236</sup> ou par le Comité des Médicaments Vétérinaires.

Un rapport européen public d'évaluation est d'abord écrit en anglais, puis certains de ses constituants sont traduits dans les autres langues (Candito & Seddah, 2012 : 323) : le résumé, l'étiquetage (du produit), la notice d'utilisation et le résumé des caractéristiques du produit, la liste des présentations autorisées.

Étant donné sa spécialité, nous avons supposé que nous trouverions dans le corpus EMEA certaines des unités lexicales appartenant à la catégorie sémantique du lexique de la santé (cf. Sous-catégories du lexique de spécialité de la santé, p.153).

#### **7.4.2.2 *Le corpus multiparallèle Europarl***

Le corpus Europarl<sup>237</sup> (Koehn, 2005) rassemble les minutes des débats qui se sont tenus au Parlement Européen entre 1996 et 2012. Ces documents sont en général en 11 langues - en 20 langues dans la version 7 de mai 2012. L'acquisition du corpus s'est faite à partir des fichiers HTML du site<sup>238</sup> du Parlement Européen. Chaque fichier rassemble les données d'un jour de débat. Un alignement a été opéré par couple de langues pour chaque texte et chaque phrase<sup>239</sup>, nous avons utilisé les ressources multilingues qui sont disponibles sur ce site.

Nous avons supposé que nous obtiendrions dans le corpus Europarl des énoncés appartenant plutôt à la catégorie sémantique du lexique scientifique transdisciplinaire. Les échanges des députés européens abordent en effet toutes les questions régissant la vie des citoyens de l'Union Européenne, dont des questions qui ressortissent à la science.

---

235 Jusqu'en 2004, l'Agence européenne des médicaments (European Medicines Agency) s'appelait en anglais European Agency for the Evaluation of Medicinal Products, EMEA, d'où le nom du corpus.

236 <https://frama.link/A8Q8GBXq> consulté le 23/05/16.

237 Accessible sur <http://statmt.org/europarl/>

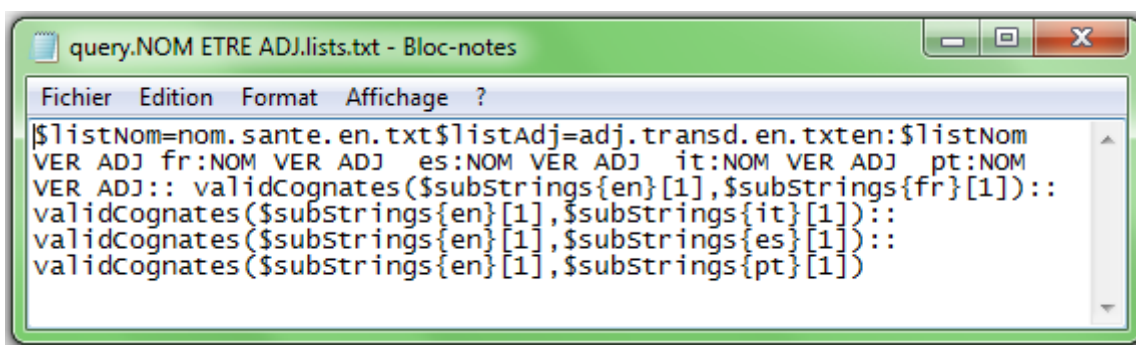
238 <http://www.europarl.eu.int/>

239 Tout cela est fait automatiquement, avec des risques d'erreurs. Ainsi, il a fallu prévoir une liste d'abréviations usuelles pour distinguer les cas où le point « . » indique la fin d'une phrase ou entre dans une abréviation.

### 7.4.3 Extractions de chaînes de caractères

Nous avons téléchargé le corpus EMEA<sup>240</sup>, décompressé les fichiers et modifié leur nom. Les fichiers d'EMEA sont nommés en fonction de la langue, par exemple 005502en, 005502es, 005502fr, 005502it, 005502pt. L'utilisation du script Cesalign exigeait que les fichiers par langue se trouvent dans cinq dossiers distincts mais qu'ils aient le même nom.

Nous voulions des concordances en anglais, espagnol, français, italien et portugais. À cette fin, un multialignement des fichiers des cinq dossiers correspondant à ces cinq langues a été obtenu en utilisant un script. Ce dernier a scindé chacun des fichiers de façon à générer autant de fichiers d'alignement qu'il existe de fichiers dans chacune des langues. Les fichiers d'alignement générés sont en-fr.xml, en-es.xml, en-pt.xml<sup>241</sup>, autrement dit, ils mettent en correspondance chaque segment du texte en langue source (généralement une phrase) avec le segment de sa traduction dans la langue cible. Pour ces quatre fichiers, les segments sources sont en anglais et les segments cibles dans les quatre langues romanes qui nous intéressent.



```
query.NOM ETRE ADJ.lists.txt - Bloc-notes
Fichier  Edition  Format  Affichage  ?
|$listNom=nom.sante.en.txt$listAdj=adj.transd.en.txten:$listNom
VER ADJ fr:NOM VER ADJ es:NOM VER ADJ it:NOM VER ADJ pt:NOM
VER ADJ:: validCognates($subStrings{en}[1],$subStrings{fr}[1]):
validCognates($subStrings{en}[1],$subStrings{it}[1]):
validCognates($subStrings{en}[1],$subStrings{es}[1]):
validCognates($subStrings{en}[1],$subStrings{pt}[1])
```

Image 12 : Copie de la requête Nom être Adjectif

Pendant cette recherche, nous avons cherché des patrons élémentaires tels que N être Adj (cf. image 58259) ou N adj. Dans la section suivante, nous présentons le premier fruit de ces extractions.

240 <http://opus.lingfil.uu.se/EMEA.php>

241 Nous nous inspirons des explications fournies dans cette page du Département de Traitement Informatique Multilingue de l'Université de Genève : <http://www.issco.unige.ch/en/research/projects/ecolore/tms/AlignDefScenario.htm>

## 7.4.4 Résultats partiels d'extractions dans les deux corpus multialignés

### 7.4.4.1 Écart entre la langue, le discours et la didactisation

Nous observons un contraste entre les analogies qui existent en langue, celles qu'on rencontre dans les textes et leur didactisation : les équivalents de traduction sont certes analogues dans les cinq langues pour 47 % des acceptions fréquentes dans ScienText, mais il faut beaucoup de temps pour sélectionner des concordances dans d'immenses corpus alignés et les rendre exploitables avec des apprenants. Nous ne présentons donc de cette recherche que des résultats embryonnaires.

Les résultats de nombreuses redites liées au genre très normé des documents de l'EMEA nous ont conduit à nous tourner vers le corpus Europarl.

### 7.4.4.2 Exemples de concordances didactisables du corpus EMEA

Voici des concordances trouvées dans le corpus EMEA. On voit qu'une concordance comporte des phrases qui sont des équivalents de traduction dans plusieurs langues. Les activités n'incluront pas la phrase en anglais.

Concordance 1)

(eng) The diagnosis must be confirmed using specific tests, such as the HIPAA test.

(spa) El diagnóstico deberá ser confirmado mediante pruebas específicas, como el test HIPAA

(fra) Le diagnostic doit être confirmé par des tests spécifiques, tels que le test HIPAA.

(ita) La diagnosi deve essere confermata da test specifici, quali il test HIPAA.

(por) O diagnóstico deve ser confirmado por testes específicos, nomeadamente HIPAA.

La phrase en français de la concordance 1) contient plusieurs mots de l'interlexique et présente une syntaxe analogue à celle de l'anglais : *diagnostic, confirmé, tests, spécifiques, test*. Dans cette phrase, le verbe devrait poser un problème d'interprétation : dans ce contexte, *doit être* pourrait être compris par inférence comme signifiant *has been* par quelqu'un qui n'a

pas appris le français. Le sigle *HIPAA*, associé à *test*, serait éventuellement compris comme le nom du *test*, mais il faudrait déjà avoir vu la détermination.

#### Concordance 2)

(eng) Refludan is given by injection and intravenous infusion (drip into a vein).

(spa) Refludan se administra mediante inyección e perfusión intravenosa (goteo en vena).

(fra) Refludan est administré par injection et par perfusion intraveineuse (goutte-à-goutte dans une veine).

(ita) Refludan è somministrato per iniezione e per infusione endovenosa (flebo in una vena).

(por) O Refludan é administrado por injeção ou perfusão intravenosa (gota a gota numa veia).

Dans la concordance 2), la préposition *par* serait le mot le plus difficile. Le mot *given* a été traduit par *administré*, qui a l'avantage d'être analogue à (eng) *administer*. La parenthèse pourrait être enlevée. Cette concordance peut être utilisée pour découvrir le passif (cf aussi Analogies et dissemblances morphosyntaxiques des verbes de l'anglais et des langues romanes, p.59), mais cette notion ne figure pas parmi les notions simples que nous avons listées.

#### Concordance 3)

(eng) The recommended dose is 0.4 mg/ kg body weight as a single intravenous injection followed by 0.15 mg/ kg body weight per hour as a continuous intravenous infusion for 2 to 10 days, or longer if needed.

(spa) La dosis recomendada es de 0,4 mg/ kg de peso corporal en una sola inyección intravenosa, seguida de 0,15 mg/ kg de peso corporal/ hora en perfusión intravenosa continuada de 2 a 10 días, o más tiempo si es preciso.

(fra) La dose recommandée est de 0,4 mg/ kg de poids corporel administrée en une seule injection intraveineuse, suivie de 0,15 mg/ kg de poids corporel par heure en perfusion intraveineuse continue pendant 2 à 10 jours, voire davantage si nécessaire.

(ita) La dose raccomandata è di 0,4 mg/ kg di peso corporeo con un'unica iniezione endovenosa seguita da 0,15 mg/ kg di peso corporeo all'ora come infusione endovenosa continua per 2-10 giorni o più a lungo se necessario.

(por) A dose recomendada é de 0,4 mg/ kg peso corporal por via intravenosa numa única injeção, seguidos de 0,15 mg/ kg peso corporal/ hora sob a forma de perfusão contínua endovenosa durante 2 a 10 dias ou durante um período mais prolongado, caso necessário.

La concordance 3) présente plusieurs intérêts : la plupart des mots sont analogues et il existe d'autres indices qui mettent sur la voie, comme des chiffres et la ponctuation. Pour faire deviner les mots opaques *suivie de*, *pendant* et *si*, le travail des inférences serait nécessaire. On pourrait tabler pour le second sur la séquence collocative *if necessary*. En revanche, *voire davantage* ne serait probablement pas compris. Ce document nous semble cependant didactisable.

#### Concordance 4)

(eng) The active substance of Econor is valnemulin, an antibiotic of the pleuromutilin group. It works by blocking the production of proteins in the bacteria's ribosomes and inhibits the growth of bacteria.

(spa) El principio activo de Econor es la valnemulina, un antibiótico del grupo de las pleuromutilinas que actúa bloqueando la producción de proteínas en los ribosomas bacterianos e inhibe el crecimiento de las bacterias. La valnemulina es activa frente a diversas bacterias, entre ellas las responsables de las enfermedades intestinales y respiratorias antes citadas.

(fra) Le principe actif d'Econor, la valnémuline, est un antibiotique du groupe des pleuromutilines. Son action consiste à bloquer la production de protéines dans les ribosomes de la bactérie et à inhiber la croissance des bactéries.

(ita) Il principio attivo di Econor è valnemulina, un antibiotico del gruppo delle pleuromutiline, che agisce bloccando la produzione di proteine nei ribosomi batterici e, in tal modo, inibisce la crescita batterica.

(por) A substância activa do Econor é a valnemulina, um antibiótico do grupo das pleuromutilinas, que actua através do bloqueio da produção de proteínas nos ribossomas bacterianos e inibe o crescimento de bactérias.

La phrase en portugais, dont le début contient une collocation analogue avec l'anglais, serait utile pour faire comprendre le français *principe actif*. Les nombreuses autres analogies rendent cette concordance 4) intéressante à traiter.

Nous ne donnons que quatre exemple de concordances d'EMEA. Le contenu, on le voit, est très spécialisé.

#### **7.4.4.3 Exemples de concordances didactisables du corpus Europarl**

Nous proposons ci-après des concordances tirées du corpus Europarl.

Concordance 5)

(eng) safety is a priority objective for the Commission

(spa) la seguridad es un objetivo prioritario de la Comisión

(fra) la sécurité est un objectif prioritaire de la Commission

(ita) la sicurezza è un obiettivo prioritario della Commissione

(por) a segurança constitui um objetivo prioritário da Comissão

Indépendamment de la différence entre (eng) *safety* et (fra) *sécurité*, la concordance 5) contient des termes simples. Cependant, elle ne permet pas de contextualisation. Peut-être pourrait-elle être utilisée comme phrase d'application de l'ordre N adj qui aurait été découvert précédemment au moyen d'un autre contexte. La question de l'acquisition préalable de la syntaxe N de N se poserait aussi.



Concordance 6)

(eng) social exclusion is a serious issue

(spa) la exclusión social es una cuestión grave

(fra) l'exclusion sociale est un problème grave

(ita) l'esclusione sociale è un grave problema

(por) a exclusão social é um problema muito grave

Dans la concordance 6), la syntaxe est très simple. Le sens de l'adjectif *grave* pourrait être proposé dans une liste à choix multiple, mais il entrerait peut-être en conflit avec le nom (eng) *grave* (la tombe). La difficulté principale serait de faire comprendre que l'anglais n'utilise pas de déterminant dans ce cas, au contraire des langues romanes, ceci à des locuteurs de langues qui ne connaissent éventuellement pas de déterminants.

La réflexion portant sur ces deux concordances et celles extraites dans le corpus EMEA pose la question de la progression dans une telle méthode. Au sujet des compétences acquises au moyen d'une approche intercompréhensive, Caddéo et Jamet notent :

« Un lecteur formé à l'intercompréhension pourra trouver tout à fait accessibles des textes de niveau B2/C1 du Cadre et rencontrer des difficultés face à des documents de niveau A2 à cause de leur brièveté ou de l'opacité de certains termes appartenant au domaine du quotidien ».

La progression s'éloignera donc probablement des niveaux du CECRL et pour plutôt prendre en compte les traits communs au français et à l'anglais, les traits divergents, et la possibilité de recourir aux autres langues romanes dans les cas d'opacité du français.

#### **7.4.4.4 Exemple de didactisation de la structure attributive**

Les textes scientifiques, dans leur description des faits, recourent très largement à la structure attributive, en témoigne la première place du verbe *be* dans la liste des 2000 mots, c'est pourquoi elle figure parmi les premières structures que nous avons listées. Par ailleurs, l'article indéfini anglais *an*, employé devant un nom commençant par une voyelle, est le seul

déterminant analogue avec les langues romanes : (spa) / (fra) / (ita) *un*. Soit la phrase : *La chimiothérapie est un traitement*. On peut poser les questions suivantes :

Which French words look like English words ? Please tick  them.

la     chimiothérapie  est  un  traitement

These sentences are in **Spanish**, **French**, Portuguese and Italian. Please observe them.

La quimioterapia    es    un tratamiento.

La chimiothérapie    est    un traitement.

A quimioterapia    é    um tratamento.

La chemioterapia    è    un trattamento.

Without using a dictionary, can you guess the English word for the **Spanish word** *es* ?

YES / NO

If so, please write this English word here :

Compare the **Spanish** *es* and the **French** *est*. What does the French *est* mean ?

Please write your answer here :

Comme on le voit dans ce questionnaire, la même phrase est présentée en parallèle dans les 4 langues romanes. Elles sont écrites en différentes couleurs, selon la langue. C'est l'espagnol *es* qui est le plus analogue avec l'anglais *is*, c'est pourquoi la première langue dans laquelle est écrit l'énoncé est l'espagnol, quand bien même le français est la langue cible. L'apprenant doit cocher les mots qu'il comprend, inférer le sens, rapprocher des mots de plusieurs langues romanes afin de deviner le sens du mot français, montrer s'il a compris en écrivant la signification dans une case.

#### **7.4.4.5 Autre piste prometteuse : des requêtes en deux langues**

Une autre piste sera de mener des requêtes non pas dans les cinq langues, mais dans deux. En anglais et français dans le cas général, en anglais et dans l'une des autres langues romanes dans les cas d'opacité. Nous aurions alors probablement plus de retours en raison de la plus

grande analogie entre l'anglais et le français (66%) et entre les autres paires de langues : anglais-italien : 65 %, anglais - espagnol : 63 % et anglais-portugais : 58 %.

On trouve dans Europarl les grandes questions plus ou moins consensuelles qui sont débattues par les député.e.s de l'Union Européenne. On rencontre par exemple des phrases telles que : *Le progrès économique et le progrès social ne sont pas deux directions opposées*, etc. Elles sont empruntées à des discours écrits oralisés, essentiellement déclaratifs ou injonctifs.

Dans ce corpus, on peut chercher des collocations qui contiennent justement ces formules. Par leur caractère collocatif, elles se prêtent à une activité telle que mettre les mots dans l'ordre. Ainsi, ayant fait observer une phrase dans laquelle on a des occurrences de la structure Det N Adj, on pourrait proposer dans l'ordre alphabétique les mots qui composent les collocations suivantes : *des effets positifs, le rôle essentiel, une cohérence interne, l'effort commun, des conséquences à long terme, un projet international, l'image positive*, etc.

On donne alors la consigne : *Put these words in the correct order in French*. On vérifie ainsi si la structure est acquise. Comme ces collocations sont analogues en anglais et en français, elles devraient pouvoir être reconnues.

## 7.5 Conclusion du chapitre 7

Nous avons montré dans ce chapitre le rapport entre l'étude lexicologique dont rendent compte les chapitres 4 à 6 et un projet didactique dans lequel l'anglais est la langue passerelle vers le français. Nous avons expliqué les principes sur lesquels nous fondons ce projet : pouvoir lire une langue sans l'avoir apprise en allant du connu à l'inconnu, orienter le contenu, stimuler la perception de la proximité par un questionnement métalinguistique. Puis nous avons fait le lien entre les analogies syntaxiques et morphosyntaxiques étudiées dans le chapitre 1 et les patrons syntaxiques que nous appliquons lors de la recherche de contextes dans des corpus multialignés EMEA et Europarl. Nous avons présenté ces corpus et donné des exemples de concordances que nous y avons trouvées.

Le travail de fouilles prend beaucoup de temps et présente une technicité certaine. Quelques concordances trouvées dans le corpus EMEA sont très intéressantes, mais en général, ses textes sont fortement redondants. Les requêtes dans Europarl semblent donner des résultats plus riches, mais relevant essentiellement du LST. Un élargissement des requêtes à d'autres corpus pourrait fournir un éventail de concordances qui correspondrait davantage à l'interlexique et aux lexiques bilingues. Le corpus ACQUIS (Steinberger *et al.* : 2006), qui recueille des textes de loi de l'Union Européenne, donnerait d'autres contextes. On peut aussi chercher des énoncés authentiques qui contiennent des séquences collocatives constituées d'un pivot anglais qui admettrait une traduction analogue dans une seule langue romane, et d'un de ses collocatifs analogues. Les possibilités d'énoncés plurilingues similaires seraient accrues, même si les collocatifs ne comptent pas parmi les mots fréquents que nous avons identifiés dans notre étude.

Le critère de la fréquence a été au cœur de cette thèse : nous pensons que les mots fréquents de l'anglais (L2) seront connus et compris des apprenants de FLE (L3), mais d'autres paramètres entrent en compte pour la compréhension d'un texte, à l'instar des articulateurs du discours, qui organisent les écrits scientifiques, lesquels sont essentiellement explicatifs et argumentatifs. Ces paramètres devront être pris en compte dans la didactisation. Nous nous interrogerons sur l'efficacité pédagogique et le niveau d'applicabilité de ce dispositif plurilingue de découverte d'une L3 romane à partir des analogies interlinguistiques avec l'anglais. Une expérimentation devra montrer si les activités inspirées de l'intercompréhension que nous avons esquissées permettraient d'accéder au français par une approche contrastive dans laquelle l'anglais L2 est une langue passerelle.

## *Conclusion générale*

Lors de cette recherche, nous avons établi une liste des adjectifs, noms et verbes anglais fréquents du corpus ScienText. Nous les avons traduits en espagnol, français, italien et portugais afin de mesurer l'analogie qui existe dans le discours scientifique de la santé entre le lexique de l'anglais et le lexique de ces quatre langues romanes. La mesure de cette analogie interlexicale servait deux objectifs principaux : d'une part, contribuer à la lexicologie, d'autre part, réfléchir à la valorisation de l'analogie lexicale et (morpho)syntaxique entre une L2 pont, l'anglais, et une L3 cible, le français.

Nous avons opté pour une approche sur corpus plutôt que pour des textes arbitrairement choisis. Bien que le corpus ScienText rassemble 35 millions de mots, nous devons considérer les résultats obtenus avec prudence, car certaines disciplines de la santé, à l'instar de la pharmacie ou de l'odontologie, ne sont pas présentes, ce qui pose la question de la représentativité des données. Nous avons recouru au profil combinatoire afin de délimiter les acceptions fréquentes en ce que les collocatifs des mots anglais fréquents ont permis de repérer des séquences collocatives qui suggèrent des acceptions.

Pour définir et traduire, nous avons utilisé plusieurs types de ressources, dont des ressources collaboratives telles que *Wikipédia* et le *Wiktionary*. Il serait intéressant d'étudier l'influence de ces outils sur les usages. L'étude des mots de la santé qui pénètrent peu à peu les dictionnaires généralistes, ainsi que la convergence des langues par les contacts entre leurs utilisateurs dans ce domaine offrent d'autres pistes.

Les données que nous avons élaborées pourront enrichir les dictionnaires interlingues, et servir ainsi à la traduction automatique dans les systèmes qui utilisent des bases lexicales interlingues par acceptions (Sérasset & Blanc, 1993). Elles pourront être complétées par la traduction dans d'autres langues romanes dans lesquelles sont menées des recherches dans le domaine de la santé (catalan, roumain). La liste des acceptions du corpus relevant du LST pourra aussi être comparée à des travaux que nous n'avons pas utilisés, par exemple l'Academic Word List (AWL, Coxhead, 2002).

La traduction proposée présente des limites parce que nous ne sommes ni terminologue, ni traducteur, ni expert de la santé. Notre objectif était avant tout didactique : délimiter des bases

de transferts interlinguistiques. Afin de donner une image fidèle de l'analogie, il faudrait travailler sur des corpus comparables : des corpus de spécialité dans chaque langue pourraient être constitués à partir de textes qui auraient des caractéristiques identiques. En comparant de tels corpus, nous n'étudierions pas l'analogie qui peut exister entre des mots anglais fréquents du discours scientifique et leur traduction possible dans des langues romanes, mais bien l'analogie interlinguistique du discours scientifique.

Afin de mesurer la similarité, le script CompareCols.csv.pl a été utilisé en plusieurs fois. Les paramètres de cet outil sont faciles à ajuster. De l'interlexique et des lexiques bilingues, on peut retenir la forte similarité formelle dans ce type de discours entre l'anglais et les langues romanes, surtout entre l'anglais et le français et entre l'anglais et l'italien. Les traductions peuvent orienter le travail des didacticiens des langues romanes, mais aussi des didacticiens de l'anglais qui voudraient prendre en compte la L1 ou la L2 de leurs étudiants locuteurs de langues romanes.

Nous avons formulé dans l'introduction quatre hypothèses sur le lexique des discours scientifiques de la santé. Nous allons les réexaminer.

#### Hypothèse 1 :

Dans le discours scientifique de la santé écrit en anglais, le lexique de spécialité fréquent contient une majorité de mots analogues en anglais, espagnol, français, italien et portugais.

Cette hypothèse est confirmée si on prend en compte la moyenne, mais les seuls adjectifs, dépassent les 50 %, contrairement aux noms, et aux verbes. Les résultats demeurent néanmoins très intéressants dans la perspective d'un enseignement plurilingue.

#### Hypothèse 2 :

Les mots anglais polysémiques sont moins analogues à leurs équivalents de traduction dans les langues romanes que les mots monosémiques.

Cette hypothèse est confirmée, mais la différence n'est pas significative : qu'un mot ait admette une ou plusieurs acceptions, les mots des langues romanes admettent souvent les mêmes acceptions que le mot anglais, autrement dit les cinq langues présentent une forte convergence sémantique.

### Hypothèse 3 :

Dans le discours scientifique de la santé écrit en anglais, le lexique de spécialité présente une plus grande analogie avec l'espagnol, le français, l'italien et le portugais que le LST.

Cette hypothèse est confirmée pour les trois catégories grammaticales, d'autant plus si on tient compte des occurrences.

Au moyen de la linguistique de corpus, nous avons pu étudier le discours scientifique sur la base des combinaisons libres des mots d'écrits de la santé dans le corpus anglais de spécialité de ScienText. Nous avons également extrait automatiquement des séquences de mots analogues : les recherches en cours dans les corpus multilingues alignés EMEA et Europarl nécessitent une préparation des fichiers et une maîtrise des patrons syntaxiques qui les rendent encore peu aisés pour le novice, mais le nombre de contextes obtenus montre la pertinence de ces recherches. Les premières extractions de concordances en fonction de chaînes de caractères donnent beaucoup de bruit, mais une piste sera de chercher par couples de langues, en partant éventuellement de séquences collocatives ou de collocatifs de mots anglais fréquents.

L'originalité de notre approche plurilingue réside dans une observation d'énoncés présents simultanément en plusieurs langues et dans un questionnement métalinguistique contrastif portant sur des traits morphosyntaxiques.

La L3 cible dans cette thèse est FLE. Néanmoins, l'interlexique et les lexiques bilingues peuvent être utilisés pour enseigner l'espagnol, l'italien et le portugais à un public de spécialistes de la santé sur le principe d'une intercompréhension de spécialité. Bien plus, le recours à une autre L2 pont que l'anglais pour enseigner une L3 appartenant à une autre famille de langues nous paraît prometteur dans le cas de L3 qui partagent avec la L2 une partie du même fond lexical : les langues slaves partagent avec les langues romanes et les langues germaniques un lexique d'origine gréco-latine (et germanique) ; le turc et le farsi L3 pourraient être rendus plus proches à des apprenants qui ont appris l'arabe en L2 ; le japonais et le coréen L3 pourraient être enseignés en partant d'acquis en chinois L2, et vice versa, etc. De manière générale, une scénarisation qui valorise des analogies multiples entre la L2 de l'apprenant et la L3 devrait permettre l'apprentissage d'une langue distante avec une pédagogie des langues apparentées.

# Références bibliographiques

Les références bibliographiques qui suivent sont classées selon trois types de références. En premier lieu, dans la sitographie, nous indiquons les ressources en ligne telles que les sites de plateformes, d'outils et d'institutions. Nous avons placé en second lieu les ressources que nous avons utilisées spécifiquement à des fins de traduction. En dernier lieu sont classées toutes les autres ressources dans la bibliographie générale.

## Sitographie

Ces sites étaient consultables au 30/06/17.

- Association Réseau pour la nouvelle orthographe du français : <http://www.renouvo.org>  
Académie française : [http://www.academie-francaise.fr/sites/academie-francaise.fr/files/rectifications\\_1990.pdf](http://www.academie-francaise.fr/sites/academie-francaise.fr/files/rectifications_1990.pdf)  
Agence européenne des médicaments : <http://www.ema.europa.eu/ema>  
Centre européen pour les langues vivantes : <http://www.ecml.at/Home/tabid/59/language/fr-FR/Default.aspx>  
Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues : Apprendre, enseigner, évaluer*, Didier, Paris. Consulté le 30/06/17 sur [www.coe.int/t/dg4/linguistic/Cadre1\\_fr.asp](http://www.coe.int/t/dg4/linguistic/Cadre1_fr.asp)  
Corpus OPUS : <http://opus.lingfil.uu.se/>  
Corpus Europarl : <http://statmt.org/europarl>  
Corpus JRC-Acquis : [http://optima.jrc.it/Acquis/index\\_2.2.html](http://optima.jrc.it/Acquis/index_2.2.html)  
Diário de Notícias : <http://www.dn.pt/>  
Éduscol, portail du Ministère français de l'Éducation nationale : <http://eduscol.education.fr>  
Faculté Computing in the Humanities and Social Sciences de l'Université de Toronto : <http://www.utoronto.ca/index.html>  
Lexicoscope : <http://turing.u-grenoble3.fr/~emolex/lexicoscope/index.php>  
Miriadi : <https://www.miriadi.net>  
Multilingual Information Processing Department, Université de Genève : <http://www.issco.unige.ch/en>  
Organisation Internationale de Normalisation : <https://www.iso.org/fr/home.html>  
ScienQuest : <http://corpora.aiakide.net/scientext18/>  
ScienText : [scientext.msh-alpes.fr/scientext](http://scientext.msh-alpes.fr/scientext)  
SciTE : <http://www.scintilla.org/SciTE.html>  
Take care, healthcare language : <http://takecareproject.eu> et le guide pour migrants : <http://takecareproject.eu/visibility>



## Ressources utilisées pour la traduction

Azzaretti, Michel (2008). *Dictionnaire médical en six langues : français-anglais-allemand-espagnol-italien-russe*. Paris : La Maison du Dictionnaire.

Cours-pharmacie.com : <http://www.cours-pharmacie.com>

DeCS (Descritores em Ciências da Saúde) : <http://decs.bvs.br/cgi-bin/wxisl660.exe/decserver/>

Dicciomed : <http://dicciomed.eusal.es>

Dictionnaire de portugais *Dicio* : <http://www.dicio.com.br/>

*Dictionnaire médical de l'Académie de Médecine*, version 2016-1 : <http://dictionnaire.academie-medecine.fr>

*Dictionnaires de iH* : <http://www.informationhospitaliere.com/dicodoc.php>

*Diccionario de genotoxicología* : <http://www2.udec.cl/~digentox/index.html>

*Diccionario de la Lengua Española* : <http://www.rae.es/recursos/diccionarios/drae>

Diccionario-medico : <http://www.cun.es/diccionario-medico>

Duizabo, Danielle (2009). *Lexique médical. Anglais-français, français-anglais*. Issy-les-Moulineaux : Elsevier Masson.

Encyclopédie *Treccani* : <http://www.treccani.it/vocabolario/>

Freitas e Costas, Manuel (2005). *Dicionário de termos médicos*. Porto : Porto Editora.

Garnier, Marcel ; Delamare, Jacques ; Delamare, François ; Gélis-Malville, Élisabeth ; Delamare, Laurent (2012). *Dictionnaire illustré des termes de médecine*, 31<sup>e</sup> édition. Paris : Maloine.

Geysen, Raymond (1985). *Dictionnaire des formes analogues en 7 langues : latin, italien, espagnol, français, anglais, néerlandais, allemand avec résumé de grammaire comparée*. Paris : Duculot.

Glossario du site BIOteach : <http://bioteach.ariel.ctu.unimi.it/v3/home/PreviewArea.aspx?name=Glossario>

*Glosario Hablado de Términos Genéticos* : <https://www.genome.gov/glossarys/index.cfm?textonly=&search=targeting>

Hillis, David M.; Sadava, David; Heller, H.Craig ; Price, Mary V. (2013). *Fondamenti di biologia*. Bologna : Zanichelli Editore. [http://online.universita.zanichelli.it/hillis-files/appendici/Hillis\\_Glossario.pdf](http://online.universita.zanichelli.it/hillis-files/appendici/Hillis_Glossario.pdf)

IATE :

<http://iate.europa.eu/SearchByQueryLoad.do;jsessionid=2k8fJLZWTnZLs97WWTt17JqyxBvNglkvVHvGPzvQWqDqMMLcrKdC!1759093768?method=load>

Portail HeTOP (Health Terminology/Oncology Portal) du CHU de Rouen : <http://www.hetop.eu/hetop/RESTR>

Juilland, Alphonse ; Brodin, Dorothy ; Davidovitch, Catherine ; Ignatius, Mary Ann ; Juilland, Ileana ; Szklarczyk, Lilian (1970). *Frequency dictionary of French words*. La Hague ; Paris : Mouton.

Juilland, Alphonse ; Traversa, Vincenzo Paolo (1973). *Frequency dictionary of Italian words*. La Hague ; Paris : Mouton.

Juilland, Alphonse ; Chang-Rodríguez, Eugenio (1964). *Frequency dictionary of Spanish words*. La Hague : Mouton.

Kernbaum, Serge (2008). *Dictionnaire de médecine Flammarion, 8<sup>e</sup> édition*. Paris : Flammarion.

Linguee : <http://www.linguee.es/espanol-ingles>

Maniez, François (Éd.) (2009). *Dorland Dictionnaire médical bilingue français-anglais / anglais-français*. Issy-les-Moulineaux : Elsevier-Masson.

Manuila, L., Manuila, A, Lewalle, P., et al. (2011). *Dicionário Médico*. Lisboa : Climespi Editores.

MedlinePlus : <https://medlineplus.gov/spanish/encyclopedia.html>

Nigro, Vincenzo (2014) *Lezioni di genetica medica*. <http://www.vincenzonigro.it/Lezioni%20genmed%20Spec.2014a.pdf>

Oxford English Dictionary : <http://www.oed.com/>

Portail TermSciences : <http://www.termosciences.fr>

Quevauvilliers, Jacques (2007). *Dictionnaire médical de poche, 2<sup>e</sup> édition*. Paris : Masson.

Rey, Alain (1988). *Le Grand Robert de la langue française : Dictionnaire alphabétique et analogique de la langue française*. Paris : Le Robert.

Reverso Context : <http://context.reverso.net>

Saladriga, Verónica ; Claros, Gonzalo (2002). Vocabulario inglés-español de bioquímica y biología molecular. *Panacea@* Vol.3, n°9-10. Décembre, 2002. [http://www.medtrad.org/panacea/IndiceGeneral/Pana9\\_tradyterm\\_mvsgc.pdf](http://www.medtrad.org/panacea/IndiceGeneral/Pana9_tradyterm_mvsgc.pdf)

TERMIUM Plus : <http://www.btb.termiumplus.gc.ca>

Trésor de la Langue Française informatisée : <http://atilf.atilf.fr>

*United Medical Dictionary* : <http://www.emro.who.int/fr/Unified-Medical-Dictionary.html>

Universidade Federal de Pernambuco : [https://www.ufpe.br/biolmol/aula3\\_RNAtranscri.htm](https://www.ufpe.br/biolmol/aula3_RNAtranscri.htm)

Università di Roma 1, cours en ligne : [https://elearning2.uniroma1.it/pluginfile.php/97686/mod\\_resource/content/4/12%20splitting%20RNA.pdf](https://elearning2.uniroma1.it/pluginfile.php/97686/mod_resource/content/4/12%20splitting%20RNA.pdf)

Wikipedia : <https://www.wikipedia.org/>

Wikisource : [https://wikisource.org/wiki/Main\\_Page](https://wikisource.org/wiki/Main_Page)

Wiktionary : [https://en.wiktionary.org/wiki/Wiktionary:Main\\_Page](https://en.wiktionary.org/wiki/Wiktionary:Main_Page)

Wiktionnaire : [https://fr.wiktionary.org/wiki/Wiktionnaire:Page\\_d%E2%80%99accueil](https://fr.wiktionary.org/wiki/Wiktionnaire:Page_d%E2%80%99accueil)

## Bibliographie générale

- Adams, James Noel (2007). *The regional diversification of Latin 200 BC-AC 600*. Cambridge : Cambridge University Press.
- Adron, Nisa & Hourck, Johanne (2011). *Fais ta valise ! Création d'un jeu pour les dyslexiques phonologiques visant à la remédiation des confusions auditives de type sourde-sonore*. Mémoire en vue de l'obtention du Certificat de Capacité d'Orthophonie. Université Lille 2.
- Alinei, Mario (1996). *Origini delle lingue d'Europa. 1, La teoria della continuità*. Bologne : Il Mulino.
- Aparicio, Xavier. *Mécanismes cognitifs du changement de langue chez les multilingues : Etudes comportementales et électrophysiologiques*. Thèse en Psychologie, Montpellier III : 2003. Editions Universitaires Européennes, 2011. Consulté le 30/06/17 sur <http://www.theses.fr/2010MON30078>
- Aristote (1996). *Poétique*. Trad. J. Hardy. Paris : Gallimard.
- Aristote (1997). *Éthique à Nicomaque*. Trad. J. Tricot. Paris : Librairie philosophique J. Vrin.
- Bacquet, Paul (1974). *Le vocabulaire anglais*. Paris : Presse Universitaire de France.
- Bailly, S., Boulton, A., Château, A., Duda, R. & Tyne, H. (2009). *L'anglais langue d'appui pour l'apprentissage du français langue étrangère*. In Gilles F. (Éd.), *L'anglais et le plurilinguisme. Pour une didactique des contacts et des passerelles linguistiques*. Paris : L'Harmattan, 2009.
- Banniard, Michel (1997). *Du latin aux langues romanes*. Paris : Nathan.
- Bär, Marcus (2004). *Europäische Mehrsprachigkeit durch rezeptive Kompetenzen : Konsequenzen für Sprach- und Bildungspolitik*. Aachen : Ed. EuroCom, 18.
- Baugh, Albert Croll & Cable, Thomas (1994). *A history of the English language*. London : Routledge.
- Bède (731). *Historia ecclesiastica gentis Anglorum*. Consulté le 30/06/17 sur [https://la.wikisource.org/wiki/Historia\\_Ecclesiastica\\_gentis\\_Anglorum\\_-\\_Liber\\_Primum](https://la.wikisource.org/wiki/Historia_Ecclesiastica_gentis_Anglorum_-_Liber_Primum)
- Bellassen, Joël (2011). Is Chinese Eurocompatible ? Is the Common European Framework Common ? : The Common European Framework of References for Languages Facing Distant Language. In *New prospect for Foreign Language Teaching in Higher Education – Exploring the Possibilities of Application of CECR-*, Tokyo, World Language and Society Education Center (WoLSEC), 23-31.
- Bento, Margaret. (2009). S'appuyer sur la comparaison entre les langues. *Le français dans le monde*, septembre-octobre 2009, (365), 30-31.
- Berger, Ruth (2014). L'arrivée des langues indo-européennes en Europe. *Dossier pour la Science*, (82), janvier – mars 2014, 22-25.
- Berglund, Björn E. (2003). Human impact and climate changes – synchronous events and a clausal link ?, *Quaternary International*, 105(1). New-York : Elsevier, 7-12.
- Bergounioux, Gabriel (2010). La langue des Indo-européens ? *Dossiers d'archéologie*, (338), mars/ avril 2010, 14-21.

- Bielenia-Grajewska, Magdalena (2009). Linguistic borrowing in the English language of economics. In *Journal in English Lexicology*, (3). 107-136. Consulté le 30/06/17 sur <https://lexis.revues.org/643>
- Blanche-Benveniste, Claire (Éd.) (1997). *Eurom4. Méthode d'enseignement simultané des langues romanes : Portugues, Español, Italiano, Français*. Florence : La Nuova Italia Editrice.
- Blanchet, Philippe, Clerc, Stéphanie, Rispaïl, Marielle (2014). *Réduire l'insécurité linguistique des élèves par une transposition didactique de la pluralité sociolinguistique. Pour de nouvelles perspectives sociodidactiques avec l'exemple du Maghreb*. In *Études de linguistique appliquée*, Klincksieck (Didier Erudition jusqu'en 2003), Insécurité linguistique en éducation : approches sociologiques comparées des élèves issus du Maghreb. (175), 283-302. Consulté le 30/06/17 sur [hal-01486080](https://hal-01486080)
- Blumenthal, Peter (2002). Profil combinatoire des noms. Synonymie distinctive et analyse contrastive, *Zeitschrift für Französische Sprache und Literatur*, (112), 115-138.
- Boch, Françoise ; Rinck, Fanny (2010). Pour une approche énonciative de l'écrit scientifique, *Lidil* [En ligne], (41), 5-14. Consulté le 30/06/17 sur <http://lidil.revues.org/3004>
- Bogaards, Paul (1994). *Le vocabulaire dans l'apprentissage des langues étrangères*. Paris : Didier.
- Bono, Mariana (2008). Influences interlinguistiques dans l'apprentissage d'une L3 : quand les langues secondes l'emportent sur la langue première. In Danielle Moore, et Véronique Castellotti (Éd.), *La compétence plurilingue : regards francophones*. Berne : Peter Lang, 147-166.
- Bonvino, Elisabetta ; Caddéo, Sandrine ; Vilagines Serra, Eulàlia ; Pippa, Salvador (2011). *EuRom5 : Lire et comprendre cinq langues romanes*. Milan : Ulrico Hoepli Editore.
- Bopp, Franz (1816). *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprachen*. Francfort : in der Andreäischen Buchhandlung.
- Bourcier, Georges (1978). *Histoire de la langue anglaise du moyen-âge à nos jours*. Paris : Bordas.
- Bramanti, Barbara (2013). The use of DNA analysis in the archeology of death and burial. In Liv Nilsson Stutz & Sarah Tarlow (Ed.) : *The Oxford handbook of the archaeology of death and burial*. Oxford : Oxford University Press, 99-122.
- Braun, Peter (2002). Internationalismen : Fallstudie. In David Alan Cruse, Franz Hundsnurscher, Michael Job. *Lexicology. An international Handbook on the Nature and Structure of Words and Vocabularies*. Berlin, New York : W. de Gruyter. 1380-1384.
- Caddéo, Sandrine ; Jamet, Marie-Christine (2013). *L'intercompréhension : une autre approche pour l'enseignement des langues*. Paris : Hachette, français langue étrangère.
- Candelier, Michel. (2008). Approches plurielles, didactiques du plurilinguisme : le même et l'autre. In *Cahiers de l'ACEDLE*, (5), 65-90.
- Candelier, Michel et al. (2012). *Le CARAP – Compétences et ressources. / FREPA – Competences and resources*. Strasbourg & Graz : Conseil de l'Europe. Consulté le 30/06/17 sur <http://carap.ecml.at/>

- Candito, Marie ; Seddah, Djamé (2012). Le corpus Sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2012*, 2, TALN, 321-334.
- Cantera, Jesús ; Ramón Trives, Francisco ; Heras-Diez, Florentino (1998). *Diccionario francés-español de falsos amigos*. Alicante : Universidad de Alicante.
- Carnet, Didier (2000). L'adjectif composé en anglais médical : tentative de classification. In Monique Mémet et Michel Petit (Éd.), *L'anglais de spécialité en France*. Bordeaux : GERAS Éditeur, 85-103.
- Carver, M.O.H. (1990). Pre-Viking trafic in the North Sea. In S. McGrail (Éd.) : *Maritime Celts, Frisians and Saxons*. CBA Research Report (71), 117-125.
- Cassen, Bernard (Coord.) (2008). *La bataille des langues*. Manière de voir ; févr./mars 2008, (97). Paris : Le Monde diplomatique.
- Castagne, Éric (2007A). Transparences lexicales entre langues voisines. In Éric Castagne (Éd.) : *Les enjeux de l'intercompréhension*. Colloque ICE 2. Reims : Épure, 155-166. Consulté le 30/06/17 sur <http://logatome.eu/publicat/Gap2005.pdf>
- Castagne, Éric (2007B). L'intercompréhension : un concept qui demande une approche multidimensionnelle. In *Actes du Colloque Diálogos em Intercompreensão*, Lisbonne, 6, 7 et 8 septembre 2007, 461-473.
- Castagne, Éric (2008). Les langues anglaises et françaises : amies ou ennemies ?, In *Études de linguistique appliquée*, (49). Paris : Didier Érudition, 31-42.
- Castellotti, Véronique & Moore, Danielle (2002). *Représentations sociales des langues et enseignements. Guide pour l'élaboration des politiques éducatives en Europe – De la diversité linguistique à l'éducation plurilingue*. Division des Politiques Linguistiques : DGIV : Direction de l'éducation scolaire, extrascolaire et de l'enseignement supérieur. Strasbourg : Conseil de l'Europe.
- Caure, Mélisandre (2009). *Caractérisation de la transparence lexicale, extension de la notion par ajustements graphophonologiques et microsémantiques, et application aux lexiques de l'anglais, de l'allemand et du néerlandais*, 412p. Thèse : Sciences du langage : Université Reims Champagne-Ardenne.
- Cavalla, Cristelle ; Loiseau, Mathieu ; Diwersy, Sascha ; Lascombe, Valérie ; Socha, Joanna (2013). EmoProf. *Journées Lig-Lidilem*, Jul 2013, Eybens (Grenoble), France. <[hal-01099027](http://hal-01099027)>
- Cavalli-Sforza, Luigi Luca. (2000). La génétique au service de l'histoire des populations, In *Sciences Humaines*, (109), 38-41.
- Cenoz, Jasone, 2001. The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In Jasone Cenoz, Britta Hufeisen & Ulrike Jessner (Éd.) : *Crosslinguistic Influence in Third Language Acquisition : Psycholinguistic Perspectives*. Clevedon : Multilingual Matters, 8-20.
- Charaudeau, Patrick (1992). *Grammaire du sens et de l'expression*. Paris : Hachette.
- Chavanes, Jacques (2014). *Les similitudes et les différences entre le français et l'arabe*. Consulté le 30/06/17 sur [http://casnav.ac-orleans-tours.fr/fileadmin/user\\_upload/casnav/Fran%C3%A7aisarabe.pdf](http://casnav.ac-orleans-tours.fr/fileadmin/user_upload/casnav/Fran%C3%A7aisarabe.pdf)



- Chia Kuo Ku, Ivan (2009). *Language contact in Viking Age in England. A sociolinguistic perspective*. Master de culture nordique viking et médiévale. Oslo : University of Oslo. Consulté le 30/06/17 sur [https://www.duo.uio.no/bitstream/handle/10852/26692/KU\\_MASTER.pdf?sequence=1](https://www.duo.uio.no/bitstream/handle/10852/26692/KU_MASTER.pdf?sequence=1)
- Clyne, Michael (1992). *Pluricentric languages : different norms in different nations*. Berlin ; New-York : Mouton de Gruyter.
- Codex Argenteus*. Consulté le 30/06/17 sur [www.alvin-portal.org](http://www.alvin-portal.org) : <http://www.alvin-portal.org/alvin/view.jsf?pid=alvin-record%3A60279&dswid=-1732> et <https://wikisource.org> : [https://wikisource.org/wiki/Bible,\\_Gothic,\\_Ulfila#II](https://wikisource.org/wiki/Bible,_Gothic,_Ulfila#II)
- Corvez, Pol (2011). *Étymologie du lexique biomédical et de la santé*. Paris : Ellipses.
- Coşeriu, Eugeniu (1970). *Einführung in die strukturelle Betrachtung des Wortschatzes*. Tübingen : TBL.
- Coxhead, Averil (2002). The Academic Word List : A corpus-based word list for academic purposes. In : Bernard Ketteman and Georg Marko (Ed.) *Teaching and Language Corpora (TALC) 2000 Conference Proceedings*. Rodopi, Atlanta.
- Crépin, André (1994). *Deux mille ans de langue anglaise*. Paris : Nathan.
- Cressman, Edmund D. (1926). What is the English Vocabulary ?. In *The Classical Journal*. 21(9), 671-677.
- Crubézy, Éric, Ludes, Bertrand & Keyser, Christine (2010). Indo-Européens et anthropologie biologique, In Jean-Paul Demoule (Éd.) : *Les Indo-Européens. Dossiers d'Archéologie*, (338), 62-67.
- Cuet, Christine (2001). Enseigner le français en Chine, méthodologies nouvelles, perspectives. In *Synergies Chine*, 2011 (6), 95-104.
- Dabène, Louise (Éd.) (1994). *Recherches sur l'intercompréhension entre locuteurs de langues romanes : programme européen GALATEA*. Grenoble : Université Stendhal Grenoble 3, Lidilem.
- Dabène, Louise (Éd.), (2003). Galatea [Ressource électronique]: méthode d'entraînement à la compréhension des langues romanes. Italien/ L'équipe de chercheurs du Centre didactique des Langues (Laboratoire LIDILEM) de l'Université Stendhal Grenoble 3. Chambéry : Génération 5.
- De Carlo M. (coord.). (2015). *Deux Référentiels de compétences en intercompréhension, Miriadi*. Lyon : Centre de Recherche en Terminologie et Traduction, Université Lyon 2.
- Degache, Christian (1997). Développer l'intercompréhension dans l'espace linguistique roman : le programme Galatea / Socrates. Assises de l'enseignement du et en français, séminaire de Lyon, Aupelf-Uref, 23-25 septembre 1997, 1-8.
- Degache, Christian (2006). *Didactique du plurilinguisme, Travaux sur l'intercompréhension et l'utilisation des technologies pour l'apprentissage des langues*, Mémoire d'habilitation à diriger des recherches, Université Stendhal – Grenoble 3, 232 p.
- Degache, Christian (2009). Nouvelles perspectives pour l'intercompréhension (Afrique de l'Ouest et Caraïbe) et évolutions du concept. In M.H. Araújo e Sá, R. Hidalgo Downing, S. Melo-Pfeifer, A. Séré & Vela Delfa, C. (Éd.). *A intercompreensão em línguas românicas: conceitos, práticas, formação*. Aveiro: Oficina Digital, 81–102.

- Delmonte, Violeta & Fernández Lagunilla, Marina (1987). *Sintaxis de las lenguas románicas*. Madrid : Éd. El Arquero.
- Demoule, Jean-Paul (2014). *Mais où sont passés les Indo-Européens ? Le mythe d'origine de l'Occident*. Paris : Seuil.
- Dinu, Claudia Elena, Crețu, Iona, Colibaba, Anca (2016). Multipurpose materials design for multilingual healthcare communication : the 'Take Care' approach. In Marinescu & Mureșan (Éd.): *The Impact of Internationalisation on Professional Communication and Academic Research, SYNERGY, 12(1)/2016*. Consulté le 30/06/17 sur <http://www.synergy.ase.ro/issues/2016-vol12-no-1/9-Dinu-Cretu-Colibaba.pdf>
- Do, Thi-Ngoc-Diep (2011). *Extraction de corpus parallèle pour la traduction automatique depuis et vers une langue peu dotée*. Autre [cs.OH]. Université de Grenoble, 2011. Français. Consulté le 30/06/17 sur <https://tel.archives-ouvertes.fr/tel-00680046>
- Drinka, Bridget (2017). *Language contact in Europe. The periphrastic perfect through history*. Cambridge : Cambridge University Press.
- Drouin, Patrick (2007). Identification automatique du lexique scientifique transdisciplinaire. *Revue française de linguistique appliquée* 2007/2, vol. XII, p. 45-64.
- Dumănescu, Luminița ; Mârza, Daniela & Eppel Marius (Éd.) (2014). *Intermarriage throughout History*. Newcastle upon Tyne : Cambridge Scholars Publishing.
- Ejstrud, Bo ; Hunnincke, Tomas A. ; Husum, Christine ; Korre, Aristeia ; Maarleveld, Thijs J. ; Vafeiadou, Konstantina (2008). *The Migration Period, Southern Denmark and the North Sea. Esbjerg : Maritime Archeology Programme*. Consulté le 30/06/17 sur <https://fr.scribd.com/document/14806111/The-Migration-Period-southern-Denmark-and-the-North-Sea>
- Éloy, J.-M. (2004). Des langues collatérales : problèmes et propositions. In Jean-Michel Éloy (Éd.) *Des langues collatérales*, Vol. I. Paris : L'Harmattan.
- Eppele, John W. ; Lewis, M. Paul ; Regmi, Dan Raj ; Yadava, Yogendra P. (2012). *Linguistic survey of Nepal (LinSuN)*. Kathamandu, Nepal : Tribhuvan University.
- Escudé et Janin, *Le point sur l'intercompréhension*, CLE international, 2010.
- Falaise, Achille ; Tutin, Agnès (2016). *Mode d'emploi de l'interface guidé ScienText*. Consulté le 30/06/17 sur [http://corpora.aiakide.net/scientext18/static/doc/mde\\_sciquest\\_2016.pdf](http://corpora.aiakide.net/scientext18/static/doc/mde_sciquest_2016.pdf)
- Falaise Achille, Tutin Agnès, Kraif Olivier (2011). Une interface pour l'exploitation de corpus arborés par des non-informaticiens : la plate-forme ScienQuest du projet ScienText, *TAL*. 52 (3)/2011, 103-128.
- Pascaline Faure (2010). Des discours de la médecine multiples et variés à la langue médicale unique et universelle, *ASp*, (58)/2010, 73-86.
- Ferguson, Charles (1959). Diglossia. *Word*, 15, 325-340.
- Ferrand, Ludovic (1999). 640 homophones et leurs caractéristiques. In : *L'Année psychologique*, 99(4), 687-708. Consulté le 30/06/17 sur [http://www.persee.fr/doc/psy\\_0003-5033\\_1999\\_num\\_99\\_4\\_28503](http://www.persee.fr/doc/psy_0003-5033_1999_num_99_4_28503)
- Forlot, Gilles (Éd.) (2009). *L'anglais et le plurilinguisme. Pour une didactique des contacts et passerelles linguistiques*. Paris : L'Harmattan.

- Fries, Charles Carpenter (1945). *Teaching and Learning English as a foreign language*. Ann Arbor : University of Michigan Press.
- Fung, Pascale ; Cheung, Percy (2004). Multi-level bootstrapping for extracting parallel sentences from a quasi-comparable corpus. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING)*, p. 1051-1057, Genève.
- Gadet, Françoise & Ludwig, Ralph (2014). *Le français au contact d'autres langues*. Paris : Ophrys.
- Galsterer, Hartmut (Bonn), "Coloniae" (2006). In: *Der Neue Pauly*, Herausgegeben von: Hubert Cancik, Helmuth Schneider (Antike), Manfred Landfester (Rezeptions- und Wissenschaftsgeschichte). Consulté le 14/07/17 sur [http://dx.doi.org/10.1163/1574-9347\\_dnp\\_e1002720](http://dx.doi.org/10.1163/1574-9347_dnp_e1002720)
- Garbarino, Sandra (2015). Intercomprensione, informatica, interazione, internazionalizzazione, interdisciplinarietà: Miriadi di « i » per la formazione al plurilinguismo. In Doina Spiță, Mihaela Lupu, Dana Nica, Iulia Nica (Éd.). *Les approches plurielles dans l'éducation aux langues : L'intercompréhension, en présence et en ligne*, Juin 2013, Iasi, Roumanie. Université Alexandru Ioan Cuza, pp.252.
- Gebühr, Michael (2002). Hügelgrab und Nydamboot – Die Ur- und Frühgeschichte Schleswig-Holsteins. In Jann Markus Witt und Heiko Vosgerau (Éd.), *Schleswig-Holstein von den Ursprüngen bis zur Gegenwart – Eine Landesgeschichte*, 11-62. Hamburg : Convent-Verlag.
- Geckeler, Horst & Kattenbusch, Dieter (1992). *Einführung in die italienische Sprachwissenschaft*. Romanische Arbeitshefte, 28. Tübingen : Max Niemeyer Verlag.
- Genet, Jean-Philippe (2016). Et l'Angleterre devint européenne, *L'Histoire*, (424), juin 2016, 48-53.
- Genetti, Carol (2014). *How languages work : an introduction to language and linguistics*. Cambridge : Cambridge University Press.
- Gilles, Fabrice (2011). *Intercompréhension au collège (6<sup>ème</sup>) : trois langues romanes découvertes simultanément (espagnol italien portugais) lors de séances d'ICE*. Mémoire de master 2, Université Reims Champagne Ardenne.
- Gimbutas, Marija (1956). *The Prehistory of Eastern Europe. Part I : Mesolithic, Neolithic and Copper Age Cultures in Russia and the Baltic Area*. Cambridge, MA : Peabody Museum.
- Godard, Danièle (2003). (Éd.) *Les langues romanes : problèmes de la phrase simple*. Paris : CNRS éditions.
- González Rey, Maria Isabel (2007). *La didactique du français idiomatique*. Fernelmont : EME.
- González Rey, Maria Isabel (2008). *La didactique du français idiomatique*. Fernelmont : E.M.E.
- González Rey, Maria Isabel (2012). *Outils et méthodes d'apprentissage en phraséodidactique*. Fernelmont : EME.
- Gosme, Julien (2012). *Énumération exhaustive et détection spécifique des analogies : étude pour les modèles de langue et la traduction automatique*. Informatique et langage



- [cs.CL]. Université de Caen, 2012. Français. Consulté le 30/06/17 sur <https://tel.archives-ouvertes.fr/tel-00700559/>
- Gosselin, Laurent (2010). *Les modalités du français. La validation des représentations*. Amsterdam ; New-York (N.Y.) : Rodopi.
- Grabar, Natalia & Hamon, Thierry (2016). A Large Rated Lexicon with French Medical Words, In *Proceedings of LREC 2016*, Portorož, Slovenia. Consulté le 30/06/17 sur [http://www.lrec-conf.org/proceedings/lrec2016/pdf/557\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2016/pdf/557_Paper.pdf)
- Granger, Sylviane ; Paquot, Magali (2010). Customising a general EAP dictionary to meet learner needs. In Sophie Granger & Magali Paquot (Éd.) : *Lexicography in the 21st century: New challenges, new applications*. Proceedings of ELEX2009. Cahiers du CENTAL. Louvain-la-Neuve, Presses universitaires de Louvain, 87-96.
- Gray, Russel. D. & Atkinson, Quentin D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*, 426, 435-438.
- Greenough, James ; Kittredge, George Lyman & Potter, Simeon (1962). *Words and their ways in English speech*. Boston : Beacon press.
- Grimal, Pierre (1984). *La civilisation romaine*. Paris : Arthaud.
- Grimm, Jakob (1822). *Deutsche Grammatik*. Göttingen : Dieterich.
- Gross, Gaston (1996). *Les expressions figées en français*. Paris : Ophrys.
- Gross, Peter (2009). *Anglais médical*. Paris : Maloine.
- Grzega, Joachim (2005). The Role of English in Learning and Teaching European Intercomprehension Skills, *Journal for EuroLinguistics*, 2, 1-18
- Andreas Grünwald (2014). Same same but different? ‘Sprachenübergreifende Vernetzung des Englisch- und Spanischunterrichts zur Förderung von Mehrsprachigkeit. Zeitschrift für Interkulturellen Fremdsprachenunterricht. Didaktik und Methodik im Bereich Deutsch als Fremdsprache - 6545 Jahrgang 19, Nummer 1, April 2014, 25-42. Consulté le 30/06/17 sur <http://tujournals.ulb.tu-darmstadt.de/index.php/zif/article/view/14/756>
- Guinet, Louis (1982). *Les emprunts gallo-romans au germanique*. Paris : Klincksieck.
- Haak, W. et al. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe, In *Nature*, (522), 11 June 2015, 207-211.
- Harris, Martin B. (1976). *Romance syntax : synchronic and diachronic perspectives*. Salford : University of Salford.
- Hatier, Sylvain (2016). *Identification et analyse linguistique du lexique scientifique transdisciplinaire. Approche outillée sur un corpus d'articles de recherche en SHS*. Thèse de doctorat : sciences du langage spécialité informatique et sciences du langage. Université Grenoble Alpes.
- Havran, Heidrun/Raasch, Albert/Schmitt, Brigitte/Schulz, Ursula (1992). Verstehen, ohne lernen zu müssen? Internationale Wörter im Fremdsprachenunterricht für Erwachsene. In: *Neusprachliche Mitteilungen aus Wissenschaft und Praxis*, 45. Jahrgang, Heft 2, p. 102-108.
- Hemming, Erik ; Horst G. Klein ; Reissner, Christina (2011). *English – the Bridge to the Romance Languages*. Aachen : Shaker, Editiones EuroCom.

- Higham, Nick ; Ryan, Martin J. (2013). *The Anglo-Saxon World*. New Haven ; London : Yale University Press.
- Higham Nick (2007). *Britons in Anglo-Saxon England*. Woodbridge : The Boydell Press.
- Hufeisen, Britta & Neuner, Gerhard (2003). *Mehrsprachigkeitskonzept – Tertiärsprachenlernen – Deutsch nach Englisch*. Strasbourg : Council of Europe Publishing.
- Hufeisen, Britta (2004). L1, L2, L3, L4, Lx – alle gleich ? Linguistische, lerninterne und lernerne Faktoren in Modellen zum multiplen Spracherwerb. In Nicole Baumgarten, Claudia Böttger, Markus Motz & Julia Probst (Éd.) : *Übersetzen, Interkulturelle Kommunikation, Spracherwerb und Sprachvermittlung – das Leben mit mehreren Sprachen* : Festschrift für Juliane House zum 60. Geburtstag. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* [Online], 8(2/3), 1-13.
- Ionin, Tania, Montrul, Silvina & Santos, H. (2011). An Experimental Investigation of the Expression of Genericity in English, Spanish and Brazilian Portuguese. *Lingua* 121, pp. 963-985.
- Jackson, Kenneth Hurlstone (1953). *Language and history in early Britain : A chronological survey of the Brittonic languages, first to twelfth century A.D.* Edinburgh : University Press.
- Jamet, Marie-Christine (2007). La transparence sonore du lexique. In F. Cappucho, A.A.P. Martins, C. Degache & M. Tost (Éd.) : *Diálogos em Intercompreensão*. Lisboa, Portugal : Universidade Católica Editora, 335-355.
- Johanson, Lars (2002). Do languages die of 'structuritis'? On the role of code-copying in language endangerment. *Rivista di Linguistica*, 14(2), 249-270.
- Jones, Sir William (1799). *The Works of Sir William Jones in six volumes. I*, Londres : Robinson and Evans.
- Jucquois Guy (2009). La linguistique génétique revisitée : la controverse du proto-mondial. *La linguistique*, 45(1)/2009, 133-151. Consulté le 30/06/17 sur <http://www.cairn.info/revue-la-linguistique-2009-1-page-133.htm>, DOI : 10.3917/ling.451.0133
- Katamba, Francis (1994). *English words*. London : Routledge.
- Kauffer, Maurice (2016). Les « faux amis » : théories et typologie. In : Albrecht Jörn & René Métrich (Éd.) : *Manuel de traductologie*. Berlin/Boston : Walter de Gruyter, 333-348.
- Kellerman, Eric (1979). Giving learners a break: native language intuitions as a source of predictions about transferability. *Working Papers on Bilingualism*, 15, 1979, p. 37-57.
- Kersaudy, Georges (2001). *Langues sans frontières. À la découverte des langues de l'Europe*. Paris : Éditions Autrement.
- Kessler, Brett (1995). Computational dialectology in Irish Gaelic. In *Proceedings of the European Association for Computational Linguistics*. Dublin : ACL, 60-67.
- Klein, Horst Günter (2002). 7 Siebe. Consulté le 30/06/17 sur <http://www.eurocom.uni-frankfurt.de/siebe/7Siebe/BIN/index.htm>

- Klein, Horst Günter (2004). L'intercompréhension (eurocom), une méthode de compréhension des langues voisines. *ELA*, 4(136)/2004, 403-418. Consulté le 30/06/17 sur <https://www.cairn.info/revue-ela-2004-4-page-403.htm#re13no13>
- Klein, Horst Günter (2008). L'anglais, base possible de l'intercompréhension romane ?, *ELA*, 1(149)/2008, 119-128. Consulté le 30/06/17 sur <https://www.cairn.info/revue-ela-2008-1-page-119.htm>
- Klein, Horst Günter & Rutke, Dorothea (1997). EuroComRom: pour un plurilinguisme européen, *Sociolinguistica*, (11), 178-183.
- Klein, Horst Günter & Stegmann, Tilbert Diego (2000). *EuroComRom – Die sieben Siebe : Romanische Sprachen sofort lesen können*. Aachen : Shaker, Editiones EuroCom vol. 1.
- Klein, Horst Günter & Reissner, Christine (2006). *Basismodul Englisch. Englisch als Brückensprache in der romanischen Interkomprehension*. Editiones EuroCom, vol. 26. Aachen : Shaker.
- Koehn, Philipp (2005). Europarl : A Parallel Corpus for Statistical Machine Translation. In : MT Summit X, Phuket, Thailand, September 13-15, 2005, Conference Proceedings : the tenth Machine Translation Summit, 79-86.
- Koessler, Maxime ; Derocquigny, Jules (1928). *Les faux Amis ou les Trahisons du vocabulaire Anglais*. Paris : Librairie Vuibert.
- Kossinna, Gustaf (1902). Die indogermanische Frage archäologisch beantwortet, In *Zeitschrift für Ethnologie*, (34), 161-222.
- Krahe, (1964). *Unsere ältesten Flußnamen*. Harrassowitz Verlag : Wiesbaden.
- Kraif Olivier (2008). Comment allier la puissance du TAL et la simplicité d'utilisation ? l'exemple du concordancier bilingue ConcQuest, In *Actes des 9<sup>e</sup> Journées d'analyse statistique des données textuelles, JADT 2008*, Lyon, Presses universitaires de Lyon, 2008, 625-634. Consulté le 30/06/17 sur [http://olivier.kraif.u-grenoble3.fr/index.php?option=com\\_content&task=view&id=42](http://olivier.kraif.u-grenoble3.fr/index.php?option=com_content&task=view&id=42)
- Kraif, Olivier (2014). *Corpus parallèles, corpus comparables : quels contrastes ?* Synthèse d'Habilitation à Diriger des Recherches. Poitiers : Université de Poitiers.
- Kraif, Olivier & Diwersy, Sacha (2012). Le Lexicoscope : un outil pour l'étude de profils combinatoires et l'extraction de constructions lexico-syntaxiques. In *Actes de la conférence TALN 2012*, 399-406.
- Kraif, Olivier (2015). Multi-alignement vs bi-alignement : à plusieurs, c'est mieux !. Actes de TALN 2015. Consulté le 30/06/17 sur [http://www.atala.org/taln\\_archives/TALN/TALN-2015/taln-2015-long-022.pdf](http://www.atala.org/taln_archives/TALN/TALN-2015/taln-2015-long-022.pdf)
- Krifi-Papoz, Sonia (2010). *Traitement phonémique en lecture : lecture normale et dyslexie développementale*. Thèse de doctorat : neuropsychologie. Université Lumière Lyon 2.
- Kroeber, Alfred & Chrétien, C. Douglas (1960). Statistics, Indo-European and taxonomy. *Language*, 36, 1-21.
- Kumano, Tadashi ; Tanaka, Hideki & Tokunaga, Takenobu. (2007) Extracting phrasal alignment from comparable corpora by using joint probability SMT model. In *Proceedings of TMI*, 95-103.
- Ladmiral, Jean-René (2014). *Sourcier ou cibliste*. Paris : Édition Les Belles Lettres.

- Lambert, Pierre-Yves (1997). *La langue gauloise*. Paris : Éditions Errance.
- Lamberterie, Charles de (2011). L'indo-européen. In Emilio Bonvini, Joëlle Busuttill et Alain Peyraube (Éd.) : *Dictionnaire des langues*. Paris : Presses universitaires de France.
- Langlais, Philippe ; Yvon, François & Zweigenbaum, Pierre. (2008). Analogical translation of medical words in different languages. In : *Proceedings of the 6th international conference on Advances in Natural Language Processing*. Consulté le 30/06/17 sur [https://perso.limsi.fr/pz/FTPapiers/Langlais\\_GOTAL2008.pdf](https://perso.limsi.fr/pz/FTPapiers/Langlais_GOTAL2008.pdf)
- Lass, Roger (1987). *The shape of English*. London : J.M. Dent & Sons.
- Lass, Roger (1995). *Old English : a historical linguistic companion*. Cambridge : Cambridge University Press.
- Lautier, Nicole ; Allieu-Mary, Nicole (2008). La didactique de l'histoire. *Acquisitions et progressions scolaires : recherches en psychologie*. Revue française de pédagogie, (162), janvier-mars 2008, 95-131.
- Lavallée, Jean-François & Langlais, Philippe. (2010). Unsupervised morphological analysis by formal analogy. In : Carol Peters *et al.* (Éd.), *Multilingual Information Access Evaluation I. Text Retrieval Experiments. CLEF 2009. Lecture Notes in Computer Science, 6241*. Springer, Berlin, Heidelberg, 617-624.
- Le Bohec, Yves (2012). *Naissance, vie et mort de l'Empire romain : de la fin du I<sup>er</sup> siècle avant notre ère jusqu'au V<sup>e</sup> siècle de notre ère*. Paris : Picard.
- Le Goff, Jacques (2008). *La civilisation de l'Occident médiéval*. Paris : Flammarion.
- Lebedynsky, Iaroslav (2006). *Les Indo-Européens. Faits, débats, solutions*. Paris : Éditions Errance.
- LeBlanc, Raymond ; Séguin, Hubert (1987). Les homographes et les parographes anglais-français dans l'enseignement de la langue seconde. *Actes du colloque AILA '87*, Sydney.
- Lee, Donald Woodward (1948). *Functional change in early English*. Menasha : Wis., G. Banta.
- Legendre, Jean-Pierre (2016). La nostalgie des « Grandes Invasions ». L'archéologie face au pangermanisme et au nazisme. In Isabelle Bardiès-Fronty (Éd.), *Les Temps des Mérovingiens. Exposition Musée de Cluny*. Dossier d'Archéologie, hors série n°31, octobre 2016.
- Lelièvre, Dominique (2004). *Voyageurs chinois à la découverte du monde. De l'Antiquité au XIX<sup>e</sup> siècle*. Genève : Éditions Olizane.
- Lepage, Yves (2003). *De l'analogie rendant compte de la commutation en linguistique*. Mémoire d'habilitation à diriger des recherches. Interface homme-machine [cs.HC]. Université Joseph-Fourier - Grenoble I, 2003.<tel-00004372>. Consulté le 30/06/17 sur <https://halshs.archives-ouvertes.fr/tel-00004372/document>
- Levenstein, Vladimir Iossifovitch (1965). Dvoitchnye kody s ispravleniem vypadenij, vstavok i zamechtchenij simvolov. *Doklady Akademij Nauk URSS*. 163 (4): 845–8.
- Lieberman, Anatoly (2005). *Word origins : and how we know them : etymology for everyone*. New York : Oxford University Press.
- Mallory, James Patrick (1997). *À la recherche des Indo-Européens*. Paris : Seuil.

- Marchello-Nizia, Christiane (2011). Le français. In Emilio Bonvini, Joëlle Busuttill et Alain Peyraube (Éd.) : *Dictionnaire des langues*. Paris : Presses universitaires de France.
- Martin Kostomaroff, Éric (2010). Galanet : un didacticiel d'interaction et de réception en langues romanes. In Doina Spiță & Claudia Tărnăuceanu (Éd.), « Galapro » sau Despre intercomprehensiune în limbi romanice. Actes du colloque Formare de formatori pentru intercomprehensiune în limbi romanice, 22-24 octobre 2009. Iași : Editura Universității Al. I. Cuza.
- Meillet, Antoine (1908). *Les dialectes indo-européens*. Paris : H. Champion.
- Meißner, Franz-Joseph (1998). Transfer beim Erwerb einer weiteren romanischen Fremdsprache : das mehrsprachige mentale Lexikon. In Franz-Joseph Meißner & M. Reinfried : *Mehrsprachigkeitsdidaktik. Konzepte, Analysen, Lehrerfahrungen mit romanischen Fremdsprachen*. Tübingen : Narr, 45-69.
- Meißner, Franz-Joseph (2000). Zwichensprachliche Netzwerke. Mehrsprachigkeitsdidaktische Überlegungen zur Wortschatzarbeit, *Französisch heute*, 31(1), 55-67.
- Meißner, Franz-Joseph (2004). Transfer und Transferieren. Anleitungen zum Interkomprehensionsunterricht. In Horst Günter Klein & Rutke, Dorothea (Éd.), *Neuere Forschungen zur Europäischen Interkomprehension*. Aachen : Shaker.
- Meißner, Franz-Joseph (2016). L'approche intercompréhensive expliquée en 24 points. *Le français à l'université.*, 21-03 | 2016. Mise en ligne le: 07 octobre 2016, consulté le : 13 juillet 2017.
- Mendelsohn, P (1996). Le concept de transfert. In Philippe Meyrieu, Michel Develay, Christiane Durand, Yves Mariani (Ed.) *Le concept de transfert de connaissance en formation initiale et en formation continue*, Lyon, CRDP, 11-20.
- Mills, Anthony David (2004). *A Dictionary of British Place-names*. Oxford : Oxford University Press.
- Mira Mateus, Helena, Villalva, Alina (2006). *Essential sobre linguística*. Lisboa : Caminho.
- Morris, William (1981). *The American heritage dictionary of the English language*. Boston : Houghton Mifflin, cop. 1981.
- Mossé, Fernand (1947). *Esquisse d'une histoire de la langue anglaise*. Lyon : IAC.
- Mullen, Alex (2016). Sociolinguistics, In Martin Millett, Louise Revell et Alison Moore (Éd.), *The Oxford Handbook of Roman Britain*. Oxford : Oxford University Press.
- Mullen, Alex (2013). New thoughts on British Latin : a curse tablet from Red Hill, Ratcliffe-on-Soar (Nottinghamshire). *Zeitschrift für Papyrologie und Epigraphik*, Band 187.
- Muller, Claude (2002). *Les bases de la syntaxe : syntaxe contrastive, français – langues voisines*. Pessac (Gironde) : Presses universitaires de Bordeaux.
- Muftah, Mohamed (2014). *Représentations et pratiques du contact des langues en classe de FLE : place et rôle de l'anglais et de l'arabe dans l'enseignement et apprentissage du français à l'université en Libye : étude à partir d'observations et d'entretiens*. Thèse de doctorat en sciences du langage soutenue à l'Université de Lorraine le 12/12/14.
- Murillo Puyal, J., Piccaluga, M. Harmegnies, B. (2008). Pour un français partenaire dans l'aire romane : aspects psycholinguistiques et didactiques de l'intercompréhension entre locuteurs de langues latines. In V. Conti & F. Grin (Ed.), *S'entendre entre langues voisines : vers l'intercompréhension*, 341-370. Genève : Georg Éditeur.

- Nerbonne, John ; Heeringa, Wilbert ; Kleiweg, Peter (1999). Edit distance and dialect proximity. In David Sankoff & Joseph Kruskal (Éd.) : *Time Warps, String Edits and Macromolecules : the Theory and Practice of Sequence Comparison*. Stanford : CSLI, 5-15.
- Neuner, Gerhard (2003). Le concept de plurilinguisme et la didactique de langue tertiaire. In Britta, Hufeisen & Gerhard, Neuner (Éd.) : *Le concept de plurilinguisme – Apprentissage d'une langue tertiaire, l'allemand après l'anglais*. Strasbourg : Éditions du Conseil de l'Europe. Consulté le 30/06/17 sur <http://archive.ecml.at/documents/pub112F2004HufeisenNeuner.pdf>
- Nicolaï, Robert (2010). De l'arbre généalogique à la saisie du contact des langues, *Dossiers d'archéologie* (338), mars/ avril 2010, 68-71.
- Nielsen, Hans Frede (1981). *Old English anf the continental Germanic languages : a survey of morphological and phonological interrelations*. Innsbruck : Institut für Sprachwissenschaft der Universität Innsbruck.
- Niobey, Georges ; Galiana, Thomas de ; Jouannon, Guy ; Lagane, René (1980). *Nouveau dictionnaire analogique*. Paris : Larousse.
- Ó Laoire, Muiris, Singleton, David (2006). Psychotypologie et facteur L2 dans l'influence translexicale. Une analyse de l'influence de l'anglais et de l'irlandais sur le français L3 de l'apprenant. In : Rebekah Rast, Pascale Trévisiol (Éd.), *L'acquisition d'une langue 3, Acquisition et Interaction en Langue Étrangère. AILE*, (24), 101-117. Paris : Encrages.
- Ó Laoire, Muiris, Singleton, David (2009). The role of prior knowledge in L3 learning and use. Further evidence of psychotypological dimensions. In Larissa Aronin & Britta Hufeisen (Éd.), *The exploration of multilingualism : development of research on L3, multilingualism and multiple language acquisition*, AILA Applied Linguistics Series, 6, 79-102.
- Ó Muirithe, Diarmaid (2010). *From the Viking word-hoard : a dictionary of Scandinavian words in the languages of Britain and Ireland*. Dublin ; Portland, Or. : Four Courts Press.
- Ollivier, Christian, Strasser, Margareta (2013). *Interkomprehension in Theorie und Praxis*. Wien : Praesens Verlag.
- Pakendorf, Brigitte (2014). La génétique appliquée aux langues, In *Dossier Pour la Science*, (82), janvier-mars 2014, 46-51.
- Péchoin, Daniel (1999). *Thesaurus : des idées aux mots, des mots aux idées*. Paris : Larousse.
- Pecman, Mojca (2005). Les apports possibles de la phraséologie à la didactique des langues étrangères. *Apprentissage des Langues et Systèmes d'Information et de la Communication*, 8 (I), 109-122. Consulté le 30/06/17 sur <https://alsic.revues.org/334>
- Perrenoud, Philippe (1997). Vers des pratiques pédagogiques favorisant le transfert des acquis scolaires hors de l'école. *Pédagogie collégiale* (Québec), 10(3)/1997, 5-16. Consulté le 30/06/17 sur [http://www.unige.ch/fapse/SSE/teachers/perrenoud/php\\_main/php\\_1997/1997\\_04.html](http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_1997/1997_04.html)
- Perret, Michèle (1999). *Introduction à la langue française*. Paris : SEDES.
- Phal, André (1971). *Vocabulaire général d'orientation scientifique (VGOS). Part du lexique commun dans l'expression scientifique*. Paris : Didier.



- Potter, Timothy W. & Johns, Catherine (1992). *Roman Britain*. London : British Museum
- Pyles, Thomas & Algeo, John (1970). *English : an introduction to language*. New York ; Chicago ; San Francisco : Harcourt, Brace & World.
- Pyles, Thomas & Algeo, John (2010). *The origins and development of the English language*. 6th edition. Boston, MA : Wadsworth Cengage Learning.
- Rastier, François (2002). *Enjeux épistémologiques de la linguistique de corpus*. En ligne : [http://www.revue-texto.net/Inedits/Rastier/Rastier\\_Enjeux.html](http://www.revue-texto.net/Inedits/Rastier/Rastier_Enjeux.html)
- Reinheimer, Sanda ; Tasmowski, Liliane (1997). *Pratique des langues romanes : espagnol, français, italien, portugais, roumain*. Paris : L'Harmattan.
- Reinheimer Rîpeanu, Sanda (1993): *Structuri morfosintactice de bază în limbile romanice*. Bucarest : Universitatea București.
- Reissner, Christina (2012). *Den Sprachenunterricht vernetzen : Das Englische als Brückensprache zum Spanischen*. In Eva Leitzke-Ungerer, Gabriele Blell & Ursula Vence (Éd.), *English-Español : Vernetzung im kompetenzorientierten Spanischunterricht*. Stuttgart, 181-201.
- Renfrew, A. Colin (1987). *Archeology and Language : The Puzzle of the Indo-European Origins*. Londres : Penguin Books.
- Riemer, Claudia (2002). *Wie lernt man Sprachen ?* In Jürgen Quetz & Gerhard Von der Handt (Éd.). *Neue Sprachen lehren und lernen. Fremdsprachenunterricht in der Weiterbildung*. Bielefeld : Bertelsmann, 49-82.
- Ringbom, H. (2007). *Cross-linguistic Similarity in Foreign Language Learning*. Clevedon : Multilingual Matters.
- Robert, J.-M. (2004). *Langues voisines, langues proches et langues lointaines : implications didactiques*. In *Langues voisines et langues étrangères : Approches didactiques et sociolinguistiques*, Alliance française de Paris, 2004. Consulté le 30/06/17 sur <http://www.u-picardie.fr/LESCLaP/spip.php?article347>
- Robert, Jean-Michel (2008). *L'anglais, comme langue proche du français*. In Jean-Michel Robert & Gilles Forlot (Éd.), *L'anglais, langue passerelle vers le français ?*. ELA, (149)/1. Paris : Klincksieck.
- Robert, Jean-Michel (2009A). *Manières d'apprendre. Pour des stratégies d'apprentissage différenciées*. Paris : Hachette.
- Robert, Jean-Michel (2009B). *De l'anglais vers le français par l'intercompréhension*. In Forlot, Gilles (Éd.) *L'anglais et le plurilinguisme. Pour une didactique des contacts et des passerelles linguistiques*. Paris : L'Harmattan, 199-210.
- Ronjat, Jules (1913). *Essai de syntaxe des parlers provençaux modernes*. Thèse de Doctorat : Lettres. Université de Paris.
- Sala, Marius ; Bîrlădeanu, Mihaela ; Iliescu, M. ; Macarie, Liliana ; Nichita ; Ioana ; Ploae-Hanganu, Mariana ; Theban, Maria ; Vintilă-Rădulescu, Iona (1988). *Vocabularul reprezentativ al limbilor romanice*. Bucarest : Editura Științifică și Enciclopedică.
- Salamanca, Francisco José Zamora (1984). *La tradición histórica de la analogía lingüística*. In *Revista Española de Lingüística*, 14(2), 367-419.
- Salway, Peter Hayes (1997). *A History of Roman Britain*. Oxford : Oxford University Press.

- Sawyer, Peter Hayes (1998). *From Roman Britain to Norman England*. Londres et New-York : Routledge.
- Schleicher, August (1861). *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. Weimar : H. Böhlau.
- Schmid, Helmut (1995). Improvements in Part-of-Speech Tagging with an Application to German. *Proceedings of the ACL SIGDAT-Workshop*. Dublin, Ireland. Consulté le 17/07/17 sur <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger2.pdf>
- Schmidt, Johannes (1872). *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar : H. Böhlau.
- Schmidt, Wilhelm (1969). *Geschichte des deutschen Sprache*. Berlin : Verlag Volk und Wissen Volkseigener Verlag.
- Schrijver, Peter (2002). The rise and fall of British Latin : evidence from English and Brittonic. In Markku Pilppula, Juhani Klemola & Heli Pitkänen (Éd.) *The Celtic roots of English*. Joensuu : Univ. of Joensuu, Faculty of Humanities, 87-110.
- Schuchardt, Hugo (1884). Slawo-deutsches und Slawo-italienisches. *Dem Herrn Franz von Miklosisch zum 20. November 1883*. Graz : Leuschner & Lubensky.
- Séguin, Hubert (1994). Les congénères anglais-français. Étude de leur similarité graphique et suggestions d'exploitation pédagogique. In Claude Buridant & Jean-Christophe Pellat (Éd.) : *L'écrit en français langue étrangère, réflexions et propositions*. Strasbourg : Presses universitaires de Strasbourg, 193-199.
- Sérasset, Gilles ; Blanc, Étienne (1993). Une approche par acceptions pour les bases lexicales multilingues. Clas, André and Bouillon, Pierrette. T-TA-TAO 93, journées scientifiques du réseau Lexicologie, Terminologie, Traduction, Sept. 1993, Montréal, Canada. Actualité scientifique, AUPELF-UREF, pp.65-84, 1993. <hal-00966437>
- Sergent, Bernard (1995). *Les Indo-Européens*. Paris : Payot.
- Sériot, Patrick (1996). *N.S. Troubetzkoy. L'Europe et l'humanité. Écrits linguistiques et paralinguistiques*. Liège : Mardaga.
- Sima, Qian (DL 2015). *Les mémoires historiques de Se-Ma Ts'ien = Shiji / traduits et annotés par Édouard Chavannes, Max Kaltenmark, Jacques Pimpaneau*. Paris : Éditions You Feng Libraire & éditeur.
- Simard, Michel ; Foster, George & Isabelle, Pierre (1992). Using cognates to align sequence. In *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation*, 67-81. Montréal, Canada.
- Sinner, Carsten (2014). *Varietätenlinguistik. Eine Einführung*. Tübingen : Narr.
- Singleton, David (1999). *Exploring the Second Language Mental Lexicon*. Cambridge : Cambridge University Press.
- Sörös, Ana (2008). *Typologie et linguistique contrastive. Théories et applications dans la comparaison des langues*. Bern ; Berlin ; Bruxelles [etc.] : P. Lang.
- Stedje, Astrid (1976). Interferenz von Muttersprache und Zweitsprache auf eine dritte Sprache. *Zielsprache Deutsch*, 14, 15-21.



- Stefenelli, Arnulf (1992). Die Transferierbarkeit des lateinischen Wortschatzes beim Erwerb romanischer Sprachen. In: *französisch heute* 3, 379-387.
- Steinberger, Ralf *et al.* (2006). The JRC-Acquis : A multilingual aligned parallel corpus with 20+ languages. In *Actes de la 5<sup>e</sup> International Conference on Language Resources and Evaluation (LREC'2006)*. Gênes, 24-26 mai 2006.
- Strabon (23). *Géographie*. Consulté le 30/06/17 sur <http://remacle.org>
- Strang, Barbara M H (1970). *A history of English*. London : Methuen.
- Streabeck, Arval L. (1966). *A short introduction to Germanic linguistics*. Boulder, Colo. : Pruett Press.
- Swales, John (1990). *Genre Analysis : English in Academic and Research Settings*. Boston : Cambridge University Press.
- Sweet, Henry (1898). *A new English Grammar : logical and historical. Part II, Syntax*. Oxford : Clarendon Press.
- Tabouret-Keller, Andrée (2006). À propos de la notion de diglossie. La malencontreuse opposition entre 'haute' et 'basse' : ses sources et ses effets. *Langage & Société*, (118), 109-128. Consulté le 30/06/17 sur <http://www.cairn.info/revue-langage-et-societe-2006-4-page-109.html>
- Tagliavini, Carlo (1969). *Le origine delle lingue neolatine ; introduzione alla filologia romanza*. Bologna : R. Pàtron.
- Tasmowski, Liliane & Reinheimer, Sanda (2003). Variations dans le radical du verbe roman. In Danièle Godard (Éd.) : *Les langues romanes. Problèmes de la phrase simple*. Paris : CNRS éditions.
- Tesnière, Lucien (1959). *Éléments de syntaxe structurale*. Paris : Librairie C. Klincksieck.
- Teubert, Wolfgang (1996). Comparable or Parallel Corpora ? *International Journal of Lexigraphy*, 9 (3) : 238-264.
- Teyssier, Paul (2012). *Comprendre les langues romanes : du français... à l'espagnol, au portugais, à l'italien & au roumain : méthode d'intercompréhension*. Nouvelle édition revue. Paris : Chandeigne.
- Thody, Philip M. W. ; Evans, Howard ; Rees, Gwilym (1985). *Faux amis and key words – A dictionary-guide to French language, culture and society through lookalikes and confusables*. London : The Athlone Press.
- Tiedemann, Jörg (2009). News from OPUS – A collection of Multilingual Parallel Corpora with Tolls and Interfaces. In N. Nicolov and K. Bontcheva and G. Angelova and R. Mitkov (Éd.) *Recent Advances in Natural Language Processing, V*, 237-248, John Benjamins, Amsterdam/Philadelphia.
- Tognini-Bonelli, Elena (2001). *Corpus Linguistics at work*. Amsterdam / Philadelphia : John Benjamins Publishing Company.
- Trần, Đỗ Đạt, (2007). *Synthèse de la parole à partir du texte en langue vietnamienne*. Thèse de doctorat : Traitement de la parole. Université de Grenoble.
- Treiman, Rebacca (1993). *Beginning to spell. A study of First-Grade Children*. New York, Oxford : Oxford University Press.

- Trevelyan, George Macaulay (1949). *Gechichte Engels. 2. Band, von 1603 bis 1918*. Munich : Leibniz Verlag.
- Troubetzkoï, Nikolai S. (1923). *Vavilonskaja bašnja i smešenie jazykov, Evrazijskij vremennik*, 3. Paris : Evrazijskoe Knigoizdatel'stvo.
- Tutin, Agnès ; Grossmann, Francis (2013). *L'écrit scientifique : du lexique au discours*. Rennes : Presses Universitaires de Rennes.
- Tutin, Agnès (2007). Autour du lexique et de la phraséologie des écrits scientifiques. *Revue française de linguistique appliquée*, vol. XXI(2), 55.
- Tutin, Agnès (2008). L'apport des corpus annotés pour l'élaboration semi-automatique d'une base de collocations de la langue scientifique générale. In François Maniez, Pascaline Dury, Nathalie Arlin, and Claire Rougemont (Éd.) *Corpus et dictionnaires de langues de spécialité*, 45–65. Travaux du CRTT. Grenoble : Presses Universitaires de Grenoble.
- Valkhoff, Marius François (1932). *Latijn, romaans, roemeens*. Amersfoort : Valkhoff.
- Vennemann, Theo (2003). *Europa Vasconica – Europa Semitica*. Berlin, New-York : Mouton de Gruyter.
- Victorri, Bernard (2008). Langues et langage (origine des), In *Encyclopædia Universalis*. Paris : Encyclopædia Universalis.
- Viel, Michel (1981). *La phonétique de l'anglais*. Paris : Presses universitaires de France.
- Virgile (98). *De vita et moribus Iulii Agricolaë*. Consulté le 30/06/17 sur [https://la.wikisource.org/wiki/De\\_vita\\_et\\_moribus\\_Iulii\\_Agricolaë](https://la.wikisource.org/wiki/De_vita_et_moribus_Iulii_Agricolaë)
- Walter, Henriette (1994). *L'aventure des langues en Occident*. Paris : Laffont.
- Walter, Henriette (1988). *Le français dans tous les sens*. Paris : Robert Laffont.
- Walter, Henriette (2001). *Honni soit qui mal y pense*. Paris : Robert Laffont.
- Walter, Henriette (2003). Le latin des naturalistes, annonce d'un bilinguisme bien particulier, *La Linguistique, Plurilinguismes*, 41, fasc. 2.
- Walter, Henriette (2014). *Minus, lapsus et mordicus. Nous parlons tous latin sans le savoir*. Paris : Robert Laffont.
- Walter, Heribert (1991): Spanisch als Sprungbrett für Leseverständnis Portugiesisch. In: *Neusprachliche Mitteilungen* 44, 155-159.
- Wang Szilas, Jue (2016). *Les enjeux de l'intégration de l'eTandem en didactique des langues-cultures étrangères : interactions entre apprenants et dynamique institutionnelle dans un dispositif universitaire sino-francophone*. Thèse de doctorat : Université de Genève.
- Wartburg, Walther von (1950). *Die Ausgliederung der Romanischen Sprachräume*. Berne : A. Francke.
- Watkins, Calvert (2000). Indo-European and the Indo-Europeans. In *The American Heritage dictionary of the English language*. 4th edition. Boston ; New-York : Houghton Mifflin Company, 2000, 1598-1603.
- Williams, Geoffrey C. (1999). Les réseaux collocationnels dans la construction et l'exploitation d'un corpus dans le cadre d'une communauté de discours scientifique. Thèse d'anglais : Nantes, Université de Nantes.

- Williams, Geoffrey C. & Millon, C. (2009). The General and the Specific ; Collocational resonance of scientific language. In *Proceedings of the Corpus Linguistics Conference CL2009*, M. Mahlberg, V. Gonzalez-Diaz & C. Smith (Éd.). Liverpool : University of Liverpool.
- Winkler, William E. (1990). *String comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage*. Washington : Educational Resources Information Center (ERIC). Consulté le 30/06/17 sur <http://files.eric.ed.gov/fulltext/ED325505.pdf>
- Wulff, Henrik R. (2004). The language of Medicine. *Journal of the Royal Society of Medicine*, 97.
- Zwanenburg, Wiechert (1983). *Productivité morphologique et emprunt. Étude des dérivés déverbaux savants en français moderne*. Amsterdam ; Philadelphia : John Benjamins Publishing Company.

# Annexes

Les annexes qui figurent dans la version papier se composent d'extraits des données car l'ensemble des données est trop volumineux. Le format portrait permet de faire figurer au mieux les données de colonnes multiples. Les fichiers complets seront accessibles en ligne sur le portail HAL dans les documents de l'auteur de cette thèse.

## Annexe 1 : les 2000 lemmes anglais fréquents du corpus ScienText

abdominal, ability, able, abnormal, abnormality, absence, absent, absolute, abundance, academic, accept, access, accompany, account, account, accumulate, accumulation, accuracy, accurate, achieve, acid, acid, acquire, acquisition, act, act, actin, action, activate, activation, activator, active, activity, actual, acute, adapt, adaptation, add, addition, additional, address, adenocarcinoma, adequate, adhesion, adjacent, adjust, adjustment, administer, administration, admission, admit, adult, adult, advance, advanced, advantage, adverse, affect, affected, affinity, African, age, aged, agent, agonist, agree, agreement, aim, aim, air, airway, albumin, alcohol, algorithm, align, alignment, allele, allergy, allow, alpha, alter, alteration, altered, alternative, alternative, alveolar, American, amino, acid, amount, amplification, amplify, analogue, analysis, analytic, analyze, androgen, angiogenesis, animal, animal, annotate, annotation, annual, antagonist, anterior, antibiotic, antibody, antigen, antioxidant, antisense, anxiety, apoptosis, apoptotic, apparent, appear, appearance, application, apply, approach, appropriate, approve, area, arise, arm, array, arrest, arterial, artery,

arthritis, article, ask, aspect, aspirin, assay, assay, assembly, assess, assessment, assign, assist, assistance, associate, association, assume, assumption, asthma, atherosclerosis, atrial, attempt, attempt, attention, attitude, attribute, author, autoantibody, autoimmune, automated, availability, available, average, average, avoid, axis, bacillus, background, bacterial, bacterium, bad, balance, band, barrier, basal, base, base, baseline, basic, basis, be, bear, become, begin, behavior, behavioral, believe, belong, beneficial, benefit, beta, bias, bind, binding, biochemical, biological, biology, biopsy, biosynthesis, birth, black, bladder, BLAST, block, block, blood, blot, blue, body, bone, boundary, bovine, box, brain, branch, breast, British, broad, brown, buffer, buffer, build, burden, *Caenorhabditis elegans*, calcium, calculate, calculation, call, cAMP, can, Canadian, cancer, candidate, capable, capacity, capture, carbon, carcinoma, cardiac, cardiovascular, care, carrier, carry, cartilage, case, catalytic, category, catheter, cause, cause, cDNA, cell, cellular, center, central, centrifugation, cerebral, certain, cervical, chain, challenge, chamber, change, change, channel, characteristic, characteristic, characterization, characterize, check, chemical, chemokine, chemotherapy, chest, child, childhood, chip, choice, cholesterol, choose, chromatin, chromosomal, chromosome, chronic, cigarette, circulate, circulation, city, class, classification, classify, clear, clearance, cleavage, clinic, clinical, clinician, clone, clone, close, cluster, cluster, cm, code, code, coding, codon, coefficient, cognitive, cohort, cold, collagen, colleague, collect, collection, college, colon, colony, color, colorectal, column, combination, combine, combined, come, comment, committee, common, communication, community, comparable, comparative, compare, comparison, compartment, compete, complement, complementary, complete, complete, complex, complex, complexity, compliance, complication, component, compose, composition, compound, comprehensive, comprise, computational, compute, computer, conceive, concentration, concept, concern, concern, conclude, conclusion, condition, conduct, confidence, confirm, conformation, conjugate, connect, consecutive, consensus, consent, consequence, conservation, conserve, consider, considerable, consideration, consist, consistent, constant, constitute, constraint, construct, construct, construction, consumption, contact, contain, content, context, continue, continuous, contraction, contrast, contribute, contribution, control, control, controlled, conventional, convert, coordinate, coordination, copy, cord, core, coronary, correct, correct, correction, correlate, correlate, correlation, correspond, corresponding, cortex, cost, count, count, country, couple, course, cover, coverage, create, criterion, critical, cross, cross, crucial, culture, culture, cultured, current, curve, cycle, cyclin,

cysteine, cytochrome, cytokine, cytoplasm, cytoplasmic, cytotoxic, daily, damage, database, dataset, datum, day, death, decade, decision, declare, decline, decrease, decrease, defect, defense, deficiency, deficient, define, definition, degradation, degree, dehydrogenase, delay, deletion, deliver, delivery, demographic, demonstrate, dendritic, density, department, depend, dependence, dependent, depression, derive, describe, description, design, design, detail, detailed, detect, detectable, detection, determinant, determination, determine, develop, development, developmental, deviation, device, diabetes, diabetic, diagnose, diagnosis, diagnostic, diameter, die, diet, dietary, differ, difference, different, differential, differentiate, differentiation, difficult, difficulty, digest, dilute, dilution, dimension, direct, direct, direction, disability, discharge, discover, discovery, discuss, discussion, disease, disorder, display, display, disruption, distal, distance, distinct, distinguish, distribute, distribution, district, divergence, diverse, diversity, divide, division, DNA, do, doctor, document, dog, dollar, domain, dominant, donor, dopamine, dose, double, downstream, draft, draw, drive, drosophila, drug, duplication, duration, dye, dynamic, dynamics, dysfunction, early, easy, economic, edge, edit, education, educational, effect, effective, effectiveness, effector, efficacy, efficiency, efficient, effort, egg, elderly, electron, electrophoresis, element, elevated, elevation, elicit, eliminate, embryo, embryonic, emergency, empirical, employ, enable, encode, end, endogenous, endometrial, endothelial, energy, enhance, enhanced, enrich, ensure, enter, entire, entry, envelope, environment, environmental, enzyme, epidemiological, epidemiology, episode, epithelial, epithelium, epitope, equal, equation, equivalent, error, Escherichia coli, essential, EST, establish, estimate, estimate, estimation, estrogen, ethanol, ethic, ethical, eukaryote, eukaryotic, European, evaluate, evaluation, event, evidence, evident, evolution, evolutionary, evolve, exact, examination, examine, example, exception, exchange, exclude, exclusion, exercise, exhibit, exist, existence, existing, exon, expand, expansion, expect, experience, experience, experiment, experimental, expert, explain, explanation, explore, expose, exposure, express, expression, extend, extension, extensive, extent, external, extracellular, extract, extract, extraction, eye, facilitate, facility, fact, factor, fail, failure, fall, false, family, fast, fat, fatigue, fatty, feature, feed, feedback, feel, female, female, fetal, fever, fiber, fibroblast, fibrosis, field, fig, figure, file, filter, final, financial, find, finding, first, fish, fit, fix, fixed, flow, fluid, fluid, fluorescence, fluorescent, focus, focus, follicle, follow, follow-up, following, food, force, form, form, format, formation, foundation, fraction, fracture, fragment, frame, framework, free, frequency, frequent, full, full-length, function, function,

functional, further, fusion, future, gain, gamma, gap, gas, gastric, gastrointestinal, gel, gender, gene, general, generate, generation, genetic, genetics, genome, genomic, genomics, genotype, geographical, get, give, gland, glass, global, glucose, glutamate, glycoprotein, go, goal, good, government, grade, gradient, grant, graph, great, green, group, grow, growth, guideline, half, hand, haplotype, have, head, health, healthy, heart, heat, height, helix, help, help, heparin, hepatic, hepatitis, heterogeneity, high, highlight, hip, histological, histone, history, hit, HIV, hold, home, homologous, homologue, homology, hormone, hospital, hospitalization, host, hour, household, human, human, hybrid, hybridization, hybridize, hypertension, hypothesis, hypothesize, hypoxia, idea, identical, identification, identify, identity, ill, illness, illustrate, image, imaging, immune, immunity, immunization, immunodeficiency, immunoglobulin, immunohistochemical, impact, impair, impairment, implement, implementation, implicate, implication, imply, importance, important, improve, improved, improvement, in vitro, in vivo, inactivation, incidence, include, inclusion, income, incorporate, increase, increase, increased, incubate, incubation, independent, index, indicate, indication, indicator, indirect, individual, individual, induce, inducible, induction, infant, infarction, infect, infected, infection, infectious, infer, inflammation, inflammatory, influence, influence, inform, information, infusion, inhibit, inhibition, inhibitive, inhibitor, initial, initiate, initiation, inject, injection, injury, input, insect, insert, insertion, insight, institute, institution, instruction, instrument, insulin, intact, intake, integrate, integration, integrin, intensity, intensive, interact, interaction, interest, interesting, interface, interferon, interleukin, intermediate, intern, internal, international, internet, interpret, interpretation, interval, intervention, interview, intestinal, intracellular, intravenous, introduce, introduction, intron, invasion, invasive, invest, investigate, investigation, investigator, involve, involvement, ion, iron, ischemia, ischemic, island, isoform, isolate, isolated, isolation, issue, item, joint, journal, junction, keep, key, kg, kidney, kilobyte, kinase, kind, kit, km, knee, know, knowledge, lab, label, labor, laboratory, lack, lack, lane, language, large, laser, last, late, lateral, law, layer, lead, leaf, learn, leave, left, length, lesion, leukemia, level, library, life, ligand, light, light, likelihood, likely, limit, limit, limitation, limited, line, lineage, linear, link, link, linkage, lipid, lipoprotein, liquid, list, list, literature, live, liver, load, local, localization, localize, locate, location, locus, log, logistic, long, long-term, longitudinal, look, loop, lose, loss, low, lower, lung, lymph, lymphocyte, lysate, macrophage, magnetic, magnitude, main, maintain, maintenance, major, majority, make, malaria, male, male, malignant, mammal, mammalian, mammary, man,

manage, management, manner, manual, manufacturer, manuscript, map, map, mapping, marker, marrow, mass, match, match, material, maternal, matrix, maturation, mature, maximal, maximum, mean, mean, mean, measure, measure, measurement, mechanical, mechanism, median, mediate, mediator, medical, medication, medicine, medium, meet, melanoma, member, membrane, memory, mental, mention, metabolic, metabolism, metabolite, metastasis, metastatic, method, methodology, methylation, mg, MI, microarray, microscope, microscopy, migration, mild, milk, min, minimal, minimize, minimum, minor, minute, miss, mitochondrial, mitochondrion, mix, mixed, mixture, ml, mm, mmol, mode, model, model, moderate, modification, modified, modify, modulate, modulation, module, molecular, molecule, monitor, monitoring, monkey, monoclonal, monocyte, month, morbidity, morphologic, morphology, mortality, mosquito, mother, motif, motor, mouse, move, movement, mRNA, ms, multiple, multivariate, murine, muscle, mutant, mutant, mutation, myocardial, name, national, native, natural, nature, necessary, neck, necrosis, need, need, negative, neonatal, nerve, nervous, network, neural, neuron, neuronal, neutrophil, new, next, nitric, nitrogen, nm, node, noise, normal, normalization, normalize, north, note, novel, nuclear, nucleic, nucleotide, nucleus, null, number, numerous, nurse, nursing, nutrition, obesity, objective, observation, observe, obtain, obvious, occur, occurrence, odds, offer, office, oil, old, oligonucleotide, oncogene, only, onset, ontology, oocyte, open, operate, operation, operon, opportunity, optimal, optimize, option, oral, order, organ, organism, organization, orientation, origin, original, ortholog, osteoarthritis, other, outcome, output, ovarian, ovary, overall, overexpression, overlap, own, oxidation, oxidative, oxide, oxygen, package, page, pain, pair, pair, pancreatic, paper, parallel, parameter, parasite, parent, part, partial, participant, participate, participation, particle, particular, partner, past, pathogen, pathogenesis, pathological, pathology, pathway, patient, pattern, PC, PCR, peak, pediatric, people, peptide, perceive, percent, percentage, perception, perform, performance, perfusion, period, peripheral, permit, person, personal, perspective, pH, phage, phase, phenomenon, phenotype, phenotypic, phosphatase, phosphate, phosphorylation, phylogenetic, phylogeny, physical, physician, physiological, physiology, PI, pig, place, place, placebo, plant, plaque, plasma, plasmid, Plasmodium falciparum, plate, platelet, platform, play, plot, pneumonia, point, point, policy, polymerase, polymorphism, pool, poor, population, portion, position, positive, positive, possess, possibility, possible, posterior, postoperative, potent, potential, potential, power, practical, practice, practitioner, precursor, predict, prediction, predictive, predictor, preference,



pregnancy, pregnant, preliminary, preparation, prepare, prescribe, prescription, presence, present, present, presentation, press, pressure, prevalence, prevent, prevention, previous, primary, primate, prime, primer, principal, principle, prior, priority, private, probability, probe, problem, procedure, process, process, processing, produce, product, production, professional, professional, profile, progesterone, prognosis, prognostic, program, progress, progression, progressive, project, proliferation, prolonged, promote, promoter, property, proportion, propose, prospective, prostaglandin, prostate, protease, protect, protection, protective, protein, protocol, prove, provide, provider, proximate, psychiatric, psychiatry, psychological, public, publication, publish, pulmonary, purchase, purification, purify, purpose, putative, pylorus, qualitative, quality, quantification, quantify, quantitative, quantity, query, question, questionnaire, rabbit, radiation, raise, random, randomize, range, range, rapid, rare, rat, rate, rating, ratio, raw, reach, reaction, reactive, reactivity, read, reading, reagent, real, real-time, reason, receive, recent, receptor, recognition, recognize, recombinant, recombination, recommend, recommendation, record, record, recover, recovery, recruit, recruitment, recurrence, recurrent, red, reduce, reduced, reduction, refer, reference, reflect, regard, regimen, region, regional, regression, regular, regulate, regulation, regulator, regulatory, rehabilitation, relate, related, relation, relationship, relative, release, release, relevance, relevant, reliability, reliable, rely, remain, remaining, removal, remove, renal, repair, repeat, repeat, repeated, replace, replacement, replicate, replication, report, report, reporter, represent, representation, representative, representative, reproductive, require, requirement, research, researcher, resident, residual, residue, resistance, resistant, resolution, resolve, resource, respective, respiratory, respond, respondent, response, responsible, responsive, responsiveness, restrict, restriction, result, result, resulting, retain, return, reveal, reverse, reverse, review, review, rheumatoid, rhythm, ribosomal, rich, right, right, rise, risk, RNA, robust, role, room, root, round, routine, rule, run, rural, *Saccharomyces cerevisiae*, safety, same, sample, sample, sampling, satisfaction, say, scale, scan, scenario, scheme, schizophrenia, school, science, scientific, score, score, screen, screen, screening, search, search, second, second, secondary, secrete, secretion, section, sector, see, seed, seek, seem, segment, select, selected, selection, selective, sense, sensitive, sensitivity, sensory, separate, separate, separation, sepsis, sequence, sequence, series, serine, serious, serum, serve, server, service, session, set, set, setting, severe, severeness, sex, sexual, shape, share, shift, shock, short, show, side, sign, signal, signaling, significance, significant, similar, similarity, simple, simulated, simulation, single,

site, situation, size, skeletal, skill, skin, sleep, slide, slope, slow, small, smoker, smoking, smooth, SNP, social, society, sodium, software, solid, soluble, solution, somatic, source, south, space, spatial, special, species, specific, specificity, specify, specimen, spectrum, sperm, spinal, splice, splice, spontaneous, spot, stability, stable, staff, stage, stain, staining, standard, standard, standardized, start, start, state, state, statement, statistic, statistical, status, stem, step, steroid, stimulate, stimulation, stimulus, store, strain, strand, strategy, strength, stress, stroke, strong, structural, structure, student, study, study, subgroup, subject, subject, subject, subsequent, subset, substance, substantial, substitution, substrate, subtype, subunit, success, successful, suffer, sufficient, suggest, suitable, sulfate, sum, summarize, summary, superfamily, supernatant, supplement, support, support, suppress, suppression, suppressor, surface, surfactant, surgery, surgical, surveillance, survey, survival, susceptibility, symptom, syndrome, synovial, synthase, synthesis, synthesize, synthetic, system, systematic, systemic, systolic, T-cell, table, tag, take, target, target, task, team, technical, technique, technology, temperature, template, temporal, tend, term, terminal, test, test, testing, testis, text, thank, theoretical, theory, therapeutic, therapy, think, third, threshold, thyroid, time, tissue, tolerance, tool, top, topology, total, total, toxicity, tract, traditional, train, training, trait, transcript, transcription, transcriptional, transduction, transfect, transfection, transfer, transfer, transform, transformation, transgenic, transient, transition, translation, translocation, transmission, transplantation, transport, transporter, trauma, treat, treatment, tree, trend, trial, trigger, true, tube, tuberculosis, tumor, type, typical, tyrosine, ultrasound, unable, unclear, undergo, underlie, underlying, understand, understanding, undertake, unique, unit, university, unknown, unlikely, unpublished, untreated, upper, uptake, urban, urinary, usage, use, use, useful, user, uterine, utility, utilization, utilize, vaccination, vaccine, valid, validate, validation, validity, value, variability, variable, variable, variance, variant, variation, variety, various, vary, vascular, vector, vein, velocity, venous, ventilation, ventricular, verify, version, vessel, viability, view, viral, virus, visit, visual, visualization, visualize, vitamin, volt, volume, wall, wash, water, wave, way, weak, web, website, week, weight, well, white, whole, wide, wild-type, window, woman, word, work, work, worker, world, write, year, yeast, yield, young, zinc

## Annexe 2 : Les 2000 lemmes anglais du corpus ScienText traduits

Les colonnes contiennent la classe grammaticale du lemme, son nombre d'occurrences dans ScienText, le lemme anglais, sa traduction en espagnol français, italien et portugais, sa définition et sa catégorie sémantique (Lspé pour lexique de spécialité de la santé; LST pour lexique scientifique transdisciplinaire. Là encore, le fichier complet sera à terme accessible sur le portail hal.

A	2000	abdominal	abdominal	abdominal	addominale	abdominal	Of or pertaining to the abdomen, ventral	Lspé
N	8029	ability	capacidad	capacité	capacità	capacidade	The quality of being able (...)	LST
A	6459	able	capaz	capable	capace	capaz	Free from constraints preventing completion of task	LST
A	2415	abnormal	anormal	anormal	anomalo	anormal	Deviating from the usual or normal type	LST
N	1361	abnormality	malformación	malformation	malformazione	malformação	1) abnormality : Something abnormal ; an aberration ; an abnormal occurrence or feature	Lspé
N	1361	abnormality	anomalía	anomalie	anomalia	anomalia	2) abnormality : The state or quality of being abnormal ; variation ; irregularity	LST
N	7676	absence	ausencia	absence	assenza	ausência	Lack ; deficiency ; nonexistence	LST
A	1844	absent	ausente	absent	assente	ausente	Not existing, lacking	LST
A	2642	absolute	absoluto	absolu	assoluto	absoluto	(mathematics) Indicating an expression that is true for all real numbers	LST
N	2242	abundance	abundancia	abondance	abbondanza	abundância	Frequency, amount, ratio of something within a given environment or sample	LST
A	1614	academic	académico / universitario	académique / universitaire	accademico / universitario	acadêmico / universitário	Belonging to an academy or other higher institution of learning; also a scholarly society or organization	LST
V	2226	accept	aceptar	accepter	accettare	aceitar	To receive, especially with a consent, with favor, or with approval	LST
N	2390	access	acceso	accès	accesso	acesso	1) access : (Internet) Connection to or communication with a computer program or to the Internet	LST
N	2390	access	acceso	accès	accesso	acesso	2) access : The right or ability of approaching or entering ; admittance ; admission ; accessibility	LST
V	1675	accompany	acompañar	accompagner	accompagnare	acompanhar	To be found at the same time	LST
N	3101	account	cuenta	compte	conto	conta	A reason, grounds, consideration, motive	LST

V	3431	account	representar	représenter	rappresentare	representar	To constitute in amount or portion	LST
V	1855	accumulate	acumularse	s'accumuler	accumularsi	acumular-se	To grow or increase in quantity or number; to increase greatly	LST
N	3755	accumulation	acumulación	accumulation	accumulazione	acumulação	The process of growing into a heap or a large amount	LST
N	4827	accuracy	exactitud / precisión	exactitude / précision	esattezza / precisione	exatidão / precisão	Exact conformity to truth, or to a rule or model ; degree of conformity of a measure to a true or standard value	LST
A	2762	accurate	exacto / preciso	exact / précis	esatto / preciso	exato / preciso	Precise or correct	LST
V	5995	achieve	alcanzar / lograr / conseguir	atteindre / réussir	conseguire / riuscire	atingir / alcançar	To carry out successfully; to accomplish	LST
N	24992	acid	ácido	acide	acido	ácido	Any of several classes of compound (...)	Lspé
A	2792	acid	ácido	acide	acido	ácido	Of, or pertaining to an acid	Lspé
V	1424	acquire	adquirir	acquérir	acquire	adquirir	1) acquire : (medicine) To contract	Lspé
V	1424	acquire	adquirir	acquérir	acquire	adquirir	2) acquire : (computing) To sample signals and convert them into digital values	LST
N	1625	acquisition	adquisición	acquisition	acquisizione	aquisição	(computing) The process of sampling signals that measure real world physical conditions and converting these signals into digital numeric values that can be manipulated by a computer	LST
N	1893	act	acto	acte	atto	ato	The process of doing something	LST
V	4095	act	actuar	agir	agire	atuar / agir	To have an effect	LST
N	3913	actin	actina	actine	actina	actina	(biochemistry) A globular structural protein (...)	Lspé
N	7232	action	acción	action	azione	ação	A way of motion or functioning	LST
V	12447	activate	activar	activer	attivare	ativar	To put into action, to put to work	Lspé
N	17257	activation	activación	activation	attivazione	ativação	(biochemistry) The process through which molecules are made able to react.	Lspé
N	2295	activator	activador	activateur	insufficienza	ativador	(biology) Any chemical or agent which regulates one or more genes by increasing the rate of transcription	Lspé
A	8829	active	activo	actif	attivo	ativo	Having the quality or power of acting	LST
N	39778	activity	actividad	activité	attività	atividade	Something done as an action or movement	Lspé
A	2590	actual	efectivo	effectif	effettivo	efetivo	Existing in act or reality, not just potentially, occurring in fact	LST
A	10609	acute	agudo	aigu	acuto	agudo	Of a short-lived condition, in contrast to a chronic condition	Lspé
V	1580	adapt	adaptarse	s'adapter	adattarsi	adaptar-se	To change oneself so as to be adapted	LST
N	1614	adaptation	adaptación	adaptation	adattamento	adaptação	Adjustment to extant conditions (...)	LST
V	9159	add	añadir	ajouter	aggiungere	aditar	To join or unite (...) so as to increase the number, augment the quantity or enlarge the magnitude, or so as to form into one aggregate	LST

N 4069	addition	adición	addition	addizione	adição	The act of adding anything	LST
A 17738	additional	adicional	additionnel	addizionale	adicional	Supplemented or added to	LST
V 4668	address	enfrentar / abordar	aborder	affrontare	enfrentar / abordar	To direct attention towards a problem or obstacle, in an attempt to resolve it	LST
N 1735	adenocarcinoma	adenocarcinoma	adénocarcinome	adenocarcinoma	adenocarcinoma	(oncology) Any of several forms of carcinoma that originate in glandular tissue	Lspé
A 1881	adequate	adecuado	adéquat	adeguato	adequado	Equal to some requirement	LST
N 4768	adhesion	adhesión	adhésion	adesione	adesão	(biochemistry) The binding of a cell to a surface or substrate	Lspé
A 2201	adjacent	adyacente	adjacent	adiacente	adjacente	Lying next to, close or contiguous	LST
V 4163	adjust	ajustar	ajuster	aggiustare	ajustar	To improve or rectify	LST
N 2013	adjustment	ajuste	ajustement	aggiustamento	ajuste	The result of adjusting something ; a small change ; a minor correction ; a modification or alteration	LST
V 3481	administer	administrar	administrer	amministrare	administrar	(medicine) To give a drug to a patient, be it orally or by any other means	Lspé
N 6077	administration	administración	administration	somministrazione	administração	The act of administering, or tendering something to another ; dispensation	Lspé
N 3809	admission	admisión / entrada	admission / entrée	ammissione / entrata	admissão / entrada	The act or practice of admitting	Lspé
V 1807	admit	admitir	admettre	ammettere	admitir	To concede as true; to acknowledge or assent to, as an allegation which it is impossible to deny; to own or confess	LST
N 9747	adult	adulto	adulte	adulto	adulto	A fully grown human or animal	Lspé
A 1937	adult	adulto	adulte	adulto	adulto	Fully grown	Lspé
N 1640	advance	avance	avancée	avanzamento	avanço	A forward move; improvement or progression	LST
A 2438	advanced	avanzado	avancé	avanzato	avançado	In a late stage of development	Lspé
N 3980	advantage	ventaja	avantage	vantaggio	vantagem	Any condition, circumstance, opportunity or means, particularly favorable to success, or to any desired end	LST
A 3962	adverse	adverso	indésirable	collaterale	adverso	Unfavorable, antagonistic in purpose or effect	Lspé
V 12452	affect	afectar	affecter	colpire	afetar	Of an illness or condition, to infect or harm (a part of the body)	Lspé
A 1648	affected	afectado	affecté	colpito	afetado	Acted on, influenced, or moved, either physically or materially	Lspé
N 1817	affinity	afinidad	affinité	affinità	afinidade	1) affinity : (medicine) The attraction between an antibody and an antigen	Lspé
N 1817	affinity	afinidad	affinité	affinità	afinidade	2) affinity : (chemistry) An attractive force between atoms, or groups of atoms, that contributes towards their forming bonds	Lspé
A 2001	African	africano	africain	africano	africano	Of or pertaining to Africa	LST
N 27059	age	edad	âge	età	idade	The whole duration of a being, whether animal, vegetable, or other kind ; lifetime	Lspé
A 4854	aged	de edad	âgé de	di età	de idade	Having the age of	Lspé

N 7630	agent	agente	agent	agente	agente	An active power or cause ; that which has the power to produce an effect ; as a physical, chemical or medicinal agent	Lspé
N 3001	agonist	agonista	agoniste	agonista	agonista	(biochemistry) A molecule that can combine with a receptor on a cell to produce a physiological reaction	Lspé
V 2143	agree	concordar / corresponder	concorder / correspondre	concordare / corrispondere	concordar / corresponder	To be conformable; to resemble; to coincide; to correspond	LST
N 4052	agreement	acuerdo	accord	accordo	acordo	A state whereby several parties share a view or opinion ; the state of not contradicting one another	LST
N 2658	aim	finalidad	finalité / visée	finalità / scopo	finalidade	Intention ; purpose ; design ; scheme	LST
V 2150	aim	apuntar	viser	mirare	apontar / visar	To direct the intention or purpose; to attempt the accomplishment of a purpose; to try to gain; to endeavor	LST
N 2414	air	aire	air	aria	ar	The atmospheric substance above the surface of the earth which animals breathe	Lspé
N 5492	airway	tráquea	trachée / voies aériennes	trachea	traqueia	(anatomy) The trachea	Lspé
N 1656	albumin	albúmina	albumine	albumina	albumina	(organic chemistry) Any of a class of monomeric proteins that are soluble in water, and are coagulated by heat (...)	Lspé
N 3513	alcohol	alcohol	alcool	alcol / alcole	álcool	(uncountable) Beverage containing ethanol	Lspé
N 8288	algorithm	algoritmo	algorithmme	algoritmo	algoritmo	A precise step-by-step plan for a computational procedure (...)	LST
V 2469	align	alineair	aligner	allineare	alinhar	(bioinformatics) To organize a linear arrangement of DNA, RNA or protein sequences which have regions of similarity	Lspé
N 10465	alignment	alineamiento	alignement	allineamento	alinhamento	(bioinformatics) A way of arranging DNA, RNA or protein sequences in order to identify regions of similarity	Lspé
N 8647	allele	alelo	allèle	allele	alelo	(genetics) One of a number of alternative forms of the same gene occupying a given position on a chromosome	Lspé
N 1802	allergy	alergia	allergie	allergia	alergia	(pathology, immunology) A disorder of the immune system causing adverse reactions to substances (allergens) not harmful to most and marked by the body's production of histamines and associated with atopy, anaphylaxis, and asthma	Lspé
V 13635	allow	permitir	permettre	permettere	permitir	To grant, give, admit, accord, afford, or yield; to let one have	LST
N 6384	alpha	alfa	alpha	alfa	alfa	(sciences) The name of the symbols A and $\alpha$ used in science and mathematics (...)	LST
V 4471	alter	alterar	altérer	alterare	alterar	To change the form or structure of	Lspé
N 3955	alteration	alteración	altération	alterazione	alteração	The state of being altered ; a change made in the form or nature of a thing ; changed condition	Lspé
A 1567	altered	alterado	altéré	alterato	alterado	Modified, changed	Lspé
N 2538	alternative	alternativa	alternative	alternativa	alternativa	A situation which allows a mutually exclusive choice between two or more possibilities ; a choice between two or more possibilities	LST
A 4429	alternative	alternativo	alternatif	alternativo	alternativo	Not traditional, outside the mainstream (...)	LST
A 898	alveolar	alveolar	alvéolaire	alveolare	alveolar	1) alveolar : (anatomy) Relating to the jaw ridge containing the tooth sockets	Lspé

A 898	alveolar	alveolar	alvéolaire	alveolare	alveolar	2) alveolar : (anatomy) Relating to the alveoli (...) of the lungs	Lspé
N 5803	American	americano	Américain	americano	americano	An inhabitant of the Americas (...)	LST
N 10626	amino acid	aminoácido	acide aminé	aminoacido / amminoacido	aminoácido	(chemistry) Any organic compound containing both an amino and carboxylic acid functional group	Lspé
N 9006	amount	cantidad	quantité	quantità	quantidade	A quantity or volume	LST
N 6448	amplification	amplificación	amplification	amplificazione	amplificação	(genetics) The using of the polymerase chain reaction (PCR) for isolating and exponentially amplifying a fragment or segment of DNA	Lspé
V 5331	amplify	amplificar	amplifier	amplificare	amplificar	To render larger, more extended, or more intense, and the like (...)	Lspé
N 1562	analogue	análogo	analogue	analogo	análogo	(chemistry) a structural derivative of a parent compound (...)	Lspé
N 41494	analysis	análisis	analyse	analisi	análise	1) analysis : The process of breaking down a substance into its constituent parts, or the result of this process	Lspé
N 41494	analysis	análisis	analyse	analisi	análise	2) analysis : The mathematical study of functions, sequences, series (...)	LST
A 1600	analytic	analítico	analytique	analitico	analítico	Of or relating to any form of analysis, or to analytics	LST
V 15903	analyze	analizar	analyser	analizzare	analisar	To subject to analysis	LST
N 1781	androgen	andrógeno	androgène	androgeno	andrógeno	(biochemistry) The generic term for any natural or synthetic compound, usually a steroid hormone, that stimulates or controls the development and maintenance of masculine characteristics in vertebrates	Lspé
N 2010	angiogenesis	angiogénesis	angiogenèse	angiogenesi	angiogênese	(medicine, zoology) The formation and development of new blood vessels	Lspé
N 11684	animal	animal	animal	animale	animal	(...) a multicellular organism (...)	Lspé
A 3156	animal	animal	animal	animale	animal	Of, or relating to animals	Lspé
V 2555	annotate	anotar	annoter	annotare	anotar	To add annotation	Lspé
N 5370	annotation	anotación	annotation	annotazione	anotação	(genetics) Information relating to the genetic structure of sequence bases	Lspé
A 2195	annual	anual	annuel	annuale	anual	Happening once every year	LST
N 2623	antagonist	antagonista	antagoniste	antagonista	antagonista	(biochemistry) A chemical that binds to a receptor but does not produce a physiological response, blocking the action of agonist chemicals	Lspé
A 1754	anterior	anterior	antérieur	anteriore	anterior	(anatomy) Nearer the forward end ; nearer the head of an animal or the front of a human	Lspé
N 4222	antibiotic	antibiótico	antibiotique	antibiotico	antibiótico	(pharmacology) Any substance that can destroy or inhibit the growth of bacteria and similar microorganisms	Lspé
N 21670	antibody	anticuerpo	anticorps	anticorpo	anticorpo	(immunology) A protein produced by B-lymphocytes that binds to a specific antigen	Lspé
N 10626	antigen	antígeno	antigène	antigene	antígeno	(immunology) A substance that induces an immune response, usually foreign	Lspé
N 2323	antioxidant	antioxidante	antioxydant	antiossidante	antioxidante	Any substance that acts to slow or prevent the oxidation of another chemical	Lspé
A 1999	antisense	* antisentido	* antisens	* antisenso	* anti-sentido / sentido negativo	(genetics) Which interacts with a complementary strand of nucleic acids so as to suppress its transcription	Lspé
N 1911	anxiety	ansiedad	anxiété /	ansia	ansiedade	(pathology) A state of restlessness and agitation, often accompanied by a distressing	Lspé

			angoisse			sense of oppression (...)	
N 10072	apoptosis	apoptosis	* apoptose	apoptosi	apoptose	(biology) A process of programmed cell death by which cells undergo an ordered sequence of events which lead to death of the cell (...)	Lspé
A 3347	apoptotic	* apoptótico	* apoptotique	* apoptotico	apoptótico	Of or pertaining to apoptosis	Lspé
A 3127	apparent	aparente	apparent	apparente	aparente	Appearing to the eye or mind (...); seeming	LST
V 13242	appear	aparecer	apparaitre	apparire	aparecer	To become visible to the apprehension of the mind; to be known as a subject of observation or comprehension, or as a thing proved; to be obvious or manifest	LST
N 1716	appearance	aparición / aparecimiento	apparition / survenue	apparizione	aparicação / aparecimento	The act of appearing or coming into sight; the act of becoming visible to the eye	LST
N 2944	application	aplicación	application	applicazione	aplicação	1) application : A potential use, often found by scientific research	LST
N 2944	application	aplicación	application	applicazione	aplicação	2) application : The act of applying or laying on, in a literal sense ; as the application of emollients to a diseased limb	Lspé
N 2944	application	aplicación	application	applicazione	aplicação	3) application : (computing) A computer program or the set of software that the end user perceives as a single entity as a tool for a well-defined purpose	LST
V 10497	apply	aplicar	appliquer	applicare	aplicar	To put to use; to use or employ (...) to appropriate; to devote	LST
N 17787	approach	enfoque	approche	approccio	enfoque / abordagem	A manner in which a problem is solved or policy is made	LST
A 6208	appropriate	apropiado	approprié	appropriato	apropriado	Belonging particularly, suitable ; fit ; proper	LST
V 5004	approve	aprobar	approuver	approvare	aprovar	To sanction officially; To ratify; to confirm	LST
N 17009	area	área	aire	area	área	Any particular extent of surface (...)	LST
V 2812	arise	aparecer / nacer	apparaitre / naitre	apparire / nascere	aparecer / nascere	To spring up; to come into action, being, or notice; to become operative, sensible, or visible; to begin to act a part; to present itself	LST
N 1070	arm	brazo	bras	braccio	braço	1) arm : (anatomy) The extended portion of the upper limb, from the shoulder to the elbow	Lspé
N 1070	arm	brazo	bras	braccio	braço	2) arm : (genetics) One of the two parts of a chromosome	Lspé
N 1070	arm	grupo *	groupe	gruppo	grupo	3) arm : A group of patients in a medical trial	Lspé
N 8368	array	microarreglo / *	microréseau / micromatrice / biopuce	microarray / micromatrice / biochip	microarranjo / microarray	The same as a microarray	Lspé
N 1844	arrest	paro	arrêt	arresto	paragem	The condition of being stopped, standstill	Lspé
A 3183	arterial	arterial	artériel	arteriale / arterioso	arterial	Of or relating to an artery	Lspé



N 5276	artery	arteria	artère	arteria	artéria	An efferent blood vessel from the heart, conveying blood away from the heart (...)	Lspé
N 10929	arthritis	artritis	arthrite	artrite	artrite	Inflammation of a joint or joints causing pain and/or disability, swelling and stiffness (...)	Lspé
N 4917	article	artículo	article	articolo	artigo	A story, report, or opinion piece in a newspaper, magazine, journal, etc.	LST
V 4040	ask	preguntar	demander	domandare	perguntar	To request (information, or an answer to a question)	LST
N 4683	aspect	aspecto	aspect	aspetto	aspecto	Any specific feature, part, or element of something	LST
N 1583	aspirin	aspirina	aspirine	aspirina	aspirina	(pharmacology) An analgesic drug, acetylsalicylic acid	Lspé
N 14624	assay	ensayo / análisis	dosage	dosaggio /saggio	ensaio / doseamento	The qualitative or quantitative chemical analysis of something	Lspé
V 2138	assay	ensayar / analizar	doser	dosare / saggiare	ensaiar / dosar	To analyze or estimate the composition or value of (a metal, ore etc.)	Lspé
N 1673	assembly	ensamblaje / montaje	assemblage / montage	assemblaggio / montaggio	assemblagem / montagem	1) assembly : The act of putting together a set of pieces, fragments, or elements	Lspé
N 1673	assembly	asamblea	assemblée	assemblea	assembleia	2) assembly : A congregation of people in one place for a purpose	LST
V 14553	assess	evaluar	évaluer	valutare	avaliar	To determine, estimate or judge the value of; to evaluate	LST
N 11215	assessment	evaluación	évaluation	valutazione	avaliação	An appraisal or evaluation	LST
V 4057	assign	asignar	assigner	assegnare	atribuir	To attribute or sort something into categories	LST
V 2433	assist	asistir	assister	assistere	assistir	To help compensate for what is missing with the help of a medical technique or therapy	Lspé
N 2385	assistance	asistencia	assistance	assistenza	assistência	Aid ; help ; the act or result of assisting	Lspé
V 31204	associate	asociar	associer	associare	associar	To connect or join together; combine	LST
N 8315	association	asociación	association	associazione	associação	1) association : (statistics) Any relationship between two measured quantities that renders them statistically dependent (...)	LST
N 8315	association	asociación	association	associazione	associação	2) association : A group of persons associated for a common purpose; an organization; society	LST
V 5600	assume	suponer	supposer	supporre	supor	To authenticate by means of belief; to surmise; to suppose to be true, especially without proof	LST
N 3279	assumption	suposición	supposition	supposizione	suposição	The thing supposed ; a postulate, or proposition assumed ; a supposition	LST
N 5755	asthma	asma	asthme	asma	asma	(pathology) A long-term respiratory condition in which the airways may unexpectedly and suddenly narrow, often in response to an allergen, cold air, exercise, or emotional stress (...)	Lspé
N 1892	atherosclerosis	aterosclerosis	athérosclérose	aterosclerosi	aterosclerose	(pathology) The clogging or hardening of arteries or blood vessels caused by plaques (...)	Lspé
A 1912	atrial	auricular	* atrial / auriculaire	atriale	atrial	Of or pertaining to an atrium, especially the atrium of the heart	Lspé
N 2129	attempt	tentativa	tentative	tentativo	tentativa	The action of trying at something	LST

V	1825	attempt	tentar / intentar	tenter	tentare	tentar / intentar	To try	LST
N	2675	attention	atención	attention	attenzione	atenção	Mental focus	LST
N	2036	attitude	actitud	attitude	atteggiamento	atitude	The disposition or state of mind	LST
V	2254	attribute	atribuir	attribuer	attribuire	atribuir	To ascribe to a given cause, reason etc.	LST
N	17797	author	autor *	auteur	autore	autor	The originator or creator of a work (...)	LST
N	1645	autoantibody	autoanticuerpo o	* autoanticorpo	autoanticorpo	* autoanticorpo	An antibody formed in response to an agent (autoantigen) produced by the organism itself	Lspé
A	2071	autoimmune	* autoinmune	* auto-immun	autoimmune	autoimune	Of or pertaining to autoimmunity	Lspé
A	1656	automated	automatizado	automatisé	automatizzato	automatizado	Made automatic	LST
N	2499	availability	disponibilida d	disponibilité	disponibilità	disponibilida de	The quality of being available	LST
A	15425	available	disponible	disponible	disponibile	disponível	Readily obtainable	LST
N	3805	average	media	moyenne	media	média	(statistics) Any measure of central tendency, especially any mean, the median, or the mode	LST
A	7793	average	medio	moyen	medio	médio	Constituting or relating to the average	LST
V	3376	avoid	evitar	éviter	evitare	evitar	To keep away from; to keep clear of; to endeavor not to meet; to shun; to abstain from	LST
N	1927	axis	eje	axe	asse	eixo	(geometry) An imaginary line around which an object spins (an axis of rotation) or is symmetrically arranged (an axis of symmetry)	LST
N	1563	bacillus	bacilo	bacille	bacillo	bacilo	Any of various rod-shaped, spore-forming aerobic bacteria in the genus Bacillus, some of which cause disease	Lspé
N	4626	background	contexto	contexte	contesto	contexto	Information relevant to the current situation (...)	LST
A	7312	bacterial	bacteriano	bactérien	batterico	bacteriano	Of or caused by bacteria	Lspé
N	6421	bacterium	bacteria	bactérie	batterio	bactéria	(microbiology) A single celled organism with no nucleus	Lspé
A	992	bad	mal	mauvais	male	mau	1) bad : Not suitable or fitting	LST
A	992	bad	mal	mauvais	male	mau	2) bad : Not good ; unfavorable, negative	LST
N	1720	balance	equilibrio	équilibre	equilibrio	equilíbrio	A state in which opposing forces harmonize; equilibrium	LST
N	4385	band	banda	bande	banda	banda	(genetics) Each of the lines representing the different components of a molecule and made visible after an electrophoresis in gel of a mixture	Lspé
N	1221	barrier	barrera	barrière	barriera	barreira	1) barrier : (physiology) A separation between two areas of the body where specialized cells allow the entry of certain substances but prevent other substances to enter	Lspé
N	1221	barrier	barrera	barrière	barriera	barreira	2) barrier : An obstacle or impediment	LST
A	3117	basal	basal	basal	basale / prossimale	basal	(anatomy) Associated with the base of an organism or structure	Lspé

N	3247	base	base	base	base	base	1) base : (chemistry) Any of a class of generally water-soluble compounds that turn red litmus blue, and react with acids to form salts	Lspé
N	3247	base	base	base	base	base	2) base : The starting point of a logical deduction or thought ; basis	LST
V	34492	base	basar / fundar	baser / fonder	basare / fondare	basear / fundar	To give as its foundation or starting point; to lay the foundation of	LST
N	5241	baseline	referencia	référence	riferimento	referência	A datum used as the basis for calculation or for comparison	LST
A	4030	basic	básico	basique	basico	básico	Elementary, simple, merely functional	LST
N	6661	basis	base	base	base	base	A starting point, base or foundation for an argument or hypothesis	LST
V	51005 2	be	estar / ser	être	essere	estar / ser	Copulative verb used to indicate different relations	LST
V	2605	bear	dar a luz	donner naissance	dare alla luce	dar à luz	To give birth to someone or something	Lspé
V	7832	become	volverse	devenir	diventare	tornar-se	To come about; happen; come into being; arise	LST
V	2670	begin	comenzar / empezar	commencer	cominciare	começar	To start, to initiate or take the first step into something	LST
N	4170	behavior	comportamiento	comportement	comportamento	comportamento	1) behavior : The way matter moves or acts	LST
N	4170	behavior	comportamiento del	comportement	comportamento	comportamento	2) behavior : The way a living creature behaves or acts	LST
A	2580	behavioral	comportamiento	comportementa	comportamento	comportamento	Of or relating to behavior	LST
V	3234	believe	creer	croire	credere	crer	To accept as true, particularly without absolute certainty (i.e., as opposed to knowing)	LST
V	2625	belong	pertenecer	appartenir	appartenere	pertencer	To be a part of a group	LST
A	1712	beneficial	beneficioso	bénéfique	benefico	benéfico	Helpful or good to something or someone	Lspé
N	5282	benefit	beneficio	bénéfice	beneficio	benefício	An advantage, help, sake or aid from something	Lspé
N	3132	beta	beta	bêta	beta	beta	The name of the second letter of the Greek alphabet (B,β) (...)	LST
N	5949	bias	sesgo	biais	distorsione	enviesamento	population value, which reduces the representativeness of the estimator by systematically distorting it	LST
V	15649	bind	unir	lier	legare	ligar	(transitive, chemistry) To make two or more elements stick together	Lspé
N	16002	binding	unión	liaison	legame	ligação	(chemistry) The action or result of making two or more molecules stick together	Lspé
A	3305	biochemical	bioquímico	biochimique	biochimico	bioquímico	Involving chemical processes in living organisms	Lspé
A	8956	biological	biológico	biologique	biologico	biológico	Of or relating to biology	Lspé

N 4761	biology	biología	biologie	biologia	biologia	The structure, function, and behavior of an organism or type of organism	Lspé
N 3567	biopsy	biopsia	biopsie	biopsia	biopsia	The removal and examination of a sample of tissue from a living body for diagnostic purposes	Lspé
N 1563	biosynthesis	biosíntesis	biosynthèse	biosintesi	biossíntese	(biochemistry) The synthesis of organic compounds within a living organism, especially the synthesis of large compounds from small ones	Lspé
N 5559	birth	nacimiento	naissance	nascita	nascimento	The process of childbearing (...)	Lspé
A 2062	black	negro	noir	nero	negro	Of or relating to any of various ethnic groups having dark pigmentation of the skin	Lspé
N 1649	bladder	vejiga	vessie	vescica	bexiga	(anatomy) An elastic muscular sac situated in the pelvic cavity, into which urine from the kidneys is stored prior to disposal by urination (...)	Lspé
N 3280	BLAST	BLAST	BLAST	BLAST	BLAST	(Basic Local Alignment Search Tool) An algorithm for comparing primary biological sequence information.	Lspé
N 1743	block	bloque	bloc	blocco	bloco	1) block : (chemistry) A portion of a macromolecule, comprising many units, that has at least one feature not present in adjacent portions	Lspé
N 1743	block	bloque	bloc	blocco	bloco	2) block : A substantial, often approximately cuboid, piece of any substance	LST
V 4351	block	bloquear	bloquer	bloccare	bloquear	To prevent (something from happening or someone from doing something)	Lspé
N 24167	blood	sangre	sang	sangue	sangue	A vital liquid flowing in the body (...)	Lspé
N 4550	blot	hibridación	buvardage	blotting	blotting	(biochemistry) A method of transferring proteins, DNA or RNA, onto a carrier	Lspé
A 1782	blue	azul	bleu	azzurro	azul	Of the color blue	LST
N 11361	body	cuerpo	corps	corpo	corpo	The physical structure of a human or animal seen as one single organism	Lspé
N 8524	bone	hueso	os	osso	osso	A composite material consisting largely of calcium phosphate and collagen and making the skeleton of most vertebrates	Lspé
N 1929	boundary	límite	limite	limite	limite	The dividing line or location between two areas	Lspé
A 3360	bovine	bovino	bovin	bovino	bovino	Of or pertaining to cattle	Lspé
N 2203	box	caja	boite	box / scatola	caixa	(genetics) One of two specific regions in a promoter.	Lspé
N 12148	brain	cerebro	cerveau	cervello	cérebro	The control center of the central nervous system of an animal (...)	Lspé
N 2193	branch	rama	branche / rameau	ramo	ramo	Any of the parts of something that divides like the branch of a tree	Lspé
N 20246	breast	mama	sein	mammella	mama	Either of the two organs on the front of a woman's chest, which contain the mammary glands ; also the analogous organs in men	Lspé
A 1580	British	británico	britannique	britannico	britânico	Of the United Kingdom.	LST
A 2580	broad	amplio	large	ampio	ampio	Having a large measure of any thing or quality	LST
A 3887	brown	marrón / castaño	marron	marrone	marrom / castanho	Having a brown color	LST
N 7128	buffer	tampón	tampon	tampone	tampão	(chemistry) A solution used to stabilize the pH (acidity) of a liquid	Lspé
V 2121	buffer	* tamponar	tamponner	tamponare	tamponar	(chemistry) To maintain the acidity of a solution near a chosen value by adding an acid or a base	Lspé

V	2351	build	construir	construire	costruire	construir	To develop or give form according to a plan or process	LST
N	2102	burden	carga	charge / masse	carica / massa	carga	(medicine) The total amount of toxins, parasites, cancer cells, plaque or the such present in an organism	Lspé
N	1646	Caenorhabditis elegans	Caenorhabditis elegans	Caenorhabditis elegans	Caenorhabditis elegans	Caenorhabditis elegans	A taxonomic species within the family Rhabditidae—a nematode used as a model organism in biological research	Lspé
N	4915	calcium	calcio	calcium	calcio	cálcio	(chemistry) A chemical element, atomic number 20 (...)	Lspé
V	11213	calculate	calcular	calculer	calcolare	calcular	(mathematics) To determine the value of something or the solution to something by a mathematical process	LST
N	3135	calculation	calculación / cálculo	calcul	calcolo	calculação / cálculo	(mathematics) The act or process of calculating	LST
V	3671	call	llamar	appeler	chiamare	chamar	To name or refer to	LST
N	1813	cAMP	AMPc / cAMP	AMPc	cAMP	cAMP	(medicine) Initialism of cyclic adenosine monophosphate	Lspé
V	2208	can	poder	pouvoir	potere	poder	To know how to; to be able to	LST
A	2291	Canadian	canadiense	canadien	canadese	canadense	Of or pertaining to Canada.	LST
N	55246	cancer	cáncer	cancer	cancro	câncer	A disease in which the cells of a tissue undergo uncontrolled (and often rapid) proliferation	Lspé
N	4731	candidate	candidato	candidat	candidato	candidato	A synonym for a candidate gene, i.e., a gene which may play a role in a given disease	Lspé
A	2054	capable	capaz	capable	capace	capaz	Able and efficient	LST
N	4276	capacity	capacidad	capacité	capacità	capacidade	Capability ; the ability to perform some task	LST
V	1815	capture	capturar	capturer	catturare	capturar	To store (as in sounds or image) for later revisitation	LST
N	1948	carbon	carbono	carbone	carbonio	carbono	(chemistry) An atom of this element, in reference to a molecule containing it	Lspé
N	8633	carcinoma	carcinoma	carcinome	carcinoma	carcinoma	(oncology) An invasive malignant tumor derived from epithelial tissue that tends to metastasize to other areas of the body	Lspé
A	5829	cardiac	cardíaco	cardiaque	cardiaco	cardíaco	Pertaining to the heart	Lspé
A	4674	cardiovascular	cardiovascular	cardiovasculaire	cardiovascolare	cardiovascular	Relating to the circulatory system	Lspé
N	29348	care	atención	soins	assistenza	atenção	The treatment of those in need (especially as a profession)	Lspé
N	2111	carrier	portador	entraîneur	portatore	portador	(chemistry) A catalyst or other intermediary in a chemical reaction	Lspé
V	5749	carry	efectuar / realizar	effectuer / réaliser / mener	effettuare / svolgere	efetuar / realizar	1) carry : To execute or perform; to put into operation	LST
V	5749	carry	transportar	transporter	trasportare	transportar	2) carry : To transfer from one place (...) to another	Lspé
N	3988	cartilage	cartílago	cartilage	cartilagine	cartilagem	(anatomy) A type of dense, non-vascular connective tissue (...)	Lspé
N	38760	case	caso	cas	caso	caso	An instance or event as a topic of study	LST

A	1932	catalytic	catalítico	catalytique	catalitico	catalítico	Of or relating to a catalyst ; having properties facilitating chemical reaction or change	Lspé
N	6613	category	categoría	catégorie	categoria	categoria	A group, often named or numbered, to which items are assigned based on similarity or defined criteria	LST
N	2080	catheter	catéter	cathéter	catetere	cateter	(medicine) A small tube inserted into a body cavity to remove fluid, create an opening, distend a passageway or administer a drug	Lspé
N	6314	cause	causa	cause	causa	causa	The source of, or reason for, an event or action ; that which produces or effects a result	LST
V	11950	cause	causar	causer	causare	causar	To set off an event or action	LST
N	13197	cDNA	ADNc	ADNc	cDNA	ADNc	(genetics) Complementary DNA	Lspé
N	20152 1	cell	célula	cellule	cellula	célula	The basic unit of a living organism (...)	Lspé
A	10055	cellular	celular	cellulaire	cellulare	celular	Relating to a cell or cells	Lspé
N	4864	center	centro	centre	centro	centro	1) center : A place where some function or activity occurs	LST
N	4864	center	centro	centre	centro	centro	2) center : The middle portion of something; the part well away from the edges	LST
A	2076	central	central	central	centrale	central	1) central : Being in the center	LST
A	2076	central	central	central	centrale	central	2) central : Being the most important	LST
A	2076	central	central	central	centrale	central	3) central : (anatomy) Exerting its action towards the peripheral organs	Lspé
N	1719	centrifugation	centrifugación	centrifugation	centrifugazione	centrifugação	The process in which mixtures are separated using the centripetal force generated by spinning in a centrifuge	Lspé
A	2668	cerebral	cerebral	cérébral	cerebrale	cerebral	Of, or relating to the brain	Lspé
A	4613	certain	cierto	certain	certo	certo	Sure, positive, not doubting	LST
A	2277	cervical	cervical	cervical	cervicale	cervical	(anatomy) Of the cervix	Lspé
N	6931	chain	cadena	chaîne	catena	cadeia	(chemistry) A number of atoms in a series, which combine to form a molecule	Lspé
N	3378	challenge	desafío	défi	sfida	desafio	A difficult task (...)	LST
N	1780	chamber	cámara	chambre	camera	câmara	An enclosed space	Lspé
V	3162	change	cambiar	changer	cambiare	mudar	1) change : To become something different	LST
V	3162	change	cambiar	changer	cambiare	mudar	2) change : To make something into something different	LST
N	29075	change	cambio	changement	cambiamento	mudança	The process of becoming different	LST
N	6489	channel	canal	canal	canale	canal	(biochemistry) A type of membrane protein	Lspé
N	9274	characteristic	característica	caractéristique	caratteristica	característica	A distinguishable feature of a person or thing	LST
A	1900	characteristic	característico	caractéristique	caratteristico	característico	Being a distinguishing feature or a person or thing	LST
N	6022	characterization	caracterización	caractérisation	caratterizzazione	caracterização	The act or process of characterizing	LST
V	7331	characterize	caracterizar	caractériser	caratterizzare	caracterizar	To determine the characteristics of	LST
V	1694	check	verificar /	vérifier	verificare	verificar /	To verify or compare with a source of information	LST

		averiguar			averiguar / conferir		
N 3021	chemical	producto químico *	produit chimique	prodotto chimico	produto químico	(chemistry, sciences) Any specific chemical element or chemical compound or alloy	Lspé
N 3386	chemokine	quimiocina / *	* chimiokine / * chémokine	* chemochina	quimiocina	(biochemistry, immunology) Any of various chemotactic cytokines, produced during inflammation, that organize the leukocytes	Lspé
N 3060	chemotherapy	quimioterapia	chimiothérapie	chemioterapia	quimioterapia	(oncology) Chemical treatment to kill or halt the replication and/or spread of cancerous cells in a patient	Lspé
N 2537	chest	pecho	poitrine	petto	peito	(anatomy) The portion of the front of the human body from the base of the neck to the top of the abdomen ; the thorax (...)	Lspé
N 17119	child	hijo	enfant	figlio	filho	A daughter or son ; an offspring	Lspé
N 1961	childhood	infancia / niñez	enfance	infanzia	infância	The time during which one is a child, from between infancy and puberty	Lspé
N 1997	chip	chip	puce	chip	chip	(electronics) A hybrid device mounted in a substrate, containing electronic circuitry and miniaturized mechanical, chemical and/or biochemical devices	Lspé
N 4148	choice	elección	choix	scelta	escolha	An option ; a decision ; an opportunity to choose or select something	LST
N 6246	cholesterol	colesterol	cholestérol	colesterolo	colesterol	(biochemistry) A steroid lipid synthesized by the liver and transported in the bloodstream to the membranes of animal cells (...)	Lspé
V 6117	choose	escoger	choisir	scegliere	escolher	To pick; to make the choice of; to select	LST
N 2005	chromatin	cromatina	chromatine	cromatina	cromatina	(biology) A complex of DNA, RNA and proteins within the cell nucleus out of which chromosomes condense during cell division	Lspé
A 3173	chromosomal	cromosómico	chromosomique	cromosomico	cromossômico o cromossomo / *	Of or relating to chromosomes	Lspé
N 12413	chromosome	cromosoma	chromosome	cromosoma	cromossoma / cromosoma	(biology, cytology) A structure in the cell nucleus that contains DNA (...)	Lspé
A 11345	chronic	crónico	chronique	cronico	crônico	Prolonged or slow to heal	Lspé
N 1712	cigarette	cigarro	cigarette	sigaretta	cigarro	Tobacco, marijuana, or other substances, in a thin roll wrapped with paper, intended to be smoked	Lspé
V 1694	circulate	circular	circuler	circolare	circular	To move in circles or through a circuit	Lspé
N 3135	circulation	circulación	circulation	circolazione	circulação	The movement of the blood in the blood-vascular system (...)	Lspé
N 2032	city	ciudad	ville	città	cidade	A large settlement, bigger than a town	LST

N 11645	class	clase	classe	classe	classe	A group, collection, category or set sharing characteristics or attributes	LST
N 6754	classification	clasificación	classification	classificazione	classificação	The act of forming into a class or classes ; a distribution into groups (...)	LST
V 3785	classify	clasificar	classifier / classer	classificare	classificar	To identify by or divide into classes ; to categorize	LST
A 5416	clear	claro	clair	chiaro	claro	Free of ambiguity or doubt	LST
N 1673	clearance	aclaramiento	clairance	clearance	clearance	(medicine) The removal of harmful substances from the blood	Lspé
N 1173	cleavage	* clivaje	clivage	clivaggio	clivagem	1) cleavage : (chemistry) The splitting of a large molecule into smaller ones	Lspé
N 1173	cleavage	segmentación	segmentation	segmentazione	segmentação	2) cleavage : (biology) The repeated division of a cell into daughter cells after mitosis	Lspé
N 3250	clinic	clínica	clinique	clinica	clínica	A medical facility, such as a hospital, especially one for the treatment and diagnosis of outpatients	Lspé
A 32064	clinical	clínico	clinique	clinico	clínico	Dealing with how to practically manage patients	Lspé
N 2177	clinician	clínico	clinicien	clinico	clínico	A healthcare provider	Lspé
N 8987	clone	clon	clone	clone	clone	A living organism (originally a plant) produced asexually from a single ancestor, to which it is genetically identical	Lspé
V 5168	clone	clonar	cloner	clonare	clonar	To create a clone	Lspé
A 3045	close	marcado	marqué	marcato	marcado	Marked, evident	LST
N 6726	cluster	* cluster	cluster / agrégat	cluster	agregado	1) cluster : (chemistry) An ensemble of bound atoms or molecules that is intermediate in size between a molecule and a bulk solid	Lspé
N 6726	cluster	conglomerado	amas / grappe	grappolo	o / agrupamento	2) cluster : A group or bunch of several discrete items that are close to each other	LST
V 3415	cluster	agruparse	se regrouper	raggrupparsi / aggrupparsi	agrupar-se	To form a cluster or group	LST
N 4603	cm	cm	cm	cm	cm	(metrology) Symbol for the centimeter (centimeter), an SI unit of length equal to 10 <sup>-2</sup> meters (meters)	LST
V 4878	code	codificar	coder	codificare	codificar	(genetics) To encode a protein	Lspé
N 1760	code	código	code	codice	código	1) code : (programming) Instructions for a computer (...)	LST
N 1760	code	código	code	codice	código	2) code : A set of rules for converting information into another form or representation	LST
A 1699	coding	* codificante	* codant	* codificante	* codificante	(genetics) Which has the same base sequence as the RNA transcript produced	Lspé
N 5254	codon	codón	codon	codone	códon	A sequence of three adjacent nucleotides (...)	Lspé
N 5276	coefficient	coeficiente	coefficient	coefficiente	coeficiente	(mathematics) A constant by which an algebraic term is multiplied	LST
A 2608	cognitive	cognitivo	cognitif	cognitivo	cognitivo	Relating to the part of the mental functions that deals with logic, as opposed to affective which deals with emotions	Lspé
N 5406	cohort	cohorte	cohorte	coorte	coorte	(statistics) A demographic grouping of people, especially those in a defined age group, or having a common characteristic	LST



A 2043	cold	frío	froid	freddo	frio	Having a low temperature	Lspé
N 4456	collagen	colágeno	collagène	collageno	colágeno	(biochemistry) Any of more than 28 types of glycoprotein that forms elongated fibers, usually found in the extracellular matrix of connective tissue	Lspé
N 2528	colleague	colega	collègue	collega	colega	A fellow member of a profession, staff, academic faculty or other organization ; an associate	LST
V 8188	collect	recoger	collecter / recueillir	raccogliere	coletar / recolher	To gather together; amass	LST
N 5841	collection	colección	collection	collezione	colecção	A set of items or amount of material procured or gathered together	LST
N 1920	college	facultad	faculté	facoltà	faculdade	A specialized division of a university	LST
N 2948	colon	colon	côlon	colon	cólon / colo	(anatomy) The last part of the digestive system	Lspé
N 2552	colony	colonia	colonie	colonia	colônia	A group of organisms of same or different species living together in close association	Lspé
N 2466	color	color	couleur	colore	cor	A particular set of visible spectral compositions, perceived or named as a class	LST
A 2776	colorectal	colorectal	colorectal	* coloretale	colorretal	Of or pertaining to the colorectum	Lspé
N 1899	column	columna	colonne	colonna	coluna	1) column : (chemistry) An object used to separate the different components of a liquid or to purify chemical compounds	Lspé
N 1899	column	columna	colonne	colonna	coluna	2) column : A vertical line of entries in a table, usually read from top to bottom	LST
N 8418	combination	combinación	combinaison	combinazione	combinação	The act of combining, the state of being combined or the result of combining	LST
V 4979	combine	combinar	combiner	combinare	combinar	To bring (two or more things or activities) together; to unite	LST
A 2650	combined	combinado	combiné	combinato	combinado	Resulting from the addition of several sources, parts, elements, aspects, etc. able to be united together, to converge.	LST
V 2735	come	venir	venir	venire	vir	To appear, to manifest itself	LST
N 1983	comment	comentario	commentaire	commento	comentário	A (...) remark	LST
N 3606	committee	comité	comité	comitato	comitê	A group of persons convened for the accomplishment of some specific purpose, typically with formal protocols	LST
A 6422	common	común	commun	comune	comum	1) common : Shared	LST
A 6422	common	común	commun	comune	comum	2) common : Occurring or happening regularly or frequently; usual	LST
N 1420	communication	comunicación	communication	comunicazione	comunicação	1) communication : The concept or state of exchanging data or information between entities	LST
N 1420	communication	comunicación	communication	comunicazione	comunicação	2) communication : The act or fact of communicating anything ; transmission	LST
N 7635	community	comunidad	communauté	comunità	comunidade	A group sharing a common understanding and often the same language, manner, tradition	LST
A 2985	comparable	comparable	comparable	comparabile	comparável	Able to be compared	LST
A 4609	comparative	comparativo	comparatif	comparativo	comparativo	Using comparison as a method of study, or founded on something using it	LST
V 33984	compare	comparar	comparer	comparare	comparar	To assess the similarities and differences between two or more things (...)	LST

N	19666	comparison	comparación	comparaison	comparazione	comparação	An evaluation of the similarities and differences of one or more things relative to some other or each other	LST
N	1900	compartment	compartimento	compartiment / loge	compartimento	compartiment o	(anatomy) A region in the body, delimited by a biological membrane	Lspé
V	4331	compete	competir	se concurrencer	competere	competir	To contend emulously; to seek or strive for the same thing, position, or reward for which another is striving; to contend in rivalry	Lspé
V	1942	complement	completar	compléter	completare	completar	To complete (...)	Lspé
A	2040	complementary	complementario	complémentaire	complementare	complementario	(genetics) Of the specific pairings of the bases in DNA and RNA	Lspé
V	4827	complete	completar	compléter	completare	completar	To make whole or entire	LST
A	7999	complete	completo	complet	completo	completo	With all parts included ; with nothing missing ; finished	LST
N	10530	complex	complejo	complexe	complesso	complexo	(chemistry) A structure consisting of a central atom or molecule weakly connected to surrounding atoms or molecules	Lspé
A	8135	complex	complejo	complexe	complesso	complexo	Intricate ; composite ; not simple	LST
N	2170	complexity	complejidad	complexité	complessità	complexidade	The state of being complex ; intricacy ; entanglement	LST
N	1598	compliance	cumplimiento	observance	compliance	cumprimento	(medicine) The accuracy with which a patient follows an agreed treatment plan	Lspé
N	4238	complication	complicación	complication	complicazione	complicação	(medicine) A disease or diseases, or adventitious circumstances or conditions, coexistent with and modifying a primary disease, but not necessarily connected with it	Lspé
N	11957	component	componente	composant	componente	componente	A smaller, self-contained part of a larger entity (...)	LST
V	1690	compose	componer	composer	comporre	compor	To make up the whole; to constitute	LST
N	3983	composition	composición	composition	composizione	composição	The general makeup of something	LST
N	3657	compound	compuesto	composé	composto	composto	A pure chemical substance consisting of two or more different chemical elements that can be separated into simpler substances by chemical reactions	Lspé
A	2611	comprehensive	exhaustivo	exhaustif	esaustivo	abrangente	Broadly or completely covering, including a large proportion of something	LST
V	3167	comprise	comprender	comprendre	comprendere	compreender	To be made up of; to consist of	LST
A	2362	computational	computacional	computational	computazionale	computacional	Of or relating to computation	LST
V	3372	compute	calcular	calculer	calcolare	calcular	To reckon or calculate	LST
N	3535	computer	computador	ordinateur	computer	computador	A programmable electronic device that performs mathematical calculations and logical operations, especially one that can process, store and retrieve large amounts of data very quickly (...)	LST
V	2216	conceive	concebir	concevoir	concepire	conceber	To develop an idea; to form in the mind; to plan; to devise; to originate	LST
N	22696	concentration	concentración	concentration	concentrazione	concentração	The proportion of a substance in a whole	Lspé

N 3618	concept	concepto	concept	concetto	conceito	An understanding retained in the mind, from experience, reasoning and/or imagination	LST
N 2808	concern	preocupación / inquietud	préoccupation / inquiétude	preoccupazione / inquietudine	preocupação / inquietude	The expression of solicitude, anxiety, or compassion toward a thing or person	Lspé
V 2724	concern	concernir / pertenecer	concerner	concernere	concernir / pertencer	To relate or belong to; to have reference to or connection with; to affect the interest of; to be of importance to	LST
V 3116	conclude	concluir	conclure	concludere	concluir	To come to a conclusion, to a final decision	LST
N 2966	conclusion	conclusión	conclusion	conclusione	conclusão	The end, finish, close or last part of something	LST
N 9803	condition	condición	condition	condizione	condição	1) condition : A requirement, term or requisite	LST
N 9803	condition	condición	condition	condizione	condição	2) condition : The health status of a medical patient	Lspé
V 7558	conduct	conducir / efectuar / realizar	conduire / effectuer / réaliser / mener	condurre / effettuare / svolgere	conduzir / efetuar / realizar	To carry out (something organized)	LST
N 4258	confidence	confianza	confiance	confidenza	confiança	(probability) Likelihood something is reliable	LST
V 9741	confirm	confirmar	confirmer	confermare	confirmar	To assure the accuracy of previous statements	LST
N 1654	conformation	conformación	conformation	conformazione	conformação	(chemistry) The spatial arrangement of the atoms in a molecule as a result of rotation about a single bond	Lspé
V 1782	conjugate	conjugar	conjuguer	coniugare	conjugar	(biology) To temporarily fuse, exchanging or transferring DNA	Lspé
V 1634	connect	conectar	connecter	connettere	conetar	To join (...) to be a link between two objects, thereby attaching them to each other	LST
A 1596	consecutive	consecutivo	consécutif	consequente	consecutivo	Having some logical sequence	LST
N 4104	consensus	consenso	consensus	consenso	consenso	General agreement among the members of a given group or community, each of which exercises some discretion in decision-making and follow-up action	LST
N 2739	consent	consenso	consentement	consenso	consentimento	Voluntary agreement or permission	Lspé
N 3997	consequence	consecuencia	conséquence	conseguenza	consequência	That which follows something on which it depends ; that which is produced by a cause	LST
N 2573	conservation	conservación	conservation	conservazione	conservação	(biology) The fact that a proportion of similar or identical nucleic acid sequences or protein structures have been maintained within given molecules or across species over a period of time	Lspé
V 8402	conserve	conservarse	se conserver	conservarsi	conservar-se	(physics, chemistry) To remain unchanged during a process	Lspé
V 17392	consider	considerar	considérer	considerare	considerar	To think about seriously	LST
A 2161	considerable	considerable	considérable	considerabile / considerevole	considerável	Significant, large in amount	LST
N 1237	consideration	consideración	considération	considerazione	consideração	1) consideration : The thought process of considering, of taking everything into account	LST

N 1237	consideration	consideración	considération	considerazione	consideração	2) consideration : Something considered as a reason or ground for a (possible) decision	LST
V 6521	consist	consistir	consister	consistere	consistir	To be composed, formed, or made up of	LST
A 8266	consistent	consistente	consistant	consistente	consistente	(logic) Of a set of statements, such that no contradiction logically follows them	LST
A 3748	constant	constante	constant	costante	constante	Unchanged through time or space	LST
V 1865	constitute	constituir	constituer	costituire	constituir	To make up; to compose; to form	LST
N 1852	constraint	constricción	contrainte	costrizione	constricção	A restriction	LST
N 1928	construct	construcción	construction	costruzione	construção	(genetics) A segment of nucleic acid, created artificially, for transplantation into a target cell or tissue	Lspé
V 5815	construct	construir	construire	costruire	construir	To build or form by assembling parts	LST
N 2284	construction	construcción	construction	costruzione	construção	The process of constructing	LST
N 2917	consumption	consumo	consommation	consumo	consumo	The act of eating, drinking or using	Lspé
N 4457	contact	contacto	contact	contatto	contato	The act of touching physically ; being in close association	Lspé
V 29062	contain	contener	contenir	contenere	conter	To hold inside	LST
N 3350	content	contenido	contenu	contenuto	conteúdo	1) content : The amount of material contained ; contents	LST
N 3350	content	contenido	contenu	contenuto	conteúdo	2) content : Subject matter ; substance	LST
N 4780	context	contexto	contexte	contesto	contexto	The surroundings, circumstances, environment, background or settings that determine, specify, or clarify the meaning of an event or other occurrence	LST
V 3191	continue	continuar	continuer	continuare	continuar	To proceed with (doing an activity);To prolong (an activity)	LST
A 3472	continuous	continuo	continuel	continuo	contínuo	Without break, cessation, or interruption	LST
N 624	contraction	contracción	contraction	contrazione	contração	1) contraction : (biology) A shortening of a muscle when it is used	Lspé
N 624	contraction	contracción	contraction	contrazione	contração	2) contraction : A reversible reduction in size	LST
N 624	contraction	contracción	contraction	contrazione	contração	3) contraction : (medicine) A strong and often painful shortening of the uterine muscles prior to or during childbirth	Lspé
N 9295	contrast	contraste	contraste	contrasto	contraste	A difference between two objects, people or concepts	LST
V 9209	contribute	contribuir	contribuer	contribuire	contribuir	To give something, that is or becomes part of a larger whole	LST
N 8550	contribution	contribución	contribution	contribuzione	contribuição	The taking part, often with the idea that it has lead to (scientific etc.) progress.	LST
N 39642	control	control	contrôle	controllo	controle	A separate group or subject in an experiment against which the results are compared where the primary variable is low or non-existent	LST
V 8930	control	controlar	contrôler	controllare	controlar	To verify or check	LST
A 2347	controlled	controlado	contrôlé	controllato	controlado	Resulting from a comparison with control samples	LST
A 2807	conventional	convencional	conventionnel	convenzionale	convencional	Following generally accepted principles	LST
V 2168	convert	convertir	convertir	convertire	converter	To transform or change into another form, substance, state, or product	Lspé
V 2519	coordinate	coordinar	coordonner	coordinare	coordenar	To synchronize (activities)	LST

N 2033	coordination	coordinación	coordination	coordinazione	coordenação	The resulting state of working together ; cooperation ; synchronization	LST
N 4277	copy	copia	copie	copia	cópia	(genetics) The result of gene or chromosomal duplication	Lspé
N 1817	cord	cuerda / cordón / médula	corde / cordon / moelle	corda / cordone / midollo	cordão / medula	(anatomy) Any structure having the appearance of a cord (...)	Lspé
N 1276	core	muestra	échantillon	campione	amostra	1) core : (medicine) A tiny sample of organic material obtained by means of a fine-needle biopsy	Lspé
N 1276	core	núcleo	cœur	nucleo	núcleo	2) core : (biochemistry) The central part of a protein structure consisting in mostly hydrophobic aminoacids	Lspé
A 7306	coronary	coronario	coronaire	coronario	coronário	Encircling something, especially with regard to the arteries or veins of the heart	Lspé
V 1933	correct	corregir	corriger	correggere	corrigir	To make something that was not valid become right. To remove error	LST
A 2418	correct	correcto	correct	corretto	correto	Free from error	LST
N 1739	correction	corrección	correction	correzione	correção	A substitution for an error or mistake	LST
N 1753	correlate	correlato	corrélat	correlato	correlato	Either of a pair of things related by a correlation ; a correlative	LST
V 5933	correlate	correlacionar	corréler	correlare	correlacionar / correlatar	To compare things and bring them into a relation having corresponding characteristics	LST
N 14056	correlation	correlación	corrélation	correlazione	correlação	One of the several measures of the linear statistical relationship between two random variables (...)	LST
V 6476	correspond	corresponder	correspondre	corrispondere	corresponder	To be equivalent or similar in character, quantity, quality, origin, structure, function etc.	LST
A 4740	corresponding	correspondiente	correspondant	corrispondente	correspondente	That have a similar relationship	LST
N 2376	cortex	córtex / corteza	cortex	corteccia	córtex / córtice	(anatomy) The outer layer of an internal organ or body structure, such as the kidney or the brain	Lspé
N 8395	cost	coste / costo	cout	costo	custo	Amount of money, time, etc. that is required or used	LST
N 3786	count	número	nombre	numero	número	The result of a tally that reveals the number of items in a set ; a quantity counted	LST
V 2633	count	contar	compter	contare	contar	To determine the number (of objects in a group)	LST
N 9181	country	país	pays	paese	país	(...) a political entity asserting ultimate authority over a geographical area	LST
V 2125	couple	acoplar	coupler	accoppiare	acoplar	To join (two things) together, or (one thing) to (another)	LST
N 5702	course	curso	cours	decorso	curso	A sequence of events	LST
V 2940	cover	cubrir	couvrir	coprire	cobrir	To extend over a given period of time or range, to occupy, to stretch over a given area	LST
N 1893	coverage	cobertura	couverture	copertura	cobertura	(genetics) The average number of reads representing a given nucleotide in the reconstructed sequence	Lspé
V 4927	create	crear	créer	creare	criar	To design (...)	LST
N 9937	criterion	criterio	critère	criterio	critério	A standard or test by which individual things or people may be compared and judged	LST

A	3279	critical	crítico	critique	critico	crítico	1) critical : Extremely important	LST
A	3279	critical	crítico	critique	critico	crítico	2) critical (medicine) of a patient condition that requires urgent treatment in an intensive care or critical care medical facility	Lspé
A	1559	cross	cruzado	croisé	incrociato	cruzado	Made in an opposite direction, or an inverse relation; mutually inverse; interchanged	LST
V	1919	cross	atravesar	traverser	attraversare	atravesar	To go from one side to the other	LST
A	1668	crucial	crucial	crucial	cruciale	crucial	Being essential or decisive for determining the outcome or future of something	LST
V	2913	culture	cultivar	cultiver	coltivare	cultivar	(biology) To maintain in an environment suitable for growth	Lspé
N	17379	culture	cultivo	culture	coltura	cultura	(microbiology) The process of growing a bacterial or other biological entity in an artificial medium	Lspé
A	2427	cultured	cultivado	cultivé	coltivato	cultivado	Artificially developed	Lspé
A	10038	current	actual	actuel	attuale	atual	Generally accepted, used, practiced, or prevalent at the moment	LST
N	4819	curve	curva	courbe	curva	curva	(geometry) A one-dimensional figure of non-zero length; the graph of a continuous map from a one-dimensional space.	LST
N	12834	cycle	ciclo	cycle	ciclo	ciclo	A process that returns to its beginning and then repeats itself in the same sequence	LST
N	1959	cyclin	* ciclina	* cycline	ciclina	ciclina	(biochemistry) Any of a group of proteins whose presence in a cell varies through its cycle, and that regulates this cycle by forming a complex with kinase	Lspé
N	1568	cysteine	cistina	cystéine	cisteina	cisteína	(biochemistry) A sulfur-containing nonessential amino acid C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub> S found in most animal proteins; it readily oxidizes to cystine	Lspé
N	1838	cytochrome	citocromo	cytochrome	citocromo	citocromo	Any of various related hemeproteins found in the cells of most organisms, which are an important part of cell respiration	Lspé
N	7779	cytokine	citosina	cytokine	citochina	citosina	(biochemistry) Any of various small regulatory proteins that regulate the cells of the immune system	Lspé
N	1691	cytoplasm	citoplasma	cytoplasme	citoplasma	citoplasma	(cytology) The contents of a cell except for the nucleus (...)	Lspé
A	2764	cytoplasmic	citoplasmático	cytoplasmique / cytoplasmatique	citoplasmatico	citoplasmático	Of or pertaining to cytoplasm	Lspé
A	1675	cytotoxic	* citotóxico	cytotoxique	citotossico	citotóxico	Of or relating to cytotoxicity	Lspé
A	3871	daily	cotidiano	quotidien	quotidiano	cotidiano	That occurs every day	LST
N	4816	damage	daño	dommage	danno	dano	Injury or harm; the condition or measure of something not being intact	Lspé
N	16643	database	base de datos	banque de données	base di dati	banco de dados	(computing) A collection of (usually) organized information in a regular structure (...)	LST
N	5252	dataset	conjunto de datos	ensemble de données	dataset	conjunto de dados	(statistics) As set of data to be analyzed	LST
N	86651	datum	dato	donnée	dato	dado	A measurement of something on a scale (...)	LST
N	26384	day	día	jour	giorno	día	Any period of 24 hours	LST
N	13137	death	muerte	mort	morte	morte	The cessation of life and all associated processes ; the end of an organism's existence	Lspé

						(...)	
N 1739	decade	década / decenio	décade / décennie	decade / decennio	década / decênio	A period of ten years	LST
N 5219	decision	decisión	décision	decisione	decisão	A choice or judgment	LST
V 3523	declare	declarar	déclarer	dichiarare	declarar	To affirm or state something emphatically	LST
N 987	decline	declive	déclin	declino	declínio	1) decline : Downward movement, fall	LST
N 987	decline	declive	déclin	declino	declínio	2) decline : A reduction or diminution of activity	Lspé
N 6556	decrease	decrecimiento / disminución	décroissance / diminution	decremento / diminuzione	decréscimo / diminuição	An amount by which a quantity is decreased	LST
V 11812	decrease	decrecer / disminuir	décroître / diminuer	decrescere / diminuire	descrecer / diminuir	Of a quantity, to become smaller	LST
N 3588	defect	defecto	défaut	difetto	defeito	A fault or malfunction	LST
N 1580	defense	defensa	défense	difesa	defesa	The action of defending, of protecting from attack, danger or injury	Lspé
N 2783	deficiency	deficiencia	déficit / carence	deficienza	deficiência	An insufficiency, especially of something essential to health	Lspé
A 2664	deficient	deficiente	déficient	deficiente	deficiente	Lacking something essential	Lspé
V 13907	define	definir	définir	definire	definir	To express the essential nature of something	LST
N 4216	definition	definición	définition	definizione	definição	A statement of the meaning of a word (...)	LST
N 3513	degradation	degradación	dégradation	degradazione	degradação	A deleterious change in the chemical structure, physical properties or appearance of a material from natural or artificial exposure	Lspé
N 6234	degree	grado *	degré	grado	grau	The amount that an entity possesses a certain property ; relative intensity, extent	LST
N 1700	dehydrogenase	deshidrogenasa	* déshydrogénase	deidrogenasi	desidrogenase	(biochemistry) Any of several enzymes that catalyze the removal of hydrogen (a proton) from biological compounds	Lspé
N 1960	delay	retraso / demora	retard	ritardo	atraso / demora	A period of time before an event occurs; the act of delaying (...)	LST
N 4732	deletion	delección	* délétion	delezione	deleção	(genetics) A mutation in which a gene, or other section of DNA, is removed from a chromosome	Lspé
V 1951	deliver	administrar	administrer	amministrare / sommministrare	administrar	(medicine) To administer a drug	Lspé
N 1452	delivery	administración	administration	somministrazione	administração	1) delivery (medicine) Administration of a drug	Lspé
N 1452	delivery	entrega	distribution	somministrazione	entrega	2) delivery (genetics) Process of introducing foreign DNA into host cells	Lspé

N 1452	delivery	parto	accouchement	parto	parto	3) delivery : The act of giving birth	Lspé
A 2632	demographic	demográfico	démographique	demografico	demográfico	Of, or pertaining to demography	LST
V 17637	demonstrate	demonstrar	démontrer	dimostrare	demonstrar	To show the steps taken to create a logical argument or equation	LST
A 1917	dendritic	dendrítico	dendritique	dendritico	dendrítico	Of or pertaining to or possessing dendrites	Lspé
N 7925	density	densidad	densité	densità	densidade	The ratio of one quantity to that of another quantity	LST
N 4413	department	departamento	département	dipartimento	departamento	A subdivision of an organization	LST
V 3871	depend	dependen	dépendre	dipendere	dependen	To rely on for support; to be conditioned or contingent; to be connected with (...)	LST
N 948	dependence	dependencia	dépendance	dipendenza	dependência	1) dependence : The state of being dependent, of relying upon another	LST
N 948	dependence	dependencia	dépendance	dipendenza	dependência	2) dependence : An irresistible physical or psychological need, especially for a chemical substance	Lspé
A 14329	dependent	dependiente	dépendant	dipendente	dependente	Relying upon, depending upon	LST
N 2566	depression	depresión	dépression	depressione	depressão	1) depression : (psychology) In psychotherapy and psychiatry, a state of mind producing serious, long-term lowering of enjoyment of life or inability to visualize a happy future	Lspé
N 2566	depression	depresión	dépression	depressione	depressão	2) depression : (biology, physiology) A lowering, in particular a reduction in a particular biological variable or the function of an organ, in contrast to elevation	Lspé
V 11878	derive	derivar	dériver	derivare	derivar	(chemistry) To create (a compound) from another by means of a reaction.	Lspé
V 25896	describe	describir	décrire	descrivere	descrever	To represent in words	LST
N 2845	description	descripción	description	descrizione	descrição	A sketch or account of anything in words ; a portraiture or representation in language ; an enumeration of the essential qualities of a thing or species	LST
N 11717	design	diseño / concepción / elaboración	conception / élaboration	disegno / concezione / elaborazione	desenho / elaboração / concepção	A plan (with more or less detail) for the structure and functions of an artifact, building or system	LST
V 7013	design	diseñar / concebir / elaborar	élaborer / concevoir	disegnare / concepire / elaborare	desenhar / conceber / elaborar	To plan and carry out (...)	LST
N 4247	detail	detalle / pormenor detallado /	détail	dettaglio	detalhe / pormenor detalhado /	A part distinct from the whole	LST
A 2926	detailed	pormenorizado	détaillé	dettagliato	pormenorizado	Characterized by attention to detail and thoroughness of treatment	LST
V 19136	detect	detectar	détecter	rilevare	detectar	To discover or find by careful search, examination, or probing	Lspé
A 2388	detectable	detectable	délectable	rilevabile	detectável	Which can be detected, noticeable	Lspé
N 10594	detection	detección	détection	rilevamento	detecção	The finding out of a constituent, a signal, an agent or the like, mostly by means of a specific device or method	Lspé
N 1295	determinant	factor	facteur	fattore	fator	1) determinant : A determining factor; an element that determines the nature of	LST



N 1295	determinant	determinante	déterminant	determinante	determinante	something	
N 3827	determination	determinación	détermination	determinazione	determinação	2) determinant (biology) A substance that causes a cell to adopt a particular fate	Lspé
V 24697	determine	determinar	déterminer	determinare	determinar	The act, process, or result of any accurate measurement, as of length, volume, weight, intensity, etc. (...)	LST
V 17407	develop	desenvolverse / desarrollarse	se développer	svilupparsi	desenvolver-se	To ascertain definitely; to figure out, find out, or conclude by analyzing, calculating, or investigating	LST
N 24217	development	desarrollo	développement	sviluppo	desenvolvimento	To progress through a sequence of stages	LST
A 3275	developmental	de desarrollo	développemental	dello sviluppo	desenvolvimento de	The process by which a mature multicellular organism or part of an organism is produces by addition of new cells	Lspé
N 3573	deviation	desviación	écart	deviazione / scarto	desvio	Related to development	Lspé
N 2199	device	aparajeo	appareil	apparecchio	aparelho	(statistics) For interval variables and ratio variables, a measure of difference between the observed value and the mean	LST
N 8628	diabetes	diabetes	diabète	diabete	diabetes	Any piece of equipment made for a particular purpose, especially a mechanical or electrical one.	LST
A 2751	diabetic	diabético	diabétique	diabetico	diabético	A group of metabolic disease whereby a person (or other animal) has high blood sugar due to an inability to produce or inability to metabolize sufficient quantities of the hormone insulin	Lspé
V 3529	diagnose	diagnosticar	diagnostiquer	diagnosticare	diagnosticar	Of or pertaining to diabetes	Lspé
N 11690	diagnosis	diagnóstico	diagnostic	diagnosi	diagnóstico	(medicine) To determine which disease is causing a sick person's signs and symptoms; to find the diagnosis	Lspé
A 5562	diagnostic	diagnóstico	diagnostique	diagnostico	diagnóstico	(medicine) The identification of the nature and cause of an illness	Lspé
N 2299	diameter	diámetro	diamètre	diametro	diâmetro	Of, or relating to diagnosis	LST
V 2750	die	morir	mourir	morire	morror	The length between two points on the circumference of a circle (...)	LST
N 5408	diet	dieta	diète	dieta	dieta	To stop living; to become dead; to undergo death	Lspé
A 3191	dietary	dietético	diététique	dietetico	dietético	A controlled regimen of food and drink (...)	Lspé
V 6164	differ	diferir	différer	differire	diferir	Of, or relating to diet	Lspé
N 32548	difference	diferencia	différence	differenza	diferença	Not to have the same traits, characteristics	LST
A 41196	different	diferente	différent	differente	diferente	A characteristic of something that makes it different from something else	LST
A 2389	differential	diferencial	différentiel	differenziale	diferencial	Not the same	LST
						1) differential : (mathematics) Of, or relating to differentiation, or the differential calculus	LST

A 2389	differential	diferencial	différentiel	differenziale	diferencial	2) differential : Dependent on, or making a difference; distinctive	LST
V 3096	differentiate	diferenciar	différencier	differenziare	diferenciar	(biology) To produce distinct organs or to achieve specific functions by a process of development (...)	Lspé
N 6794	differentiation	diferenciación	différenciation	differenziazione	diferenciação	The gradual formation or production of organs or parts by a process of evolution or development (...)	Lspé
A 4936	difficult	difícil	difficile	difficile	difícil	Hard, not easy	LST
N 1327	difficulty	dificultad	difficulté	difficoltà	dificuldade	1) difficulty : An obstacle that hinders achievement of a goal	LST
N 1327	difficulty	dificultad	difficulté	difficoltà	dificuldade	2) difficulty : The state of being difficult, or hard to do	LST
V 866	digest	digerir	digérer	digerire	digerir	1) digest : (chemistry) To expose to a gentle heat in a boiler or matress, as a preparation for chemical operations	Lspé
V 866	digest	digerir	digérer	digerire	digerir	2) digest : To separate (the food) in its passage through the alimentary canal into the nutritive and nonnutritive elements (...)	Lspé
V 1956	dilute	diluir	diluer	diluire	diluir	To make thinner by adding solvent to a solution; especially by adding water	Lspé
N 2765	dilution	dilución	dilution	diluizione	diluição	A solution that has had additional solvent, such as water, added to it into order to make it less concentrated	Lspé
N 1366	dimension	dimensión	dimension	dimensione	dimensão	1) dimension : A measure of spatial extent in a particular direction, such as height, width or breadth, or depth	LST
N 1366	dimension	dimensión	dimension	dimensione	dimensão	2) dimension : A single aspect of a given thing	LST
V 3349	direct	dirigir	diriger	dirigere	dirigir	To manage, control, steer	LST
A 7494	direct	directo	direct	diretto	direto	Straight, constant, without interruption	LST
N 1346	direction	dirección	direction	direzione	direção	1) direction : A theoretical line (physically or mentally) followed from a point of origin or towards a destination	LST
N 1346	direction	dirección	direction	direzione	direção	2) direction : An general trend for future action	LST
N 2825	disability	discapacidad / incapacidad	incapacité	disabilità	deficiência	State of being disabled (...)	Lspé
N 846	discharge	alta / salida	sortie	uscita	alta / saída	1) discharge : (medicine) the act of releasing an inpatient from hospital	Lspé
N 846	discharge	corrimiento	perte / écoulement	perdita / versamento	corrimento	2) discharge : (medicine) pus or exudate (other than blood) from a wound or orifice, usually due to infection or pathology	Lspé
V 1640	discover	descubrir	découvrir	discoprire / scoprire	descobrir / desvendar	To find or learn something for the first time	LST
N 2431	discovery	descubrimiento	découverte	scoperta	descoberta	Something discovered	LST
V 3928	discuss	discutir	discuter	discutere	discutir	To converse or debate concerning a particular topic	LST
N 4678	discussion	discusión	discussion	discussione	discussão	Conversation or debate concerning a particular topic	LST
N 46377	disease	enfermedad	maladie	malattia	enfermidade /	An abnormal condition of the body or mind that causes discomfort or dysfunction (...)	Lspé

N 9165	disorder	desorden / trastorno	désordre / trouble	disordine / disturbo	doença desordem / transtorno	(medicine) A physical or psychological malfunction	Lspé
N 1796	display	visualización / despliegue	affichage	visualizzazione	visualização	(computing) The presentation of information for visual or tactile reception	LST
V 4810	display	mostrar / presentar	montrer / présenter	mostrare / presentare	mostrar / apresentar	To show (...) to manifest	LST
N 1749	disruption	ruptura	rupture	rottura	ruptura	An interruption to the regular flow or sequence of something	LST
A 1865	distal	distal	distal	distale	distal	(anatomy) Remote from the point of attachment or origin; as, the distal end of a bone or muscle	Lspé
N 3267	distance	distancia	distance	distanza	distância	1) distance : The amount of space between two points (...)	LST
N 3267	distance	distancia	distance	distanza	distância	2) distance : Remoteness in succession or relation	LST
A 6136	distinct	distinto	distinct	distinto	distinto	Different from one another	LST
V 2592	distinguish	distinguir	distinguer	distinguere	distinguir	To see someone or something as different from others	LST
V 2978	distribute	distribuir / repartir	distribuer / répartir	distribuire / ripartire	distribuir / repartir	To scatter or spread	LST
N 8752	distribution	distribución	distribution	distribuzione	distribuição	1) distribution : (mathematics, statistics) The set of relative likelihoods that a variable will have a value in a given interval	LST
N 8752	distribution	distribución	distribution	distribuzione	distribuição	2) distribution : The frequency of occurrence or extent of existence	LST
N 1637	district	distrito	district	distretto	distrito	An administrative division of an area	LST
N 2147	divergence	divergencia	divergence	divergenza	divergência	The process in which two or more populations accumulate genetic changes (mutations) through time	Lspé
A 2316	diverse	diverso	divers	diverso	diverso	Consisting of many different elements	LST
N 2839	diversity	diversidad	diversité	diversità	diversidade	The quality of being diverse or different ; difference or unlikeness	LST
V 1818	divide	dividir	diviser	dividere	dividir	1) divide : To split or separate (something) into two or more parts	LST
V 1818	divide	dividir	diviser	dividere	dividir	2) divide : (arithmetic) To calculate the number (the quotient) by which you must multiply one given number (the divisor) to produce a second given number (the dividend)	LST
N 2304	division	división	division	divisione	divisão	The act or process of dividing anything	Lspé
N 5534	DNA	ADN	ADN	DNA	ADN	(biochemistry) A biopolymer of deoxyribonucleic acids (...)	Lspé
V 15205	do	hacer	faire	fare	fazer	To perform; to execute	LST
N 10707	doctor	médico	médecin	medico	médico	A physician ; a member of the medical profession ; one who is trained and licensed to heal the sick (...)	Lspé
V 2683	document	documentar	documenter	documentare	documentar	To furnish with documents or papers necessary to establish facts or give information	LST
N 1745	dog	perro	chien	cane	cão	A mammal, Canis lupus familiaris (...)	Lspé

N 2078	dollar	dolar	dollar	dollaro	dólar	The official designation for currency in some parts of the world, including Canada, Australia, the United States, Hong Kong, and elsewhere	LST
N 23923	domain	dominio	domaine	dominio	domínio	A folded section of a protein molecule that has a discrete function	Lspé
A 1188	dominant	dominante	dominant	dominante	dominante	1) dominant : Predominant, common, prevalent	LST
A 1188	dominant	dominante	dominant	dominante	dominante	2) dominant : (medicine) Designating the follicle which will survive atresia and permit ovulation	Lspé
N 3390	donor	donante	donneur	donatore	doador	One who donates (...)	Lspé
N 1596	dopamine	dopamina	dopamine	dopamina	dopamina	(biochemistry) A neurotransmitter associated with movement, attention, learning, and the brain's pleasure and reward system	Lspé
N 12962	dose	dosis	dose	dose	dose	A measured portion of medicine taken at any one time	Lspé
A 1888	double	doble	double	doppio	dobro	Twice the quantity	LST
A 2522	downstream	corriente abajo	en aval	a valle	à jusante	(biology) Towards the 3' end of a DNA molecule	Lspé
V 2705	draft	redactar	rédigier	redigere	redigir	To write a first version, make a preliminary sketch	LST
V 2381	draw	extraer / sacar / tirar	extraire / tirer	estrarre / tirare	extraire / tirer	To remove or separate or displace	Lspé
V 2414	drive	causar	causer / entrainer	causare / innescare	causar	To cause (a mechanism) to operate	LST
N 6396	drosophila	* drosófila	drosophile	drosophila	drosófila	Any fruit fly of the genus Drosophila	Lspé
N 9897	drug	medicamento	médicament	medicamento	medicamento	1) drug : (pharmacology) A substance used to treat an illness, relieve a symptom, or modify a chemical process in the body for a specific purpose	Lspé
N 9897	drug	droga	drogue	droga	droga	2) drug : A psychoactive substance (...)	Lspé
N 2917	duplication	duplicación	duplication	duplicazione	duplicação	(biochemistry) The act or process of dividing by natural growth or spontaneous action	Lspé
N 5310	duration	duración	durée	durata	duração	An amount of time or a particular time interval	LST
N 2188	dye	tintura / tinte	teinture	tintura	tintura	A colorant, especially one that has an affinity to the substrate to which it is applied	Lspé
A 2206	dynamic	dinámico	dynamique	dinamico	dinâmico	Able to change and adapt	LST
N 2485	dynamics	dinámica	dynamique	dinamica	dinâmica	(mechanics) The branch of mechanics that is concerned with the effects of forces on the motion of objects	Lspé
N 3036	dysfunction	disfunción	dysfonction	disfunzione	disfunção	A failure to function in an expected or complete manner. Usually refers to a disorder in a bodily organ (e.g. erectile dysfunction), a mental disorder, or the improper behavior of a social group	Lspé
A 7294	early	precoz	précoce	precoce	precoce	1) early : (medicine) Having begun to occur ; in its early stages	Lspé
A 7294	early	precoz	précoce	precoce	precoce	2) early : At a time in advance of the usual or expected event	Lspé
A 1970	easy	fácil	facile	facile	fácil	Requiring little skill or effort	LST
A 2342	economic	económico	économique	economico	econômico	Pertaining to an economy	LST

N 1642	edge	borde / orilla	bord	bordo / orlo / margine	bordo / margem	The boundary line of a surface	LST
V 1904	edit	editar	éditer	editare	editar	To be the editor of a publication	LST
N 5084	education	educación	éducation	educazione	educação	The process or art of imparting knowledge, skill and judgment	LST
A 1998	educational	educacional / educativo	éducatif	educativo	educacional / educativo	Of or relating to education	LST
N 60116	effect	efecto	effet	effetto	efeito	The outcome of a cause	Lspé
A 6971	effective	eficaz	efficace	efficace	eficaz	Having the power to produce a required effect or effects	LST
N 2826	effectiveness	eficacia	efficacité	efficacia	eficácia	The property of being effective, of achieving results	LST
N 1910	effector	efector	effecteur	effettore	efetor	(biology) Any small molecule that effects the function of an enzyme by binding to an allosteric site	Lspé
N 5348	efficacy	eficacia	efficacité	efficacia	eficácia	Ability to produce a desired amount of a desired effect	LST
N 3993	efficiency	eficiencia	efficience	efficienza	eficiência	The quality of producing an effect or effects	LST
A 3072	efficient	eficiente	efficient	efficiente	eficiente	Having the power to produce a required effect or effects	LST
N 3509	effort	esfuerzo	effort	sforzo	esforço	The work involved in performing an activity ; exertion	LST
N 978	egg	huevo	œuf	uovo	ovo	1) egg : (zoology) An approximately spherical or ellipsoidal body produced by birds, snakes, insects and other animals, housing the embryo during its development	Lspé
N 978	egg	óvulo	ovule	ovulo	óvulo	2) egg : (biology) The female primary cell, the ovum	Lspé
A 2935	elderly	anciano	âgé	anziano	idoso	Old ; having lived for relatively many years	Lspé
N 2200	electron	electrón	électron	elettrone	elétron	(physics) The subatomic particle having a negative charge and orbiting the nucleus; the flow of electrons in a conductor constitutes electricity.	Lspé
N 2243	electrophoresis	electroforesis	électrophorèse	elettroforesi	eletroforese	A method for the separation and analysis of large molecules (such as proteins) by migrating a colloidal solution of them through a gel	Lspé
N 11356	element	elemento	élément	elemento	elemento	One of the simplest or essential parts or principles (...)	LST
A 3940	elevated	elevado	élevé	elevato	alto	Above a normal level	Lspé
N 1599	elevation	elevación	élévation	elevazione	elevação	The act of raising from a lower place, condition, or quality to a higher	Lspé
V 1679	elicit	elicitar / provocar	éliciter / provoquer	elicitare / provocare	elicitar / provocar	To evoke, educe (...); to generate, obtain, or provoke as a response or answer	LST
V 2444	eliminate	eliminar	éliminer	eliminare	eliminar	To completely destroy (something) so that it no longer exists	Lspé
N 5653	embryo	embrión	embryon	embrione	embrião	In viviparous animals, the young animal's earliest stages in the mother's body	Lspé
A 2831	embryonic	embrionario	embryonnaire	embrionario	embrionário	Of or relating to an embryo	Lspé
N 2090	emergency	emergencia / urgencia	urgence	emergenza / urgenza	emergência / urgência	A situation which poses an immediate risk and which requires urgent attention	Lspé
A 1563	empirical	empírico	empirique	empirico	empírico	Pertaining or based on experience	LST

V 4013	employ	emplear	employer	impiegare permettere / rendere possibile	empregar	To use (...)	LST
V 2676	enable	permitir / hacer posible	permettre / rendre possible	rendere possibile	permitir / possibilitar	To make able (...)	LST
V 9605	encode	codificar	encoder	codificare	codificar	Constitute the code necessary for the biosynthesis of a protein by means of a matrix so as to transcript DNA material	Lspé
N 9772	end	fin	fin	fine	fim	The final point of something in space or time	LST
A 3671	endogenous	endógeno	endogène	endogeno	endógeno	Produced, originating or growing from within	Lspé
A 1835	endometrial	endometrial	* endométrial	endometriale	endometrial	Of or pertaining to the endometrium, the lining of the uterus	Lspé
A 6548	endothelial	endotelial	endothélial	endoteliale	endotelial	Of or pertaining to the endothelium	Lspé
N 4638	energy	energía	énergie	energia	energia	(physics) A quantity that denotes the ability to do work and is measured in a unit dimensioned in mass × distance <sup>2</sup> /time <sup>2</sup> (ML <sup>2</sup> /T <sup>2</sup> ) or the equivalent	Lspé
V 5820	enhance	amplificar	amplifier	amplificare	ampliar	To augment or make something greater	Lspé
A 2401	enhanced	amplificado	amplifié	amplificato	ampliado	Raised up, amplified	Lspé
V 1007	enrich	enriquecer	enrichir	arricchire	enriquecer	1) enrich : (chemistry) To make to rise the proportion of a given constituent	Lspé
V 1007	enrich	enriquecer	enrichir	arricchire	enriquecer	2) enrich : To add new elements, to complete	LST
V 3201	ensure	asegurar / garantizar	assurer / garantir	assicurare / garantire	assegurar / garantir	To make sure or certain of something (usually some future event or condition)	LST
V 2668	enter	entrar	entrer	entrare	entrar	To cause to go (into), or to be received (into) ; to put in ; to insert ; to cause to be admitted	LST
A 3423	entire	entero	entier	intero	inteiro	Whole ; complete	LST
N 3832	entry	entrada	entrée	entrada	entrada	The act of entering	LST
N 1717	envelope	envoltura	enveloppe	* envelope	envelope	(biology) An enclosing structure or cover, such as a membrane	Lspé
N 5150	environment	ambiente	environnement	ambiente	ambiente	All the elements that affect a system or its inputs and outputs	LST
A 4327	environmental	ambiental	environnemental	ambientale	ambiental	Pertaining to one's environment	LST
N 12407	enzyme	enzima	enzyme	enzima	enzima	(biochemistry) A globular protein that catalyzes a biological chemical reaction	Lspé
A 2101	epidemiological	epidemiológico	épidémiologique	epidemiologico	epidemiológico	Of or pertaining to epidemiology	Lspé
N 3077	epidemiology	epidemiología	épidémiologie	epidemiologia	epidemiologia	(sciences) The branch of a science dealing with the spread and control of diseases, viruses, concepts etc. throughout populations or systems	Lspé
N 2102	episode	episodio	épisode	episodio	episódio	An incident or action standing out by itself (...)	Lspé
A 7341	epithelial	epitelial	épithélial	epiteliale	epitelial	Of or pertaining to the epithelium	Lspé
N 3070	epithelium	epitelio	épithélium	epitelio	epitélio	(anatomy) A membranous tissue composed of one or more layers of cells which forms	Lspé

N 2464	epitope	* epítopo	* épitope	epitopo	epítopo	the covering of most internal and external surfaces of the body and its organs (...) (biochemistry) That part of a biomolecule (such as a protein) that is the target of an immune response	Lspé
A 3543	equal	igual	égal	uguale	igual	Exactly identical, having the same value	LST
N 3637	equation	ecuación	équation	equazione	equação	(mathematics) An assertion that two expressions are equal, expressed by writing the two expressions separated by an equal sign; from which one is to determine a particular quantity.	LST
A 1850	equivalent	equivalente	équivalent	equivalente	equivalente	Similar or identical	LST
N 4040	error	error	erreur	errore	erro	1) error : Any alteration in the DNA chemical structure occurring during DNA replication, recombination or repairing	Lspé
N 4040	error	error	erreur	errore	erro	2) error : (statistics) The difference between a measured or calculated value and a true one	LST
N 7479	Escherichia coli	Escherichia coli	Escherichia coli	Escherichia coli	Escherichia coli	A taxonomic species within the family Enterobacteriaceae (...)	Lspé
A 5768	essential	esencial	essentiel	essenziale	essencial	Of high importance	LST
N 5154	EST	EST	EST	EST	EST	(genetics) Expressed Sequence Tag. A short sub-sequence of a cDNA sequence	Lspé
V 6939	establish	establecer	établir	stabilire	estabelecer	To prove and cause to be accepted as true, to demonstrate	LST
N 7891	estimate	estimación	estimation	stima	estimação	A rough calculation or guess	LST
V 10001	estimate	estimar	estimer	stimare	estimar	To calculate roughly (...)	LST
N 1416	estimation	estimación	estimation	stima	estimação	1) estimation : The process of making an estimate	LST
N 1416	estimation	estimación	estimation	stima	estimação	2) estimation : The amount, extent, position, size, or value reached in an estimate	LST
N 5994	estrogen	estrógeno	œstrogène	estrogeno	estrogênio / estrógeno	(biochemistry) Any of a group of steroids that are secreted by the ovaries and function as female sex hormones	Lspé
N 2656	ethanol	etanol	éthanol	etanolo	etanol	(organic chemistry) A simple aliphatic alcohol formally derived from ethane by replacing one hydrogen atom with a hydroxyl group : CH <sub>3</sub> -CH <sub>2</sub> -OH.	Lspé
N 2163	ethic	ética	éthique	etica	ética	A set of principles of right and wrong behavior guiding, or representative of, a specific culture, society, group, or individual	LST
A 1620	ethical	ético	éthique	etico	ético	Of or relating to the accepted principles of right and wrong, especially those of some organization or profession.	LST
N 1638	eukaryote	eucariota / eucarionte	eucaryote	eucariota	eucarioto	Any of the single-celled or multicellular organisms of the taxonomic domain Eukaryota, whose cells contain at least one distinct nucleus	Lspé
A 2627	eukaryotic	eucariótico	* eucaryote	eucariotico	eucariótico	Having complex cells	Lspé
A 2903	European	européo	européen	européo	européu	Related to Europe or the European Union	LST
V 12006	evaluate	evaluar	évaluer	valutare	avaliar	To draw conclusions from examining; to assess	LST
N 10420	evaluation	evaluación	évaluation	valutazione	avaliação	An assessment (...)	LST
N 12379	event	accidente	accident	accidente	acidente	(medicine) An episode of severe health conditions	Lspé

N	19326	evidence	prueba / evidencia	preuve	prova	prova	Facts or observations presented in support of an assertion	LST
A	1832	evident	evidente	évident	evidente	evidente	Obviously true	LST
N	7169	evolution	evolución	évolution	evoluzione	evolução	(biology) The change in the genetic composition of a population over successive generations	Lspé
A	4575	evolutionary	evolutivo	évolutif	evolutivo	evolutivo	Of or relating to evolution	Lspé
V	1952	evolve	evolucionar	évoluer	evolvere	evolver / evolucionar / evoluir	To change; transform	LST
A	2307	exact	exacto	exact	esatto	exato	Precisely agreeing with a standard, a fact, or the truth	LST
N	6013	examination	examen	examen	esame	exame	An inspection by a medical professional to establish the extent and nature of any sickness or injury	Lspé
V	6620	examine	examinar	examiner	esaminare	examinar	1) examine : To check the health or condition of something or someone	Lspé
V	6620	examine	examinar	examiner	esaminare	examinar	2) examine : To observe or inspect carefully or critically	LST
N	6415	example	ejemplo	exemple	esempio	exemplo	Something that is representative of all such things in a group	LST
N	1985	exception	excepción	exception	eccezione	excepção	That which is excepted or taken out from others; a person, thing, or case, specified as distinct, or not included (...)	LST
N	1101	exchange	intercambio	échange	scambio	troca	1) exchange : (biochemistry) The transfer of substances or elements like gas, amino-acids, ions etc. sometimes through a surface like a membrane	Lspé
N	1101	exchange	intercambio	échange	scambio	troca	2) exchange : The action consisting in giving or borrowing something, often with an idea of reciprocity	LST
V	6052	exclude	exclure	exclure	escludere	excluir	To keep out	LST
N	1700	exclusion	exclusión	exclusion	esclusione	exclusão	The act of excluding or shutting out; removal from consideration or taking part	LST
N	4428	exercise	ejercicio	exercice	esercizio	exercício	Physical activity intended to improve strength and fitness	Lspé
V	4702	exhibit	mostrar / presentar	montrer / présenter	mostrare / presentare	mostrar / apresentar	To display or show (...)	LST
V	4533	exist	existir	exister	esistere	existir	to be; have existence; have being or reality	LST
N	1680	existence	existencia	existence	esistenza	existência	The state of being, existing, or occurring; beinghood	LST
A	2619	existing	existente	existant	esistente	existente	That exists, or has existence, especially that exists now	LST
N	7662	exon	* exón	* exon	esone	éxon / exão	(genetics) A region of a transcribed gene present in the final functional RNA molecule	Lspé
V	1585	expand	extender	étendre	estendere	estender	To increase the extent, number, volume or scope	LST
N	1940	expansion	expansión	expansion	espansione	expansão	The act or process of expanding	LST
V	9404	expect	esperar	s'attendre	aspettarsi	esperar	To look for (mentally); to look forward to, as to something that is believed to be about to happen or come (...)	LST
N	5003	experience	experiencia	expérience	esperienza	experiência	A collection of events and/or activities from which an individual or group may gather	Lspé



V 2789	experience	experimentar	expérimenter	sperimentare	experienciar / experimentar	knowledge, opinions, and skills. To observe certain events; undergo a certain feeling or process; or perform certain actions that may alter one or contribute to one's knowledge, opinions, or skills	LST
N 19012	experiment	experimento	expérience	esperimento	experimento	A test under controlled conditions made to either demonstrate a known truth, examine the validity of a hypothesis, or determine the efficacy of something previously untried	LST
A 10731	experimental	experimental	expérimental	sperimentale	experimental	Of or pertaining to an experiment	LST
N 1822	expert	experto / perito	expert	esperto / perito	experto / perito	A person with extensive knowledge or ability in a given subject	LST
V 6763	explain	explicar	expliquer	spiegare	explicar / explicar	To make plain, manifest, or intelligible; to clear of obscurity; to illustrate the meaning of	LST
N 2830	explanation	explicación	explication	spiegazione	explicação	The act or process of explaining	LST
V 3090	explore	explorar	explorer	esplorare	explorar	To examine or investigate something systematically	LST
V 4486	expose	exponer	exposer	esporre	expor	To submit to an active (mostly dangerous) substance like an allergen, ozone, nicotine, solvent, or to any other stress, in order to test the reaction, resistance, etc.	Lspé
N 10688	exposure	exposición	exposition	esposizione	exposição	The condition of being exposed, uncovered or unprotected	Lspé
V 13161	express	expresar	exprimer	esprimere	expressar	1) express: (biochemistry) To translate messenger RNA into a protein or to transcribe deoxyribonucleic acid into messenger RNA	Lspé
V 13161	express	expresar / exprimir	exprimer	esprimere	expressar / exprimir	2) express : To convey or communicate; to make known or explicit	LST
N 73939	expression	expresión	expression	espressione	expressão	The process of translating a gene into a protein	Lspé
V 1702	extend	extender	étendre	estendere	estender	1) extend : To cause to increase in extent	LST
V 1702	extend	extenderse	s'étendre	estendersi	estender-se	2) extend : To possess a certain extent	LST
N 1375	extension	extensión	étendue	estensione	extensão	1) extension : The act of extending or the state of being extended ; a stretching out ; enlargement in breadth or continuation of length ; increase ; augmentation ; expansion	LST
N 1375	extension	extensión	extension	estensione	extensão	2) extension : (weightlifting) An exercise in which an arm or leg is straightened against resistance	Lspé
A 2668	extensive	extensivo	étendu	estensivo	extensivo	In the nature of an extent, wide, widespread	LST
N 4268	extent	medida	mesure	misura	medida	A range of values or locations	LST
A 1315	external	externo	externe	esterno	externo	1) external : (anatomy) Away from the mesial plane of the body ; lateral	Lspé
A 1315	external	externo	externe	esterno	externo	2) external : Provided by something or someone outside of the entity (object, group, company etc.) considered	LST
A 3863	extracellular	* extracelular	extracellulaire	extracellulare	extracelular	Occurring or found outside of a cell	Lspé
N 1201	extract	extracto	extrait	estratto	extrato	1) extract : A decoction, solution, or infusion made by drawing out from any substance (...)	Lspé
N 1201	extract	extracto	extrait	estratto	extrato	2) extract : A solid preparation obtained by evaporating a solution of a drug (...)	Lspé

V 6143	extract	extraer	extraire	estrarre	extrair	To withdraw by expression, distillation, or other mechanical or chemical process	Lspé
N 3083	extraction	extracción	extraction	estrazione	extração	An act of extracting or the condition of being extracted	Lspé
N 2807	eye	ojo	œil	occhio	olho	An organ through which animals see	Lspé
V 3867	facilitate	facilitar	faciliter	facilitare	facilitar	To make easy or easier	LST
N 2398	facility	material / equipamiento / instalación	matériel / équipement / installation	materiale / apparecchiatur a / impianto	material / equipamento / instalação	The physical means or contrivances to make something (especially a public service) possible ; the required equipment, infrastructure, location etc.	LST
N 5081	fact	hecho	fait	fatto	fato	Something concrete used as a basis for further interpretation	LST
N 50751	factor	factor	facteur	fattore	fator	One of the elements, circumstances, or influences which contribute to produce a result	LST
V 3759	fail	fallar	ne pas parvenir	fallire	falhar	Not to achieve a particular stated goal	LST
N 3730	failure	fallo / insuficiencia	défaut / insuffisance	fallimento / mancanza / insufficienza	falha / insuficiência	1) failure : (medicine) Inability to perform its organic function	Lspé
N 3730	failure	fallo / fracaso	échec	fallimento	falha / fracasso	2) failure : State or condition of not meeting a desirable or intended objective	LST
V 2070	fall	caer	tomber	scendere	cair	To become lower	LST
A 3009	false	falso	faux	falso	falso	Untrue, not factual, wrong	LST
N 12864	family	familia	famille	famiglia	família	1) family : (biology, taxonomy) A rank in the classification of organisms, below order and above genus ; a taxon at that rank	Lspé
N 12864	family	familia	famille	famiglia	família	2) family : A group of people who are closely related (...)	Lspé
A 1962	fast	rápido	rapide	rapido	rápido	Swift, rapid	LST
N 3224	fat	grasa	graisse	grasso	gordura	A specialized animal tissue with a high oil content, used for long-term storage of energy	Lspé
N 1850	fatigue	fatiga	fatigue	fatica	fadiga	A weariness caused by exertion; exhaustion	Lspé
A 2860	fatty	graso	gras	grasso	gordo	Like fat ; greasy	Lspé
N 10257	feature	característica	caractéristique	caratteristica	característica	(...) any marked peculiarity or characteristic (...)	LST
V 1385	feed	nutrir	nourrir	nutrire	nutrir	1) feed : To give food to eat	Lspé
V 1385	feed	alimentar	alimenter	alimentare	alimentar	2) feed : To give to a machine to be processed	LST
N 1599	feedback	realimentación / retroacción / * retroalimentación	feedback / rétroaction	feedback / retroazione / retroregolazione	feedback / retroalimentação / realimentação	(cybernetics, systems) The signal that is looped back to control a system within itself	LST

V 2097	feel	estimar	estimer	stimare	estimar	To think, believe, or have an impression concerning	LST
N 2467	female	mujer	femme	donna	mulher	1) female : A human member of the feminine sex or gender	Lspé
N 2467	female	hembra	femelle	femmina	fêmea	2) female : An animal of the sex that produces eggs	Lspé
A 4966	female	femenino	fémnin / femelle	femminile	feminino	Belonging to the sex with larger, fertilizable gametes	Lspé
A 4294	fetal	fetal	foetal	fetale	fetal	Pertaining to, or connected with, a fetus	Lspé
N 1577	fever	fiebre	fièvre	febbre	febre	A higher than normal body temperature of a person (or, generally, a mammal), usually caused by disease	Lspé
N 3963	fiber	fibra	fibre	fibra	fibra	A long tubular cell found in muscle tissue ; myocyte	Lspé
N 5809	fibroblast	fibroblasto	fibroblaste	fibroblasto	fibroblasto	A cell found in connective tissues that produces fibers, such as collagen	Lspé
N 2978	fibrosis	fibrosis	fibrose	fibrosi	fibrose	(medicine) The formation of excess fibrous connective tissue in an organ	Lspé
N 2427	field	campo	champ	campo	campo	1) field : A physical phenomenon, such as a force, potential or fluid velocity, that pervades a region	Lspé
N 2427	field	campo	champ	campo	campo	2) field : The extent of a given perception	Lspé
N 2427	field	terreno	terrain	terreno / campo	terreno	3) field : A realm of practical, direct, or natural operation, contrasting with an office, classroom, or laboratory	LST
N 28198	fig	Fig.	fig.	fig.	fig.	Abbreviation of figure (diagram or illustration)	LST
N 61628	figure	figura	figure	figura	figura	A drawing or diagram conveying information	LST
N 11274	file	fichero / archivo	fichier	file	ficheiro / arquivo	(computing) An aggregation of data on a storage device, identified by a name	LST
N 2757	filter	filtro	filtre	filtro	filtro	A device which separates a suspended, dissolved, or particulate matter from a fluid, solution, or other substance ; any device that separates one substance from another	Lspé
A 9361	final	final	final	finale	final	Last, ultimate	LST
A 1619	financial	financiero	financier	finanziario	financeiro	Related to finances	LST
V 39963	find	encontrar	trouver	trovare	encontrar	To discover by study or experiment direct to an object or end	LST
N 15261	finding	resultado	résultat	risultato	resultado	A result of research (...)	LST
A 21432	first	primero	premier	primo	primeiro	Preceding all others of a series or kind (...)	LST
N 2614	fish	pez	poisson	pesce	peixe	A cold-blooded vertebrate animal that lives in water, moving with the help of fins and breathing with gills	Lspé
V 946	fit	convenir	convenir	convenire	convir	1) fit : To be suitable for	LST
V 946	fit	ajustar	ajuster	aggiustare	ajustar	2) fit : To adjust	LST
V 2713	fix	fijar	fixer	fissare	fixar	(chemistry, biology) To convert into a stable or available form	Lspé
A 1711	fixed	fijado	fixé	fissato	fixado	Chemically stable	Lspé
N 8203	flow	flujo	flux	flusso	fluxo	The movement of a real or figurative fluid	Lspé

N 3552	fluid	fluido	fluide	fluido	fluido	(Physics) Any substance which can flow with relative ease, tends to assume the shape of its container, and obeys Bernoulli's principle ; a liquid, gas or plasma	Lspé
A 1629	fluid	fluido	fluide	fluido	fluido	Of or relating to fluid	Lspé
N 4481	fluorescence	fluorescencia	fluorescence	fluorescenza	fluorescência	(physics) The emission of light (or other electromagnetic radiation) by a material when stimulated by the absorption of radiation or of a subatomic particle	Lspé
A 2625	fluorescent	fluorescente	fluorescent	fluorescente	fluorescente	Emitting visible light as a result of the excitation of phosphors by ultraviolet photons	Lspé
N 3379	focus	enfoque	focalisation	focalizzazione	enfoque	Concentration of attention	LST
V 4844	focus	enfocar / centrarse	se focaliser / se concentrer	focalizzarsi / concentrarsi	focalizar / centrar-se	To concentrate one's attention	LST
N 2011	follicle	fóliculo	follicule	follicolo	fóliculo	(anatomy) A small cavity or sac (...)	Lspé
V 12033	follow	seguir	suivre	seguire	seguir	1) follow : To go or come after in a sequence	LST
V 12033	follow	seguir	suivre	seguire	seguir	2) follow : To carry out (orders, instructions, etc.)	LST
N 5417	follow-up	seguimiento	suivi	follow-up	seguimento	(medicine) The revisiting of a patient in ambulatory care	Lspé
A 6528	following	siguiente	suisvant	seguinte	seguinte	Next in sequence or time	LST
N 5528	food	alimento	aliment	alimento	alimento	Any substance that can be consumed by living organisms, especially by eating, in order to sustain life	Lspé
N 2741	force	fuerza	force	forza	força	Strength or energy of body or mind; active power; vigour; might; capacity of exercising an influence or producing an effect	Lspé
N 11269	form	forma	forme	forma	forma	The shape or visible structure of a thing or person (...)	LST
V 7243	form	formar	former	formare	formar	To constitute, to compose, to make up	LST
N 2343	format	formato	format	format	formato	(computing) A file type	LST
N 8577	formation	formación	formation	formazione	formação	The process during which something comes into being and gains its characteristics	LST
N 2512	foundation	fundación	fondation	fondazione	fundação	A donation or legacy appropriated to support a charitable institution, and constituting a permanent fund; endowment	LST
N 6031	fraction	fracción	fraction	frazione	fração	A part of a whole, especially a comparatively small part	LST
N 2261	fracture	fractura	fracture	frattura	fratura	The act of breaking, or something that has broken, especially that in bone or cartilage	Lspé
N 7710	fragment	fragmento	fragment	frammento	fragmento	A part broken off ; a small, detached portion (...)	Lspé
N 1354	frame	marco	cadre	quadro	fase	1) frame : (genetics : reading frame) A way of dividing nucleotide sequences into a set of consecutive triplets	Lspé
N 1354	frame	cuadro	cadre	cornice	quadro	2) frame : (Internet) An individually scrollable region of a web page	LST
N 2411	framework	marco	cadre	quadro	quadro	A basic conceptual structure	LST
A 2832	free	sin	sans	senza	sem	1) Without; not containing (what is specified); exempt; clear; liberated	LST
A 2832	free	libre	libre	libero	livre	2) Unattached or uncombined	Lspé
A 2832	free	libre	libre	libero	livre	3) With no or only freedom-preserving limitations on distribution or modification	LST
N 13916	frequency	frecuencia	fréquence	frequenza	frequência	The rate of occurrence of anything ; the relationship between incidence and time period	LST

A	2771	frequent	frecuente	fréquent	frequente	frequente	Done or occurring often	LST
A	5035	full	completo	complet	completo	completo	Total, entire	LST
A	1974	full-length	completo	complet	completo	completo	Not shortened ; complete and uncut	LST
N	30722	function	función	fonction	funzione	função	The physiological activity of an organ or body part	Lspé
V	3009	function	funcionar	fonctionner	funzionare	funcionar	To carry on a function; to be in action	LST
A	15441	functional	funcional	fonctionnel	funzionante	funcional	Of a disease, such that its symptoms cannot be referred to any appreciable lesion or change of structure; opposed to organic disease, in which the organ itself is affected	Lspé
A	9000	further	adicional	supplémentaire	extra	adicional	More, additional	LST
N	4763	fusion	fusión	fusion	fusione	usão	The result of the hybridization of two genes which originally coded for separate proteins	Lspé
A	3673	future	futuro	futur	futuro	futuro	Having to do with or occurring in the future	LST
N	2433	gain	ganancia	gain	guadagno	ganho	What is gained	LST
N	1251	gamma	gama	gamma	gamma	gama	1) gamma : The name of the third letter of the Greek alphabet ( $\Gamma$ , $\gamma$ ), preceded by beta ( $B$ , $\beta$ ) and followed by delta, ( $\Delta$ , $\delta$ )	LST
N	1251	gamma	gama	gamma	gamma	gama	2) gamma : (mathematics) The Gamma function, symbolized by $\Gamma$	LST
N	1696	gap	gap	trou	gap	lacuna	1) gap : (genetics) Unsequenced region in a sequence alignment	Lspé
N	1696	gap	gap	trou	gap	lacuna	2) gap : Something missing	LST
N	2043	gas	gas	gaz	gas	gás	(chemistry) Matter in a state intermediate between liquid and plasma (...)	Lspé
A	2928	gastric	gástrico	gastrique	gastrico	gástrico	Of or relating to the stomach	Lspé
A	1802	gastrointestinal	gastrointestinal	gastrointestinal	gastrointestinal	gastrointestinal	Of or pertaining to the stomach and intestines	Lspé
N	6867	gel	gel	gel	gel	gel	A semi-solid to almost solid colloid of a solid and a liquid (...)	Lspé
N	4207	gender	género	genre	genere	gênero	Biological sex : the sum of the biological characteristics by which male and female and other organisms are distinguished	Lspé
N	134095	gene	gen	gène	gene	gene	A unit of heredity, a segment of DNA or RNA that is transmitted from one generation to the next.	Lspé
A	6545	general	general	général	generale	geral	1) general : Not specific or particular	Lspé
A	6545	general	general	général	generale	geral	2) general : Not limited in use or application; applicable to the whole or every member of a class or category	LST
V	6277	generate	generar	générer	generare	gerar	1) generate : To produce as a result of a chemical or physical process	Lspé
V	6277	generate	generar	générer	generare	gerar	2) generate : To bring into being; give rise to	LST
N	5027	generation	generación	génération	generazione	geração	The fact of creating something, or bringing something into being ; production, creation	Lspé
A	14885	genetic	genético	génétique	genetico	genético	Relating to genetics or genes	Lspé
N	3531	genetics	genética	génétique	genetica	genética	(biology, genetics) The branch of biology that deals with the transmission and variation of inherited characteristics, in particular chromosomes and DNA.	Lspé
N	28443	genome	genoma	génomme	genoma	genoma	The complete genetic information (either DNA or, in some viruses, RNA) of an	Lspé

						organism, typically expressed in number of base pairs	
A 9802	genomic	genómico	génomique	genomico	genômico	Of or pertaining to a genome	Lspé
N 1917	genomics	genómica	génomique	genomica	* genômica	(genetics) The study of the complete genome of an organism	Lspé
N 6245	genotype	genotipo	génotype	genotipo	genótipo	(genetics) The combination of alleles, situated on corresponding chromosomes, that determines a specific trait of an individual (...)	Lspé
A 2712	geographical	geográfico	géographique	geografico	geográfico	Of or relating to geography	LST
V 2333	get	obtener	obtenir	ottenere	obter	To obtain; to acquire	LST
V 21996	give	dar	donner	dare	dar	To move, shift, provide something (...)	LST
N 3089	gland	glándula	glande	ghiandola	glândula	(zoology) An organ that synthesizes a substance, such as hormones or breast milk, and releases it, often into the bloodstream (endocrine gland) or into cavities inside the body or its outer surface (exocrine gland)	Lspé
N 1664	glass	vidrio	verre	vetro	vidro	An amorphous solid, often transparent substance made by melting sand with a mixture of soda, potash and lime	Lspé
A 5426	global	global	global	globale	global	Which has to be considered in its entirety	LST
N 6084	glucose	glucosa	glucose	glucoso / glicosio / glucosio / glicoso	glucose / glicose	(biochemistry) A simple monosaccharide (sugar) (...)	Lspé
N 1803	glutamate	glutamato	glutamate	glutamato / glutammato	glutamato	(chemistry) Any salt or ester of glutamic acid	Lspé
N 1704	glycoprotein	glicoproteína	glycoprotéine	glicoproteina	glicoproteína	(biochemistry) A protein with covalently bonded carbohydrates	Lspé
V 3707	go	andar	aller	andare	andar	To move (a particular distance, or in a particular fashion)	LST
N 3024	goal	objetivo	objectif	obiettivo	objetivo	A result that one is attempting to achieve	LST
A 15649	good	bueno	bon	buono	bom	Being satisfying	LST
N 2205	government	gobierno / Estado	gouvernement / État	governo / Stato	governo / Estado	The state and its administration viewed as the ruling political power	LST
N 3063	grade	grado	grade	grado	grau	(medicine) The degree of malignity of a tumor expressed on a scale	Lspé
N 1640	gradient	gradiente	gradient	gradiente	gradiente	(sciences) The rate at which a physical quantity increases or decreases relative to change in a given variable (...)	LST
N 3790	grant	beca	bourse	borsa	bolsa	The thing or property granted ; a gift ; a boon	LST
N 1910	graph	gráfico	graphique	grafico	gráfico	A diagram displaying data; in particular one showing the relationship between two or more quantities, measurements or indicative numbers that may or may not have a specific mathematical formula relating them to each other	LST
A 13293	great	grande	grand	grande	grande	Very big, large	LST
A 3618	green	verde	vert	verde	verde	Having green as its color	LST

N 58821	group	grupo	groupe	gruppo	grupo	A number of things or persons being in some relation to one another	LST
V 2750	grow	crecer	croître / grandir	crescere	crescer	1) grow : To become bigger	LST
V 2750	grow	cultivar	cultiver	coltivare	cultivar	2) grow : To cause or allow something to become bigger, especially to cultivate (...)	Lspé
N 23669	growth	crecimiento	croissance	crescita	crescimento	An increase in size, number, value or strength	LST
N 5883	guideline	directriz	directive	direttiva	diretriz	A non-specific rule or principle that provides direction to action or behavior	LST
N 2133	half	mitad	moitié	metà	metade	One of two usually roughly equal parts into which anything may be divided, or considered as divided (...)	LST
N 2540	hand	mano	main	mano	mão	The part of the fore limb below the forearm or wrist in a human, and the corresponding part in many other animals	Lspé
N 2274	haplotype	* haplotipo	haplotype	* aplotipo	haplótipo	(genetics) A group of alleles that are transmitted together	Lspé
V 73319	have	tener	avoir	avere	ter	To possess, own, hold	LST
N 1622	head	cabeza	tête	testa / capo	cabeça	1) head : The part of the body of an animal or human which contains the brain, mouth and main sense organs	Lspé
N 1622	head	cabeza	tête	testa / capo	cabeça	2) head : The rounded part of a bone fitting into a depression in another bone (...)	Lspé
N 55877	health	salud	santé	sanità	saúde	The state of being free from physical or psychological disease, illness, or malfunction	Lspé
A 6127	healthy	sano	sain	sano	sano	Enjoying health and vigor of body, mind or spirit	Lspé
N 11951	heart	corazón	cœur	cuore	coração	(anatomy) A muscular organ that pumps blood through the body (...)	Lspé
N 3217	heat	calor	chaleur	calore	calor	Thermal energy	Lspé
N 1584	height	altura	hauteur	altezza	altura	The vertical distance from the ground to the highest part of a standing person or animal	Lspé
N 1908	helix	hélice	hélice	elica	hélice	(mathematics) A curve on the surface of a cylinder or cone such that its angle to a plane perpendicular to the axis is constant (...)	LST
N 1682	help	ayuda	aide	aiuto	ajuda	Action given to provide assistance ; aid	LST
V 2746	help	ayudar	aider	aiutare	ajudar	To contribute in some way to	LST
N 1735	heparin	heparina	héparine	eparina	heparina	(medicine, biochemistry) A glycosaminoglycan, originally isolated from liver cells, now made synthetically for medical use, used as an anticoagulant	Lspé
A 2775	hepatic	hepático	hépatique	epatico	hepático	Of or relating to the liver	Lspé
N 2907	hepatitis	hepatitis	hépatite	epatite	hepatite	Inflammation of the liver, sometimes caused by a viral infection	Lspé
N 2282	heterogeneity	heterogeneidad	hétérogénéité	eterogeneità	heterogeneidade	The quality of a substance which is not uniform	LST
A 56715	high	alto	haut	alto	alto	Elevated	LST
V 1781	highlight	enfaticar / destacar / resaltar / realzar	souligner	enfaticare / sottolineare / evidenziare	enfaticar / destacar / realçar / ressaltar	To make prominent; emphasize	LST
N 1682	hip	cadera	hanche	anca	quadril	(anatomy) The outward-projecting parts of the pelvis and top of the femur and the	Lspé

A 2880	histological	histológico	histologique	istologico	histológico	overlying tissue Pertaining to histology	Lspé
N 2035	histone	histona	histone	istone	histona	(biochemistry) Any of various simple water soluble proteins that are rich in the basic amino acids lysine and arginine and are complexed with DNA in the nucleosomes of eukaryotic chromatin	Lspé
N 3348	history	historia	histoire	storia	história	1) history : (medicine) A list of past and continuing medical conditions of an individual or family	Lspé
N 3348	history	historia	histoire	storia	história	2) history : The branch of knowledge that studies the past	LST
N 1660	hit	resultado	résultat	risultato	resultado	(computing, Internet) The result of a search of a computer system or of a search engine	LST
N 9886	HIV	VIH	VIH	VIH	HIV / VIH	(virology) Initialism of human immunodeficiency virus	Lspé
V 2272	hold	mantener	maintenir	mantenere	manter	To maintain or keep to a position or state	LST
N 1981	home	residencia / hogar / asilo	résidence / maison	residenza / casa	lar / casa / asilo	1) home : A place of refuge, rest or care; an asylum	Lspé
N 1981	home	domicilio / casa	domicile / maison	domicilio / casa	domicílio / casa / lar	2) home : One's own dwelling place ; the house or structure in which one lives (...)	LST
A 2922	homologous	homólogo	homologue	omologo	homólogo	Corresponding to a similar structure	Lspé
N 4104	homologue	homólogo	homologue	omologo	homólogo	(genetics) One of a group of similar DNA sequences that share a common ancestry	Lspé
N 1653	homology	homología	homologie	omologia	homologia	1) homology : (genetics) The presence of the same series of bases in related genes	Lspé
N 1653	homology	homología	homologie	omologia	homologia	2) homology : (chemistry) The relationship between the elements in the same group of the periodic table, or between organic compounds in a homologous series	Lspé
N 6230	hormone	hormona	hormone	ormone	hormônio	(physiology) Any substance produced by one tissue and converged by the bloodstream to another place to effect physiological activity	Lspé
N 12999	hospital	hospital	hôpital	ospedale	hospital	A building designed to diagnose and treat the sick, injured or dying (...)	Lspé
N 2004	hospitalization	hospitalización	hospitalisation	ospedalizzazione	hospitalização	The hospitalizing of a patient, the condition of being hospitalized, or the period a patient stays in hospital	Lspé
N 5559	host	huésped / hospedante	hôte	ospite	hospedeiro	(biology) A cell or organism which harbors another organism or biological entity (...)	Lspé
N 17906	hour	hora	heure	ora	hora	A time period of sixty minutes (...)	LST
N 2728	household	hogar	ménage	nucleo familiare	agregado familiar	Collectively, all the persons who live in a given house (...)	LST
N 5069	human	humano	humain	umano	humano	A human being, whether man, woman or child	Lspé
A 54183	human	humano	humain	umano	humano	Of or belonging to the species Homo sapiens	Lspé
A 1664	hybrid	híbrido	hybride	ibrido	híbrido	Consisting of diverse components	Lspé
N 7776	hybridization	hibridación	hybridation	ibridizzazione	hibridação	(chemistry) The linear combination of atomic orbitals in a molecule to form hybrid orbitals	Lspé



V 1801	hybridize	hibridar	hybrider	ibridare	hibridizar	(biochemistry) to combine complementary subunits of multiple biological macromolecules	Lspé
N 4786	hypertension	hipertensión	hypertension	ipertensione	hipertensão	(pathology) The disease or disorder of abnormally high blood pressure.	Lspé
N 7603	hypothesis	hipótesis	hypothèse	ipotesi	hipótese	(sciences) Used loosely, a tentative conjecture explaining an observation, phenomenon or scientific problem that can be tested by further observation (...)	LST
V 1913	hypothesize	hipotetizar	* hypothétiser / faire l'hypothèse	ipotizzare	* hipotetizar / fazer a hipótese	To believe or assert on uncertain grounds	LST
N 2167	hypoxia	hipoxia	hypoxie	ipossia	hipóxia	(pathology) A condition in which tissues (especially the blood) are deprived of an adequate supply of oxygen ; anoxia	Lspé
N 1983	idea	idea	idée	idea	ideia	(...) Any result of mental activity; a thought, a notion; a way of thinking	LST
A 4379	identical	idéntico	identique	identico	idêntico	Bearing full likeness by having precisely the same characteristics	LST
N 9821	identification	identificación	identification	identificazione	identificação	The act of identifying, or proving to be the same	LST
V 31051	identify	identificar	identifier	identificare	identificar	To establish the identity of someone or something	LST
N 3657	identity	identidad	identité	identità	identidade	Sameness, identicalness ; the quality or fact of (several specified things) being the same	LST
A 1943	ill	enfermo	malade	malato	enfermo / doente	Unwell in terms of health or physical condition ; sick	Lspé
N 1931	illness	enfermedad	maladie	malattia	enfermidade	1) illness : An instance of a disease or poor health	Lspé
N 1931	illness	enfermedad	maladie	malattia	enfermidade	2) illness : A state of bad health or disease	Lspé
V 3002	illustrate	ilustrar	illustrer	illustrare	ilustrar	To clarify something by giving, or serving as, an example or a comparison	LST
N 7324	image	imagen	image	immagine	imagem	An optical or other representation of a real object ; a graphic ; a picture	LST
N 3782	imaging	imagen	imagerie	imaging	imagiologia	The technique or practice of creating images of otherwise invisible aspects of an object, especially of body parts	Lspé
A 4016	immune	inmunitario	immunitaire	immunitario	imunitário	1) immune : Of or pertaining to the immune system	Lspé
A 4016	immune	immune	immun	immune	imune	2) immune : Protected by inoculation or due to innate resistance to pathogens	Lspé
N 2242	immunity	inmunidad	immunité	immunità	imunidade	(medicine) Fully protective resistance against infection	Lspé
N 1647	immunization	inmunización	immunisation	immunizzazione	imunização	The process by which an individual is exposed to a material that is designed to prime his or her immune system against that material	Lspé
N 2207	immunodeficiency	inmunodeficiencia	immunodéficience	immunodeficienza	imunodeficiência	(medicine) A depletion in the body's natural immune system, or in some component of it	Lspé
N 1565	immunoglobulin	inmunoglobulina	immunoglobuline	immunoglobulina	imunoglobulina	(biochemistry, immunology) Any of the glycoproteins in blood serum that respond to invasion by foreign antigens and that protect the host by removing pathogens; an antibody	Lspé
A 1689	immunohistochemical	* inmunohistoq	* immunohistoch	immunoistochimico	* imunohistoqu	Of, pertaining to, or by means of immunohistochemistry	Lspé

N 7759	impact	uímico	imique	impatto	ímico	A significant or strong influence ; an effect	LST
V 1772	impair	perjudicar	affaiblir	pregiudicare	prejudicar	To weaken; to affect negatively; to have a diminishing effect on	Lspé
N 2302	impairment	insuficiencia	insuffisance	insufficienza	insuficiência	The result of being impaired; a deterioration or weakening; a disability or handicap; an inefficient part or factor	Lspé
V 3286	implement	implementar	implémenter / mettre en œuvre	implementare	implementar	To bring about; to put into practice	LST
N 2590	implementation	implementación	implémentation / mise en œuvre	implementazione	implementaçã o	The process of moving an idea from concept to reality (...)	LST
V 2485	implicate	implicar	impliquer	implicare	implicar	To imply, to have as a necessary consequence or accompaniment	LST
N 4708	implication	implicación	implication	implicazione	implicação	Logical consequence	LST
V 2458	imply	implicar	impliquer	implicare	implicar	To have as a necessary consequence	LST
N 5501	importance	importancia	importance	importanza	importância	The quality or condition of being important or worthy of note	LST
A 20989	important	importante	important	importante	importante	Having relevant and crucial value	LST
V 9754	improve	mejorar	améliorer	migliorare	melhorar	To make (something) better; to increase the value or productivity (of something)	LST
A 2339	improved	mejorado	amélioré	migliorato	melhorado	That has been made better	LST
N 5296	improvement	mejoramiento	amélioration	miglioramento	melhorament o	The state of being improved ; betterment	LST
A 11160	in vitro	In vitro	in vitro	in vitro	in vitro	In an artificial environment outside the living organism	Lspé
A 7009	in vivo	in vivo	in vivo	in vivo	in vivo	Within a living organism	Lspé
N 1865	inactivation	inactivación	inactivation	inattivazione	inativação	The process of rendering something inactive	Lspé
N 7685	incidence	incidencia	incidence	incidenza	incidência	(epidemiology) A measure of the risk that a person develops a new condition within a specified period of time (...)	Lspé
V 25275	include	incluir	inclure	includere	incluir	To contain, as parts of a whole; to comprehend	LST
N 1674	inclusion	inclusión	inclusion	inclusione	inclusão	1) inclusion : An addition or annex to a group, set or total	LST
N 1674	inclusion	inclusión	inclusion	inclusione	inclusão	2) inclusion : (biology) A nuclear or cytoplasmic aggregate of stainable substances	Lspé
N 2457	income	ingreso	revenu	reddito	renda	Money one earns by working or by capitalizing on the work of others	LST
V 2816	incorporate	incorporar	incorporer	incorporare	incorporar	To include as a part	LST
N 18132	increase	acrecentamie nto	accroissement	accrescimento	acrescimento	For a quantity, the act or process of becoming larger	LST
V 27691	increase	acrecentar	s'accroître	accrescersi	acrescentar-se	To become larger	LST
A 9775	increased	mayor	accru	maggiore	maior	Made or become greater	LST
V 6988	incubate	incubar	incuber	incubare	incubar	To brood, raise, or maintain eggs, organisms, or living tissue through the provision of	Lspé

						ideal environmental conditions	
N 4125	incubation	incubación	incubation	incubazione	incubação	A period of little reaction which is followed by more rapid reaction	Lspé
A 10255	independent	independiente	indépendant	indipendente	independente	Not dependent, free	LST
N 7544	index	índice	indice	indice	índice	(sciences) A number representing a property or ratio, a coefficient	LST
V 21597	indicate	indicar	indiquer	indicare	indicar	To point out; to discover; to direct to a knowledge of; to show; to make known	LST
N 1835	indication	indicación	indication	indicazione	indicação	That which serves to indicate or point out; mark; token; sign; symptom; evidence	LST
N 3500	indicator	indicador	indicateur	indicatore	indicador	(chemistry) Any of many substances, such as litmus, used to indicate the concentration of a substance, or the degree of a reaction	Lspé
A 1732	indirect	indirecto	indirect	indiretto	indireto	Not direct	LST
N 11914	individual	individuo	individu	individuo	indivíduo	An element belonging to a population	LST
A 11348	individual	individual	individuel	individuale	individual	Relating to a single person or thing	LST
V 30367	induce	inducir	induire	indurre	induzir	To cause, bring about, lead to	LST
A 2041	inducible	* inducible	inductible	inducibile	induzível	Able to be induced or caused	LST
N 7451	induction	inducción	induction	induzione	indução	(biology) In developmental biology, the development of a feature from part of a formerly homogenous field of cells in response to a morphogen whose source determines the feature's position and extent	Lspé
N 3583	infant	lactante	nourrisson	lattante	lactente	A very young human being, from birth to somewhere between six months and two years of age, needing almost constant care and/or attention	Lspé
N 1406	infarction	infarto	infarctus	infarto	infarto / enfarte	1) infarction (pathology) An infarct	Lspé
N 1406	infarction	infarto	infarcissement	infarto	enfartamento	2) infarction : (pathology) The process which causes an infarct	Lspé
V 6385	infect	infectar	infecter	infettare	infectar	To bring into contact with a substance that causes illness (a pathogen)	Lspé
A 3696	infected	infectado	infecté	infettato	infectado	Having an infection	Lspé
N 21518	infection	infección	infection	infezione	infecção	(pathology) The act or process of infecting	Lspé
A 1269	infectious	infeccioso	infectieux	infettivo	infeccioso	1) infectious : (medicine) Transmitted from one person to another (...)	Lspé
A 1269	infectious	infeccioso / contagioso	contagieux	contagioso	contagioso	2) infectious : (medicine) Able to infect others	Lspé
V 1610	infer	inferir	inférer	inferire	inferir	To lead to (something) as a consequence; to imply	LST
N 5607	inflammation	inflamación	inflammation	infiammazione	inflamação	(medicine) A condition of any part of the body, consisting in congestion of the blood vessels, with obstruction of the blood current, and growth of morbid tissue (...)	Lspé
A 9258	inflammation and inflammator y	inflamatorio	inflammatoire	infiammatorio	inflamatório	Causing or caused by inflammation	Lspé
N 5207	influence	influencia	influence	influenza	influência	The power to affect, control or manipulate something (...)	LST
V 6177	influence	influir	influencer /	influenzare	influenciar /	To (...) exert an influence upon ; to modify, bias, or sway (...)	LST

V 1888	inform	informar	influer	informare	influir	To communicate knowledge to	LST
N 24620	information	información	information	informazione	informação	As contrasted with data, knowledge which is gathered as a result of processing data	LST
N 2462	infusion	perfusión	perfusion	perfusione	perfusão	(medicine) The administration of liquid substances directly into a vein for medical purposes	Lspé
V 9545	inhibit	inhibir	inhiber	inibire	inibir	to hinder; to restrain	Lspé
N 9054	inhibition	inhibición	inhibition	inibizione	inibição	(chemistry) The process of stopping or retarding a chemical reaction	Lspé
A 3279	inhibitive	inhibitorio	inhibiteur	inibitorio	inibitivo / inibitório	That inhibits	Lspé
N 10883	inhibitor	inhibidor	inhibiteur	inibitore	inibidor	(chemistry) Any substance capable of stopping or slowing a specific chemical reaction	Lspé
A 8054	initial	inicial	initial	iniziale	inicial	Chronologically first, early ; of or pertaining to the beginning	LST
V 2892	initiate	iniciar	initier / amorcer	iniziare	iniciar	To begin; to start	LST
N 1560	initiation	iniciación	amorçage	iniziazione	iniciação	1) initiation : (chemistry) The first step of transcription or of transduction	Lspé
N 1560	initiation	iniciación	initiation	iniziazione	iniciação	2) initiation : The act of initiating, or the process of being initiated or introduced	LST
V 2640	inject	inyectar	injecter	iniettare	injetar	To administer an injection (...), especially of medicine or drugs	Lspé
N 5258	injection	inyección	injection	iniezione	injeção	(medicine) Something injected subcutaneously, intravenously, or intramuscularly by use of a syringe and a needle	Lspé
N 7710	injury	herida	blessure	ferita	ferida	Damage to the body of a human or animal	Lspé
N 3922	input	entrada	entrée	input	entrada	Something fed into a process with the intention of it shaping or affecting the outputs of that process	LST
N 1758	insect	insecto	insecte	insetto	inseto	An arthropod in the class Insecta, characterized by six legs, up to four wings, and a chitinous exoskeleton	Lspé
V 1584	insert	insertar	insérer	inserire	inserir	To put in between or into	Lspé
N 3386	insertion	inserción	insertion	inserzione	inserção	(genetics) The addition of a nucleotide to a chromosome (...)	Lspé
N 2382	insight	comprensión	compréhension	comprensione	compreensão	(...) An extended understanding of a subject (...)	LST
N 1979	institute	instituto	institut	istituto	instituto	1) institute : An institution of learning ; a college, especially for technical subjects	LST
N 1979	institute	instituto	institut	istituto	instituto	2) institute : An organization founded to promote a cause	LST
N 1855	institution	institución	institution	istituzione	instituição	An established organization (...)	LST
N 1953	instruction	instrucción	instruction	istruzione	instrução	A set of directions provided by a manufacturer for the users of a product or service	LST
N 4462	instrument	instrumento	instrument	strumento	instrumento	A tool, implement used for manipulation or measurement	LST
N 7349	insulin	insulina	insuline	insulina	insulina	A polypeptide hormone that regulates carbohydrate metabolism	Lspé
A 2083	intact	intacto	intact	intatto	intato	Left complete or entire ; not damaged	Lspé
N 3831	intake	ingestión / toma	ingestion / prise	ingestione / assunzione /	ingestão / tomada	An act or instance of taking in (...)	Lspé

V 2066	integrate	integrar	intégrer	presa integrare	integrar	To include as a constituent part or functionality	LST
N 2403	integration	integración	intégration	integrare integrazione	integrar integração	The combination with compatible elements in order to incorporate them	LST
N 2281	integrin	* integrina	* intégrine	integrina	integrina	(biochemistry) Any of many heterodimeric transmembrane proteins that function as receptors in communication between cells	Lspé
N 7388	intensity	intensidad	intensité	intensità	intensidade	The quality of being intense	LST
A 3594	intensive	intensivo	intensif	intensivo	intensivo *	(medicine) Related to the need to manage life-threatening conditions by means of sophisticated life support and monitoring	Lspé
V 5170	interact	interactuar / interaccionar	interagir	interagire	interaccionar / interagir	To act upon each other (...)	LST
N 20011	interaction	interacción	interaction	interazione	interação	The situation or occurrence in which two or more objects or events act upon one another to produce a new effect ; the effect resulting from such a situation or occurrence	LST
N 5349	interest	interés	intérêt	interesse	interesse	1) interest : Attention that is given to or received from someone or something	LST
N 5349	interest	interés	intérêt	interesse	interesse	2) interest : A business link or involvement	LST
A 2611	interesting	interesante	intéressant	interessante	interessante	Arousing or holding the attention	LST
N 1293	interface	interfaz	interface	interfaccia	interface	1) interface : (chemistry, physics) A thin layer or boundary between different substances or two phases of a single substance	Lspé
N 1293	interface	interfaz	interface	interfaccia	interface	2) interface : (computing) The point of interconnection between systems or subsystems	LST
N 2475	interferon	interferón	interféron	interferone	interferon	(biochemistry) Any of a group of glycoproteins, produced by the immune system, that prevent viral replication in infected cells	Lspé
N 2184	interleukin	interleucina	interleukine	interleuchina	interleucina	(biochemistry) Any of a group of cytokine proteins important in the regulation of lymphocyte function	Lspé
A 1972	intermediate	intermedio	intermédiaire	intermedio	intermediário	Between two extremes, or in the middle of a range	LST
N 1793	intern	interno	interne	interno	interno	A medical student or recent graduate working in a hospital as a final part of medical training	Lspé
A 4358	internal	interno	interne	interno	interno	Within the body	Lspé
A 5627	international	internacional	international	internazionale	internacional	Concerning more than one nation	LST
N 1556	internet	internet	internet	internet	internet	The specific internet consisting of a global network of computers that communicate using Internet Protocol (IP) and that use Border Gateway Protocol (BGP) to identify the best paths to route those communications	LST
V 2687	interpret	interpretar	interpréter	interpretare	interpretar	To explain or tell the meaning of; to expound (...)	LST
N 3716	interpretation	interpretación	interprétation	interpretazione	interpretação	An act of interpreting or explaining what is obscure; a translation; a version; a construction	LST
N 7241	interval	intervalo	intervalle	intervallo	intervalo	A period of time	LST
N 10198	intervention	intervención	intervention	intervento	intervenção	(medicine) Operation	Lspé

N 3902	interview	entrevistó / entrevista	entretien	intervista	entrevista	A conversation in person (or, by extension, over the telephone, Internet etc.) (...)	LST
A 2411	intestinal	intestinal	intestinal	intestinal	intestinal	Relative to the intestines	Lspé
A 4820	intracellular	intracelular	intracellulaire	intracellulare	intracelular	Inside or within a cell	Lspé
A 1789	intravenous	intravenoso / endovenoso	intraveineux / * endoveineux	intravenoso / endovenoso	intravenoso / endovenoso	Inside the veins	Lspé
V 4040	introduce	introducir	introduire	introdurre	introduzir	To add to a system, a mixture, or a container	Lspé
N 905	introduction	introducción	introduction	introduzione	introdução	1) introduction : The act or process of introducing	Lspé
N 905	introduction	introducción	introduction	introduzione	introdução	2) introduction : A written or oral explanation of what constitutes the basis of an issue	LST
N 3588	intron	* intrón	intron	introne	* íntron	(genetics) A portion of a split gene that is included in pre-RNA transcripts but is removed during RNA processing and rapidly degraded	Lspé
N 2192	invasion	invasión	invasion	invasione	invasão	(medicine) The spread of cancer cells, bacterias and such to the organism	Lspé
A 3375	invasive	invasivo	invasif	invasivo	invasivo	That invades healthy tissue ; (of a procedure) in which part of the body is entered	Lspé
V 2790	invest	invertir	investir	investire	investir	To spend money, time, or energy into something, especially for some benefit or purpose	LST
V 9036	investigate	investigar	investiguer / étudier	investigare	investigar	To inquire into or study in order to ascertain facts or information	LST
N 5334	investigation	investigación	investigation	investigazione	investigação	The act of investigating ; the process of inquiring into or following up ; research ; study ; inquiry, especially patient or thorough inquiry or examination (...)	LST
N 3797	investigator	investigador	investigateur	investigatore	investigador	One who investigates	LST
V 20600	involve	implicar	impliquer	implicare	implicar	To imply, to have as a necessary consequence or accompaniment	LST
N 3707	involvement	participación / implicación	participation / implication	partecipazione / coinvolgimento	participação / implicação	The act of involving, or the state of being involved	LST
N 2456	ion	ion / ión	ion	ione	íon / ionte / iônio	An atom or group of atoms bearing an electrical charge such as the sodium and chlorine atoms in a salt solution	Lspé
N 2493	iron	hierro	fer	ferro	ferro	(physics, chemistry, metallurgy) A metallic chemical element having atomic number 26 and symbol Fe	Lspé
N 1933	ischemia	isquemia	ischémie	ischemia	isquemia	(pathology, cardiology) Local disturbance in blood circulation due to mechanical obstruction of the blood supply (vasoconstriction, thrombosis or embolism)	Lspé
A 1691	ischemic	isquémico	ischémique	ischemico	isquêmico	(pathology) Characteristic of, or accompanied by ischemia	Lspé
N 851	island	isla	île	isola	ilha	1) island : A contiguous area of land, smaller than a continent, totally surrounded by water	LST
N 851	island	isla	île	isola	ilha	2) island : An entity surrounded by other entities that are very different from itself	Lspé
N 3591	isoform	* isoforma	* isoforme	* isoforma	* isoforma	(biochemistry) Any of several different forms of the same protein (...)	Lspé

V 9700	isolate	aislar	isoler	isolare	isolar	(chemistry) To separate a substance in pure form from a mixture	Lspé
A 2056	isolated	aislado	isolé	isolato	isolado	(medicine) Which has been extracted from the organism	Lspé
N 1943	isolation	aislamiento	isolement	isolamento	isolamento	1) Isolation : (chemistry) The obtaining of an element from one of its compounds, or of a compound from a mixture	Lspé
N 1943	isolation	aislamiento	isolement	isolamento	isolamento	2) Isolation ; The state of being away from other people	LST
N 6911	issue	cueestión	question	questione	questão	A point in debate or controversy on which the parties take affirmative and negative positions (...)	LST
N 8370	item	ítem	item	item	item	(psychometrics) A question on a test, which may include its answers	LST
N 2373	joint	articulación	articulation	articolazione	articulação	(anatomy) Any part of the body where two bones join, in most cases allowing that part of the body to be bent or straightened	Lspé
N 9588	journal	periódico	périodique	periodico	periódico	A newspaper or magazine dealing with a particular subject	LST
N 1931	junction	unión	jonction	giunzione	junção	The boundary between two physically different materials	Lspé
V 3060	keep	mantener	maintenir	mantenere	manter	To maintain the condition of	LST
A 4628	key	clave	clé	chiave	chave	Important	LST
N 4804	kg	kg	kg	kg	kg	(metrology) Symbol for the kilogram (...)	LST
N 2918	kidney	riñón	rein	rene	rim	An organ in the body that filters the blood, producing urine	Lspé
N 2585	kilobyte	kilobyte	* kilooctet	kilobit	kilobit / quibit	(computing, chiefly RAM) A unit of storage capacity, equal to 1024 (210) bytes	LST
N 12937	kinase	cinasa	kinase	cinasi / chinasi	cinase / quinase	(biochemistry) Any of a group of enzymes that transfers phosphate groups from high-energy donor molecules, such as ATP, to specific target molecules (substrates), in a process termed phosphorylation	Lspé
N 1794	kind	género	genre	genere	gênero	A type, race or category; a group of entities that have common characteristics such that they may be grouped together	LST
N 4059	kit	kit	kit	kit	kit	Any collection of items needed for a specific purpose (...)	LST
N 2727	km	km	km	km	km	(metrology) Symbol for the kilometer (...)	LST
N 2522	knee	rodilla	genou	ginocchio	joelho	In humans, the joint or the region of the joint in the middle part of the leg between the thigh and the shank	Lspé
V 17593	know	conocer / saber	connaitre / savoir	conoscere / sapere	conhecer / saber	To be aware of; to be cognizant of	LST
N 3482	knowledge	saber	savoir	sapere	saber	1) knowledge : The total of what is known ; all information and products of learning	LST
N 3482	knowledge	conocimiento	connaissance	conoscenza	conhecimento	2) knowledge : Awareness of a particular fact or situation ; a state of having been informed or made aware of something	LST
N 1913	lab	Lab.	labo	* lab	lab.	(Colloquial) A laboratory	LST
V 6961	label	marcar	marquer	marcare	marcar	(biochemistry) To replace specific atoms by their isotope in order to track the presence or movement of this isotope through a reaction, metabolic pathway or cell	Lspé
N 785	labor	trabajo	travail	lavoro	trabalho	1) labor : Effort expended on a particular task; toil, work	LST

N 785	labor	parto	travail / accouchement	parto	parto	parto	2) labor : The act of a mother giving birth	Lspé
N 8913	laboratory	laboratorio	laboratoire	laboratorio	laboratório		A room, building or institution equipped for scientific research, experimentation or analysis	LST
N 6108	lack	falta	manque	mancanza	falta		A deficiency or need (of something desirable or necessary) ; an absence, want	LST
V 3735	lack	faltar	manquer	mancare	faltar		To be short (of or for something)	LST
N 2876	lane	carril / calle	piste	corsia	pista		(genetics) Each of the columns representing a molecule and made visible after an electrophoresis in gel of a mixture	Lspé
N 1287	language	lenguaje	langage	linguaggio	linguagem		1) language : (computing) A computer language; a machine language	LST
N 1287	language	lengua	langue	lingua	língua		2) language : A body of words, and set of methods of combining them (called a grammar), understood by a community and used as a form of communication	LST
A 23516	large	grande	grand	grande	grande		Of considerable or relatively great size or extent	LST
N 1847	laser	láser	laser	laser	laser		A device that produces a monochromatic, coherent beam of light	Lspé
A 4416	last	último	dernier	ultimo	último		Most recent, latest, last so far	LST
A 3443	late	tardío	tardif	tardo	tardio		Near the end of a period of time	LST
A 1841	lateral	lateral	latéral	laterale	lateral		To the side ; of or pertaining to the side	Lspé
N 883	law	ley	loi	legge	lei		1) law : (sciences) A well-established, observed physical characteristic or behavior of nature	LST
N 883	law	ley	loi	legge	lei		2) law : The body of rules and standards issued by a government, or to be applied by courts and similar authorities	LST
N 3294	layer	capa	couche	strato	camada		A single thickness of some material covering a surface	Lspé
V 13314	lead	ocasionar	occasionner / entraîner	occasionare	ocasionar		To produce	LST
N 1947	leaf	hoja	feuille	foglia	folha		The usually green and flat organ that represents the most prominent feature of most vegetative plants	Lspé
V 2519	learn	aprender	apprendre	imparare	aprender		To acquire, or attempt to acquire knowledge or an ability to do something	LST
V 4139	leave	dejar	laisser	lasciare	deixar		To let be or do without interference	LST
A 3476	left	izquierdo	gauche	sinistro	esquerdo		The west side of the body when one is facing north	Lspé
N 5258	length	largo	longueur	lunghezza	comprimento		1) length : The distance measured along the longest dimension of an object	LST
N 5258	length	duración	durée	durata	duração		2) length : The duration	LST
N 5905	lesion	lesión	lésion	lesione	lesão		(medicine) An infected or otherwise injured or diseased organ or part, especially such part of skin	Lspé
N 2795	leukemia	leucemia	leucémie	leucemia	leucemia		Any specific form or type of cancer of the blood-forming tissues	Lspé
N 57599	level	nivel	niveau	livello	nível		Degree or amount	LST
N 3350	library	biblioteca	bibliothèque	biblioteca	biblioteca		1) library : A collection of DNA material from a single organism or relative to a single disease	Lspé



N 3350	library	biblioteca	bibliothèque	biblioteca	biblioteca	2) library : An institution which holds books and/or other forms of stored information for use by the public or qualified people	LST
N 14420	life	vida	vie	vita	vida	(biology) A status possessed by any of a number of entities, including animals, plants, fungi, bacteria, and sometimes viruses, which have the properties of replication and metabolism	Lspé
N 6096	ligand	ligando	ligand	legante	ligando / ligante	(chemistry) An ion, molecule, or functional group that binds to another chemical entity to form a larger complex	Lspé
A 1419	light	ligero	léger	leggero	ligeiro / leve	1) light : Easily interrupted by stimulation	Lspé
N 1569	light	luz	lumière	luce	luz	1) light : Spiritual or mental illumination; enlightenment, useful information	LST
N 1569	light	luz	lumière	luce	luz	2) light : A source of illumination	Lspé
A 1419	light	ligero	léger	leggero	ligeiro / leve	2) light : Gentle; having little force or momentum	LST
N 3645	likelihood	verosimilitud	vraisemblance	verosimiglianza	verossimilhança	(statistics) The probability that some fixed outcome was generated by a random distribution with a specific parameter	LST
A 11489	likely	probable	probable	probabile	provável	Probable ; having a greater than even chance of occurring	LST
N 2529	limit	límite	limite	limite	limite	A restriction ; a bound beyond which one may not go	LST
V 5236	limit	limitar	limiter	limitare	limitar	To restrict; not to allow to go beyond a certain bound	LST
N 4759	limitation	limitación	limitation	limitazione	limitação	A restriction ; a boundary, real or metaphorical, caused by some thing or some circumstance	LST
A 3960	limited	limitado	limité	limitato	limitado	Of numbers, amounts, data : poor, small, felt as insufficient	LST
N 19177	line	línea	lignée	linea	linhagem	(genetics) Population of cells derived from a single cell and containing the same genetic makeup	Lspé
N 3009	lineage	linaje	lignée	lignaggio	linhagem	Descent in a line from a common progenitor ; progeny ; race ; descending line of offspring or ascending line of parentage	Lspé
A 2652	linear	lineal	linéaire	lineare	linear	1) linear : (mathematics) Of or relating to a class or poly nominal of the form $y = ax + b$	LST
A 2652	linear	lineal	linéaire	lineare	linear	2) linear : Having the form of a line ; straight	LST
N 1489	link	enlace / vínculo	lien	legame	enlace / vínculo	1) link : A connection between places, people, events, things, or ideas	LST
N 1489	link	enlace / vínculo	lien	link	link / vínculo	2) link : Abbreviation of hyperlink	LST
V 6000	link	ligar / relacionar	lier / relier	legare / collegare	ligar / relacionar	To connect two or more things	LST
N 5985	linkage	ligamiento	liaison	linkage	ligamento	(genetics) The property of genes being inherited together	Lspé
N 7222	lipid	lípidio	lipide	lipide	lipídio / lipide	(organic chemistry) Any of a group of organic compounds including the fats, oils, waxes, sterols and triglycerides (...)	Lspé
N 2781	lipoprotein	lipoproteína	lipoprotéine	lipoproteina	lipoproteína	Any of a large group of complexes of protein and lipid with many biochemical functions	Lspé

A 1756	liquid	líquido	liquide	liquido	líquido	Flowing freely like water, fluid	Lspé
N 5883	list	lista	liste	lista	lista	(computing, programming) A codified representation of a list, used to store data (...)	LST
V 2607	list	listar	lister	elencare	listar	To create or recite a list	LST
N 6140	literature	literatura	littérature	letteratura	literatura	All the papers, treatises etc. published in academic journals on a particular subject	LST
V 1563	live	vivir	vivre	vivere	viver	1) live : To be alive; to have life	Lspé
V 1563	live	vivir / habitar	vivre / habiter	vivere / abitare	viver / habitar	2) live : To have permanent residence somewhere, to inhabit, to reside	LST
N 10792	liver	hígado	foie	fegato	fígado	(anatomy) A large organ in the body that stores and metabolizes nutrients, destroys toxins and produces bile (...)	Lspé
N 2070	load	carga	charge	carica	carga	A unit of measure for various quantities	Lspé
A 4202	local	local	local	locale	local	1) local : From or in a nearby location	LST
A 4202	local	local	local	locale	local	2) local : (medicine) Of or pertaining to a restricted part of an organism	Lspé
N 2732	localization	localización	localisation	localizzazione	localização	1) localization : The act of localizing	LST
N 2732	localization	localización	localisation	localizzazione	localização	2) localization : The state of being localized	LST
V 2587	localize	localizar	localiser	localizzare	localizar	To determine where something takes place or is to be found	LST
V 5336	locate	localizar	localiser	localizzare	localizar	To designate the site or place of; to define the limits of (...)	LST
N 5687	location	posición	position / emplacement	posizione	posição	A particular point or place in physical space	LST
N 7496	locus	* locus	locus	locus	locus	(genetics) A fixed position on a chromosome that may be occupied by one or more genes	Lspé
N 2850	log	log	log	log	log	(mathematics) Logarithm	LST
A 2389	logistic	logístico	logistique	logistico	logístico	Relating to logistics	LST
A 3935	long	largo	long	lungo	longo	1) long : Having great duration	LST
A 3935	long	largo	long	lungo	longo	2) long : Having much distance from one terminating point (...) to another point (...)	LST
A 4464	long-term	a largo plazo	à long terme	a lungo termine	a longo prazo	Extending over a relatively long time period	LST
A 1156	longitudinal	longitudinal	longitudinal	longitudinale	longitudinal	1) longitudinal : Sampling data over time rather than merely once	LST
A 1156	longitudinal	longitudinal	longitudinal	longitudinale	longitudinal	2) longitudinal : Running in the direction of the long axis of a body	Lspé
V 2196	look	buscar	chercher	cercare	buscar	To search for, to try to find	LST
N 3100	loop	bucle	anse / boucle	ansa	laço	(biochemistry) A flexible region in a protein's secondary structure	Lspé
V 2699	lose	perder	perdre	perdere	perder	To cause to cease to be in one's possession or capability (...)	Lspé
N 9694	loss	pérdida	perte	perdita	perda	The result of an alteration in a function or characteristic of the body, or of its previous integrity	Lspé
A 36943	low	bajo	bas	basso	baixo	Not high in amount or quantity	LST
V 1907	lower	abajar / bajar	abaisser / baisser	abbassare	abaixar / baixar	To reduce the degree, intensity, strength, etc.	LST

N 15021	lung	pulmón	poumon	polmone	pulmão	(anatomy) A biological organ that extracts oxygen from the air	Lspé
N 2413	lymph	linfa	lymphe	linfa	linfa	(physiology, immunology) A colorless, watery, bodily fluid carried by the lymphatic system, that consists mainly of white blood cells.	Lspé
N 5609	lymphocyte	linfocito	lymphocyte	linfocito / linfocita	linfócito	(cytology) A type of white blood cell or leukocyte (...)	Lspé
N 1578	lysate	* lisado	lysat	* lisato	lisado	(biology) All the material formed by the lysis of cells	Lspé
N 7247	macrophage	* macrófago	macrophage	macrofago	macrófago	(immunology, cytology) A white blood cell that phagocytes necrotic cell debris and foreign material, including viruses, bacteria (...)	Lspé
A 2328	magnetic	magnético	magnétique	magnetico	magnético	Of, relating to, operating by, or caused by magnetism	Lspé
N 2382	magnitude	magnitud	ampleur	magnitudine	magnitude	The absolute or relative size, extent or importance of something	LST
A 6157	main	principal	principal	principale	principal	Principal ; prime ; chief ; leading	LST
V 6267	maintain	mantener	maintenir	mantenere	manter	To keep up; to preserve; to uphold (a state, condition etc.)	LST
N 2236	maintenance	mantenimient o	maintenance	mantenimento	mantimento	(biology) The natural process which keeps an organism alive	Lspé
A 12874	major	mayor	majeur	maggiore	maior	Of great significance or importance	LST
N 4987	majority	mayoría	majorité	maggioranza	maioria	More than half (50%) of some group	LST
V 18946	make	hacer	faire	fare	fazer	To constitute	LST
N 6737	malaria	malaria	malaria	malaria	malária	A disease spread by mosquito, in which a protozoan, Plasmodium, multiplies in blood (...)	Lspé
N 2485	male	hombre / varón	homme	uomo	homem	1) male : A human member of the masculine sex or gender	Lspé
N 2485	male	macho	mâle	maschio	macho	2) male : An animal of the sex that has testes	Lspé
A 5187	male	masculino	masculin / mâle	maschio	masculino / macho	Belonging or referring to the sex having the smaller gametes	Lspé
A 3010	malignant	maligno	malin	maligno	maligno	Tending to produce death ; threatening a fatal issue	Lspé
N 1562	mammal	mamífero	mammifère	mammifero	mamífero	An animal of the class Mammalia, characterized by being warm-blooded, having hair and feeding milk to its young	Lspé
A 4943	mammalian	mamífero	mammalien	mammaliano	mamífero	Of, or pertaining to mammals	Lspé
A 3145	mammary	mamario	mammaire	mammario	mamário	Of or relating to mamma or breast	Lspé
N 8814	man	hombre	homme	uomo	homem	An adult male human	Lspé
V 2096	manage	manejar / gestionar	gérer	gestire	gerir / gerenciar	To handle with skill	LST
N 7446	management	gestión	management /	amministrazio	gerência	Administration : the process or practice of managing	LST

N 3425	manner	manera / modo	gestion manière / façon	ne / gestione maniera / modo	maneira / modo	Mode of action ; way of performing or effecting anything ; method ; style (...)	LST
A 2222	manual	manual	manuel	manuale	manual	Performed with the hands	LST
N 2634	manufacturer	fabricante	fabriquant	fabbricante	fabricante	One who manufactures	LST
N 12082	manuscript	manuscrito	manuscrit	manoscritto	manuscrito	A single, original copy of a book, article, composition (...)	LST
N 6478	map	mapa	carte	mappa	mapa	Graphical representation of relations existing between several objects, components or themes	LST
V 2938	map	cartografiar	cartographier	mappare	cartografar	To create a visual representation of a territory, etc. via cartography	Lspé
N 3306	mapping	cartografía	cartographie	mappatura	cartografia	(biology) The process of locating genes on a chromosome	Lspé
N 6321	marker	marcador	marqueur	marcatore	marcador	1) marker : (medicine) A substance used as an indicator of a biological state, most commonly disease	Lspé
N 6321	marker	marcador	marqueur	marcatore	marcador	2) marker : (genetics) A gene or DNA sequence with a known location on a chromosome that can be used to identify individuals or species	Lspé
N 1791	marrow	médula	moelle	midollo	medula	The substance inside bones which produces blood cells	Lspé
N 2576	mass	masa	masse	massa	massa	1) mass : (Physics) The quantity of matter which a body contains, irrespective of its bulk or volume. It is one of four fundamental properties of matter. It is measured in kilograms in the SI system of measurement	LST
N 2576	mass	masa	masse	massa	massa	2) mass : (medicine) A palpable or visible abnormal globular structure ; a tumor	Lspé
V 2477	match	combinar	combiner	combinare	combinar	1) match : To fit together, or make suitable for fitting together (...)	Lspé
V 2477	match	corresponder	correspondre	corrispondere	corresponder	2) match : To agree, to be equal, to correspond to	LST
N 2047	match	correspondencia	correspondance	corrispondenza	correspondência	The equivalence ; a state of correspondence	LST
N 7484	material	material	matériel	materiale	material	A sample or specimens for study	LST
A 3336	maternal	maternal	maternel	materno	maternal	Of or pertaining to a mother	Lspé
N 3086	matrix	matriz	matrice	matrice	matriz	1) matrix : (biology) The material or tissue between the cells of animals and plants	Lspé
N 3086	matrix	matriz	matrice	matrice	matriz	2) matrix : (biology) The medium in which bacteria are cultured	Lspé
N 3086	matrix	matriz	matrice	matrice	matriz	3) matrix : (mathematics) A rectangular arrangement of numbers or terms	LST
N 2132	maturation	maduración	maturation	maturazione	maturação	(biology) The process of differentiation that produces the adult form of an organism	Lspé
A 1669	mature	maduro	mûr	maturo	maduro	Fully developed	Lspé
A 2039	maximal	máximo	maximal	massimo	máximo	Largest, greatest	LST
A 4518	maximum	máximo	maximum	massimo	máximo	To the highest degree	LST
N 2462	mean	medio	moyen	mezzo	meio	1) mean : (statistics) The average of a set of values, calculated by summing them together and dividing by the number of terms ; the arithmetic mean	LST
N 2462	mean	media	moyenne	media	média	2) mean : A method or course of action used to achieve some result	LST

V 6630	mean	significar / querer decir	signifier / vouloir dire	significare / volere dire	significar / querer dizer	To convey (a given sense); to signify, or indicate (an object or idea)	LST
A 11217	mean	medio	moyen	medio	médio	Having the mean as its value	LST
N 6795	measure	medida	mesure	misura	medida	1) measure : The quantity, size, weight, distance or capacity of a substance compared to a designated standard	LST
N 6795	measure	medida	mesure	misura	medida	2) measure : A tactic, strategy (...)	LST
V 16944	measure	medir	mesurer	misurare	medir	To ascertain the quantity of a unit of material via calculated comparison with respect to a standard	LST
N 12551	measurement	medición	mesure	misurazione / misura	medição	The act of measuring	LST
A 2383	mechanical	mecánico	mécanique	meccanico	mecânico	Related to or done by machine	LST
N 17292	mechanism	mecanismo	mécanisme	meccanismo	mecanismo	A mental, physical or chemical process	LST
A 1869	median	mediano	médian	mediano	mediano	1) median : Having the median as its value	LST
A 1869	median	mediano	médian	mediano	mediano	2) median : (anatomy) median : In the middle of an organ, structure, etc. ; towards the median plane of an organ or limb	Lspé
V 13726	mediate	mediar	médiatiser	mediare	mediar	To act as an intermediary causal or communicative agent; convey	Lspé
N 2075	mediator	mediador	médiateur	mediatore	mediador	A chemical substance transmitting information to a targeted cell	Lspé
A 17423	medical	médico	médical	medico	médico	Of or pertaining to the practice of medicine	Lspé
N 4178	medication	medicación	médication	medicazione	medicação	A medicine, or all the medicines regularly taken by a patient	Lspé
N 3826	medicine	medicina	médecine	medicina	medicina	1) medicine : The study of the cause, diagnosis, prognosis and treatment of disease or illness	Lspé
N 3826	medicine	medicina / medicamento	médicament	medicina / medicamento	medicamento	2) medicine : A substance which specifically promotes healing when ingested or consumed in some way	Lspé
N 10972	medium	medio	milieu	mezzo	médio	(microbiology) A nutrient solution for the growth of cells in vitro	Lspé
V 3187	meet	satisfacer / responder	satisfaire / répondre	soddisfare / rispondere	satisfazer / responder	To satisfy; to comply with	LST
N 2614	melanoma	melanoma	mélanome	melanoma	melanoma	(oncology, pathology) A dark-pigmented, usually malignant tumor arising from a melanocyte and occurring most commonly in the skin	Lspé
N 10370	member	miembro	membre	membro	membro	One who officially belongs to a group	LST
N 15316	membrane	membrana	membrane	membrana	membrana	A flexible or separate tissue forming a plane or film and separating two environments (...)	Lspé
N 1443	memory	memoria	mémoire	memoria	memória	1) memory : The ability of an organism to record information (...)	Lspé
N 1443	memory	memoria	mémoire	memoria	memória	2) memory : (computing) The part of a computer that stores variable executable code or data (RAM) or unalterable executable code or default data (ROM)	LST
A 3998	mental	mental	mental	mentale	mental	Of or related to the mind or an intellectual process	Lspé
V 2430	mention	mencionar	mentionner	menzionare	mencionar	To make a short reference to something	LST

A 5014	metabolic	metabólico	métabolique	metabolico	metabólico	Of or pertaining to metabolism	Lspé
N 5114	metabolism	metabolismo	métabolisme	metabolismo	metabolismo	(Physiology) The complete set of chemical reactions that occur in living cells	Lspé
N 1715	metabolite	metabolito	métabolite	metabolito / metabolita	metabólito	Any substance produced by, or taking part in, a metabolic reaction	Lspé
N 4069	metastasis	metástasis	métastase	metastasi	metástase	(medicine) The transference of a bodily function or disease to another part of the body, specifically the development of a secondary area of disease remote from the original site, as with some cancers	Lspé
A 2311	metastatic	metastásico	métastatique	metastatico	metastático	Relating to metastasis	Lspé
N 44600	method	método	méthode	metodo	método	A process by which a task is completed ; a way of doing something (...)	LST
N 2324	methodology	metodología	méthodologie	metodologia	metodologia	A collection of methods, practices, procedures and rules used by those who work in some field	LST
N 2749	methylation	metilación	méthylation	metilazione	metilação	(genetics) The addition of a methyl group to cytosine and adenine residues in DNA that leads to the epigenetic modification of DNA and the reduction of gene expression and protein production	Lspé
N 12502	mg	mg	mg	mg	mg	(metrology) Symbol for the milligram (...)	LST
N 2326	MI	IAM / IMA *	IDM	IMA	IAM / EAM	Myocardial Infarction	Lspé
N 11139	microarray	microarreglo /* micromatriz	* microréseau / * micromatrice	microarray / micromatrice	microarranjo / *microséries	Any of several devices containing a two-dimensional array of small quantities of biological material used for various types of assay ; the array itself	Lspé
N 1630	microscope	microscopio	microscope	microscopio	microscópio	An optical instrument for imaging very small objects (...)	Lspé
N 2694	microscopy	microscopia / microscopía	microscopie	microscopia	microscopia	The use of microscopes	Lspé
N 3882	migration	migración	migration	migrazione	migração	(biochemistry) The movement of cells in particular directions to specific locations	Lspé
A 2542	mild	leve	léger	lieve	leve	(medicine) Not serious or dangerous	Lspé
N 1611	milk	leche	lait	latte	leite	A white liquid produced by the mammary glands of female mammals to nourish their youngs (...)	Lspé
N 15198	min	min	min. / mn / m	min / m	min	Abbreviation of minute	LST
A 2662	minimal	mínimo	minimum	minimo	mínimo	Smallest possible amount, quantity, or degree	LST
V 1854	minimize	minimizar	minimiser	minimizzare	minimizar	To make as small or as insignificant as possible	LST
A 1973	minimum	mínimo	minimum	minimo	mínimo	Smallest possible amount, quantity, or degree	LST
A 1975	minor	menor	mineur	minore	menor	Of little significance or importance	LST
N 7691	minute	minuto	minute	minuto	minuto	A unit of time equal to sixty seconds (...)	LST
V 2953	miss	faltar	manquer	mancare	faltar	To be absent, deficient, or wanting	LST
A 4345	mitochondria	*	mitochondrial	mitocondriale	mitocondrial	Of, or relating to mitochondria	Lspé

	l	mitochondrial						
N 2102	mitochondrion	mitocondria	mitochondrie	mitocondrio	mitocôndria	(cytology) A spherical or ovoid organelle found in the cytoplasm of eukaryotic cells (...)	Lspé	
V 1637	mix	mezclar	mélanger	mescolare	mesclar	To form by mingling; to produce by the stirring together of ingredients; to compound of different parts	Lspé	
A 2174	mixed	mixto	mélangé	misto	misto	Having two or more separate aspects	LST	
N 3199	mixture	mezcla / mezclado	mixture	miscuglio / miscela	mistura	Something that consists of diverse elements	Lspé	
N 21049	ml	ml	ml	ml / mL	ml	(metrology) Symbol for the milliliter (...)	LST	
N 17751	mm	mm	mm	mm	mm	(metrology) Symbol for the millimeter (...)	LST	
N 1653	mmol	mmol	mmol	mmol	mmol	(metrology) Symbol for the millimole (...)	Lspé	
N 2678	mode	modo	mode	modo	modo	A particular means of accomplishing something	LST	
N 43961	model	modelo	modèle	modello	modelo	The structural design of a complex system	LST	
V 2367	model	modelizar	modéliser	modellare	modelar	To use as an object in the creation of a forecast or model	LST	
A 3103	moderate	moderado	modéré	moderato	moderado	Not excessive	LST	
N 3988	modification	modificación	modification	modificazione	modificação	The act or result of modifying or the condition of being modified	LST	
A 1935	modified	* modificado	modifié	modificato	modificado	Changed, altered	LST	
V 3706	modify	modificar	modifier	modificare	modificar	To make partial changes to	LST	
V 3254	modulate	modular	moduler	modulare	modular	To regulate, adjust or adapt	LST	
N 2214	modulation	modulación	modulation	modulazione	modulação	The variation and regulation of a population, physiological response, etc.	LST	
N 2609	module	módulo	module	modulo	módulo	A self-contained component of a system, often interchangeable, which has a well-defined interface to the other components	LST	
A 15927	molecular	molecular	moléculaire	molecolare	molecuar	Relating to molecules	Lspé	
N 9363	molecule	molécula	molécule	molecola	molécula	(chemistry) The smallest particle of a specific element or compound that retains the chemical properties of that element or compound ; two or more atoms held together by chemical bonds	Lspé	
V 3581	monitor	monitorizar	* monitorer / surveiller	monitorare	monitorar / monitorizar	To watch over; to guard	Lspé	
N 2204	monitoring	* monitoreo	monitoring / monitorage / surveillance	monitoraggio	monitoração / monitorizaçã o	(medicine) The continuous observation of a patient's health condition for any changes which may occur over time, using a monitor or measuring device of some sort	Lspé	
N 1923	monkey	mono	singe	scimmia	mono	Any member of the clade Simiiformes not also of the clade Hominoidea (...)	Lspé	
A 3279	monoclonal	monoclonal	monoclonal	monoclonale	monoclonal	(immunochemistry) Produced from a single clone	Lspé	
N 3303	monocyte	monocito	monocyte	monocito	monócito	(immunology, cytology) A type of blood leukocyte that differentiates into a macrophage.	Lspé	

N	10744	month	mes	mois	mese	mês	A period into which a year is divided, historically based on the phases of the moon (...)	LST
N	2925	morbidity	morbilidad	morbidity	morbosità	morbidade / morbilidade	(medicine) The incidence of a disease, as a rate of a population which is affected	Lspé
A	2034	morphologic	morfológico	morphologique	morfologico	morfológico	Of or pertaining to morphology ; morphological	Lspé
N	2258	morphology	morfología	morphologie	morfologia	morfologia	(biology) The study of the form and structure of animals and plants	Lspé
N	10488	mortality	mortalidad	mortalité	mortalità	mortalidade	(demography) The death rate of a population	Lspé
N	3248	mosquito	mosquito	moustique	zanzara	mosquito	A small flying insect of the family Culicidae, known for biting and sucking blood, leaving an itching bump on the skin (...)	Lspé
N	3013	mother	madre	mère	madre	mãe	A (human) female who (a) parents a child (b) gives birth to a baby (c) donates a fertilized egg or (d) donates a body cell which has resulted in a clone (...)	Lspé
N	7824	motif	motivo	motif	motivo	motivo	In a nucleotide or amino-acid sequence, pattern that is widespread and has, or is conjectured to have, a biological significance	Lspé
A	2678	motor	motor	moteur	motore / motorio	motor	(biology) describing neurons that create the ability to move	Lspé
N	39502	mouse	ratón	souris	topo	rato	Any small rodent of the genus Mus.	Lspé
V	1776	move	mover	déplacer	muover	mover	To change place or posture; to stir; to go, in any manner, from one place or position to another	Lspé
N	3001	movement	movimiento	mouvement	movimento	movimento	Physical motion between points in space	Lspé
N	14581	mRNA	ARNm	ARNm	mRNA	ARNm	(biochemistry) Messenger RNA	Lspé
N	5699	ms	ms	ms	ms	ms	(metrology) Symbol for the millisecond (...)	LST
A	15436	multiple	múltiple *	multiple	multiplo	múltiplo	Having more than one element, part, component, or function	LST
A	1979	multivariate	multivariado / * multivariable	multivarié / multivariable	* multivariato	multivariado / multivariável	Having or involving multiple variables	LST
A	3803	murine	* murino	murin	murino	murino / murídeo	Of or pertaining to, or characteristic of the mouse, rat or any mammal of the family	Lspé
N	11802	muscle	músculo	muscle	muscolo	músculo	A contractile form of tissue (...)	Lspé
N	4201	mutant	mutante	mutant	mutante	mutante	Something which has mutated, which has one or more new characteristics from a mutation	Lspé
A	6310	mutant	mutante	mutant	mutante	mutante	That has mutated	Lspé
N	19671	mutation	mutación	mutation	mutazione	mutação	Any heritable change of the base-pair sequence of genetic material	Lspé
A	4175	myocardial	miocárdico	myocardique	miocardico	miocárdico	Relating to the myocardium	Lspé
N	2992	name	nombre	nom	nome	nome	Any nounal word or phrase which indicates a particular person, place, class, or thing	LST
A	9964	national	nacional	national	nazionale	nacional	Of or having to do with a country	LST



A	2195	native	nativo	natif	nativo	nativo	Belonging to one by birth	Lspé
A	4786	natural	natural	naturel	naturale	natural	Relating to nature	Lspé
N	10521	nature	naturaleza	nature	natura	natureza	The innate characteristic of a thing. What something will tend by its own constitution, to be or do (...)	Lspé
A	5635	necessary	necesario	nécessaire	necessario	necessário	Needed, required	LST
N	1937	neck	cuello	cou	collo	pescoço	The part of the body connecting the head and the trunk (...)	Lspé
N	3629	necrosis	necrosis	nécrose	necrosi	necrose	(pathology) The localized death of cells or tissues through injury, disease, or the interruption of blood supply	Lspé
N	6214	need	necesidad	besoin	bisogno	necessidade	A requirement for something	LST
V	5178	need	necesitar / precisar	nécessiter / avoir besoin	necessitare / avere bisogno	necessitar / precisar	To be obliged or required (to do something)	LST
A	5733	negative	negativo	négatif	negativo	negativo	1) negative : Damaging; undesirable; unfavourable	Lspé
A	5733	negative	negativo	négatif	negativo	negativo	2) negative : Characterized by the presence of features which do not support a hypothesis	LST
A	1845	neonatal	neonatal	néonatal	neonatale	neonatal	Of or pertaining to the period of time immediately following birth or to the newborn	Lspé
N	3616	nerve	nervio	nerf	nervo	nervo	(anatomy) A bundle of neurons with their connective tissue sheaths, blood vessels and lymphatics	Lspé
A	2150	nervous	nervioso	nerveux	nervoso	nervoso	Relating to or affecting the nerves	Lspé
N	9044	network	red	réseau	rete	rede	Any interconnected group or system	LST
A	3241	neural	neural	neural	neurale	neural	(biology) Of, or relating to the nerves or the nervous system	Lspé
N	7653	neuron	neurona	neurone	neurone	neurônio / neurone	(cytology) A cell of the nervous system, which conducts nerve impulses (...)	Lspé
A	2994	neuronal	neuronal	neuronal	neuronale	neuronal	Of, or relating to a neuron	Lspé
N	3725	neutrophil	neutrófilo	neutrophile	neutrofilo	neutrófilo	(biology, medicine) (...) A white blood cell	Lspé
A	21655	new	nuevo	nouveau	nuovo	novo	Additional; recently discovered	LST
A	2394	next	siguiente	suivant	segunte	seguinte	Following in a sequence	LST
A	2690	nitric	nítrico	nitrique	nitrico	nítrico	O, pertaining to, or containing nitrogen	Lspé
N	1824	nitrogen	nitrógeno	azote	azoto	nitrogênio	A chemical element (symbol N) with an atomic number of 7 and atomic weight of 14.0067	Lspé
N	3915	nm	nm	nm	nm	nm	(metrology) Symbol for the nanometer (...)	Lspé
N	5890	node	nódulo / ganglio	nodule / ganglion	nodulo / ganglio	nódulo / gânglio	(medicine) A hard concretion or incrustation which forms upon bones attacked with rheumatism, gout, or syphilis; sometimes also, a swelling in the neighborhood of a joint	Lspé
N	1003	noise	ruido	bruit	rumore	ruído	1) noise : (genetics) The measured level of variation in gene expression among cells, regardless of source, within a supposedly identical population	Lspé
N	1003	noise	ruido	bruit	rumore	ruído	2) noise : (technology) Unwanted part of a signal	LST

A 9768	normal	normal	normal	normale	normal	1) normal : (medicine) Healthy; not sick or ill	Lspé
A 9768	normal	normal	normal	normale	normal	2) normal : (probability theory, statistics) (Of a distribution) which has a very specific Bell curve shape	LST
N 2911	normalization	normalización	normalisation	normalizzazione	normalização	Any process that makes something more normal or regular, which typically means conforming to some regularity or rule, or returning from some state of abnormality	LST
V 2745	normalize	normalizar	normaliser	normalizzare	normalizar / normatizar	To format in a standardized manner, to make consistent	LST
A 1595	north	del norte	du nord	del nord	do norte	Of or pertaining to the north; northern	LST
V 7470	note	notar	noter	notare	notar	To record in writing; to make a memorandum of	LST
A 8957	novel	nuevo	nouveau	nuovo	ново	New, original, especially in an interesting way	LST
A 9558	nuclear	nuclear	nucléaire	nucleare	nuclear	Relating to the nucleus of a cell	Lspé
A 5248	nucleic	nucleico	nucléique	nucleico	nucleico	Referring to the nucleus of something	Lspé
N 8319	nucleotide	nucleótido	nucléotide	nucleotide	nucleotídeo	(biochemistry) The monomer comprising DNA or RNA biopolymer molecules (...)	Lspé
N 5061	nucleus	núcleo	noyau	nucleo	núcleo	(cytology) A large organelle found in cells which contains genetic material	Lspé
A 1203	null	nulo	nul	nullo	nulo	1) null : (mutation) Causing a complete loss of gene function, amorphic	Lspé
A 1203	null	nulo	nul	nullo	nulo	2) null : (mathematics) of the null set	LST
N 47731	number	número	nombre	numero	número	An abstract entity used to describe quantity	LST
A 2444	numerous	numeroso	nombreux	numeroso	numeroso	Indefinitely large numerically	LST
N 3881	nurse	enfermero	infirmier	infermiere	enfermeiro	A person trained to provide care for the sick	Lspé
N 2252	nursing	enfermería	soins infirmiers	assistenza infermieristica	enfermagem	The process or profession of caring for patients as a nurse	Lspé
N 1867	nutrition	nutrición	nutrition	nutrizione	nutrição	(biology) The organic process by which an organism assimilates food and uses it for growth and maintenance	Lspé
N 2897	obesity	obesidad	obésité	obesità	obesidade	The state of being obese due to an excess of body fat	Lspé
N 2491	objective	objetivo	objectif	obiettivo	objetivo	A goal that is striven for	LST
N 9449	observation	observación	observation	osservazione	observação	The act of noting and recording some event, or the record of such noting	LST
V 21978	observe	observar	observer	osservare	observar	To notice or view, especially carefully or with attention to detail	LST
V 23330	obtain	obtener	obtenir	ottenere	obter	To get hold of; to gain possession of, to procure; to acquire, in any way	LST
A 1610	obvious	obvio	évident	ovvio	óbvio	Easily discovered or understood, self-explanatory	LST
V 15739	occur	ocurrir / acontecer / suceder	avoir lieu	avvenire	ocorrer / acontecer / suceder	(sciences) To be present or found	LST
N 3369	occurrence	ocurrencia	occurrence	occorrenza	ocorrência	Actual instance when a situation arises	LST
N 2703	odds	momios	chances	odds	chances	The probabilities of an event happening	LST

V 3282	offer	ofrecer	offrir	offrire	oferecer	To place at someone's disposal; to present to be either accepted or turned down	LST
N 819	office	consultorio	cabinet	studio	consultório	1) office : A building or room where clerical or professional duties are performed	LST
N 819	office	oficina	bureau	ufficio	escritório	2) office : A bureau, an administrative unit of government	LST
N 1690	oil	aceite	huile	olio	óleo	Liquid fat	Lspé
A 8687	old	viejo	vieux	vecchio	velho	Having existed for a relatively long period of time	Lspé
N 4045	oligonucleotide	* oligonucleótido	oligonucléotide	oligonucleotide	oligonucleotídeo / oligonucleótido	A short sequence of nucleotides (RNA or DNA), typically with twenty or fewer base pairs	Lspé
N 2699	oncogene	oncogén	oncogène	oncogene	oncógene	(oncology, genetics) Any gene that contributes to the conversion of a normal cell into a cancerous cell when mutated or expressed at high levels	Lspé
A 5097	only	solo	seul	solo	só	Alone in a category	LST
N 4231	onset	aparición	apparition	insorgenza	aparicação	(medicine) The initial phase of a disease or condition, in which symptoms first become apparent.	Lspé
N 1747	ontology	ontología	ontologie	ontologia	ontologia	(computer science, information science) A structure of concepts or entities within a domain, organized by relationships; a system model	LST
N 3027	oocyte	oocito / ovocito	ovocyte	oocito / ovocito	oócito / ovócito	(cytology) A cell that develops into an egg or ovum ; a female gametocyte	Lspé
A 1397	open	abierto	ouvert	aperto	aberto	1) open : (computing) Made public, usable without a free license	LST
A 1397	open	abierto	ouvert	aperto	aberto	2) open : Not settled or adjusted ; not decided or determined (...)	LST
A 1397	open	abierto	ouvert	aperto	aberto	3) open : (medicine) Resulting from an incision, puncture or any other process by which the skin no longer protects an internal part of the body	Lspé
V 1721	operate	operar	opérer	operare	operar	(medicine) To perform some manual act upon a human body in a methodical manner, and usually with instruments, with a view to restore soundness or health (...)	Lspé
N 1967	operation	operación	opération	operazione	operação	(medicine) A surgical procedure	Lspé
N 1791	operon	* operón	opéron	operone	* operão	A unit of genetic material that functions in a coordinated manner by means of an operator, a promoter, and structural genes that are transcribed together	Lspé
N 2125	opportunity	oportunidad / ocasión	opportunité / occasion	opportunità / occasione	oportunidade / ocasião	A favorable circumstance or occasion	LST
A 3694	optimal	óptimo	optimal	ottimale	ótimo	The best, most favorable or desirable	LST
V 1877	optimize	optimizar	optimiser	ottimizzare	otimizar	To make optimal	LST
N 2475	option	opción	option	opzione	opção	One of the choices which can be made	LST
A 4898	oral	oral	oral	orale	oral	Relating to the mouth	Lspé
N 4900	order	orden	ordre	ordine	ordem	(mathematics) The cardinality, or number of elements in a set or related structure	LST
N 5312	organ	órgano	organe	organo	órgão	(medicine) A largest part of an organism, composed of tissues that perform similar functions	Lspé

N 5757	organism	organismo	organisme	organismo	organismo	(biology) A discrete and complete living thing, such as animal, plant, fungus or microorganism	Lspé
N 3326	organization	organización	organisation	organizzazione	organização	1) organization : The way in which something is organized (...)	LST
N 3326	organization	organización	organisation	organizzazione	organização	2) organization : A group of people or other legal entities with an explicit purpose and written rules	LST
N 1626	orientation	orientación	orientation	orientamento	orientação	An inclination, tendency or direction	LST
N 4866	origin	origen	origine	origine	origem	The source (...)	LST
A 4511	original	original	original	originale	original	First in a series or copies / versions	LST
N 1928	ortholog	ortólogo	* orthologue	* ortologo	* ortólogo	(genetics) either of two or more homologous gene sequences found in different species related by linear descent	Lspé
N 1678	osteoarthritis	osteoartritis	ostéoarthrite	osteoartrite / osteoartrosi	osteoartrite / osteoartrose	(pathology) A form of arthritis, affecting mainly older people, caused by chronic degeneration of the cartilage and synovial membrane of the joints, leading to pain and stiffness	Lspé
A 51502	other	otro	autre	altro	outro	Not the one previously referred to	LST
N 15543	outcome	resultado	résultat	risultato	resultado	Information, event, object or state of being produced as a result or consequence of a plan, process, accident, effort or other similar action or occurrence	LST
N 1626	output	salida	sortie	output	saída	1) output : (computing) Data sent out to the computer, as to output device such as a monitor or printer	LST
N 1626	output	débito / gasto	débit	gittata	débito	2) output : (medicine) The flow rate of body liquids such as blood and urine	Lspé
A 4589	ovarian	ovárico	ovarien	ovarico	ovárico	Relating to the ovaries	Lspé
N 1951	ovary	ovario	ovaire	ovario / ovaia	ovário	(anatomy) A female reproductive organ, often paired, that produces ova and in mammals secretes the hormones estrogen and progesterone	Lspé
A 7632	overall	global	global	complessivo	global	All-encompassing	LST
N 2888	overexpression	sobreexpresión	* surexpression	sovraespressione	sobreexpressão	(genetics) The artificial expression of a gene in increased quantity	Lspé
V 2670	overlap	superponerse	se chevaucher	sovrapporsi	sobrepor-se	(genetics) To have some similar nucleotide sequences	Lspé
A 2703	own	propio	propre	proprio	próprio	Belonging to	LST
N 1646	oxidation	oxidación	oxydation	ossidazione	oxidação	(chemistry) The combination of a substance with oxygen	Lspé
A 3088	oxidative	* oxidativo	oxydatif	ossidativo	oxidativo	Of, relating to, or produces by oxidation	Lspé
N 2976	oxide	óxido	oxyde	ossido	óxido	(chemistry) A binary chemical compound of oxygen with another chemical element	Lspé
N 2239	oxygen	oxígeno	oxygène	ossigeno	oxigênio	1) oxygen : (medicine) A mixture of oxygen and other gases, administered to a patient to help him or her to breathe	Lspé
N 2239	oxygen	oxígeno	oxygène	ossigeno	oxigênio	2) oxygen : A chemical element (symbol O) with an atomic number of 8 and relative atomic mass of 15.9994	Lspé
N 2442	package	paquete	paquet	pacchetto	pacote	Something which consists of various components, such as a piece of computer software	LST

N	3089	page	página	page	pagina	página	(Internet) A web page	LST
N	11320	pain	dolor	douleur	dolore	dor	An ache or bodily suffering (...)	Lspé
V	1670	pair	emparejar	apparier / coupler	accoppiare	acoplar	To group into sets of two	LST
N	11063	pair	par	paire	paio	par	Two similar or identical things taken together (...)	LST
A	3092	pancreatic	pancreático	pancréatique	pancreatico	pancreático	Of or relating to the pancreas	Lspé
N	3946	paper	artículo	article	articolo	artigo	1) paper : A written document that reports scientific or academic research and is usually subjected to peer review before publication (...)	LST
N	3946	paper	papel	papier	carta	papel	2) paper : A sheet material (...)	LST
A	1738	parallel	paralelo	parallèle	parallelo	paralelo	Equally distant from one another at all points	LST
N	11740	parameter	parámetro	paramètre	parametro	parâmetro	A variable kept constant during an experiment, calculation or similar	LST
N	4847	parasite	parásito	parasite	parassita	parasita	(biology) an organism that lives on or in another organism, deriving benefit from living on or in that other organism	Lspé
N	3156	parent	padre	parent	genitore	pai	One of the two persons from whom one is immediately biologically descended (...)	Lspé
N	12971	part	parte	partie	parte	parte	A fraction of a whole	LST
A	3187	partial	parcial	partiel	parziale	parcial	Existing as a part or portion ; incomplete	LST
N	8173	participant	participante	participant	partecipante	participante	One who participates	LST
V	9674	participate	participar	participer	partecipare	participar	To join in, to take part, to involve oneself	LST
N	2137	participation	participación	participation	partecipazione	participação	The act of participating, of taking part to a process	LST
N	3654	particle	partícula	particule	particola	partícula	A very small piece of matter (...)	Lspé
A	5444	particular	particular	particulier	particolare	particular	Specialized ; characteristic of a specific person or thing	LST
N	2290	partner	socio / compañero	partenaire	partner	sócio / parceiro	Someone who is associated with another in a common activity or interest	LST
A	2671	past	pasado	passé	passato	passado	Having already happened ; in the past ; finished	LST
N	2335	pathogen	agente patogénico	agent pathogène	agente patogeno	agente patogênico	(pathology, immunology) Any organism or substance, especially a microorganism, capable of causing disease, such as bacteria, viruses, protozoa or fungi (...)	Lspé
N	2765	pathogenesis	patogenia	pathogénie	patogenesi	patogenia	The mechanism whereby something causes a disease	Lspé
A	2221	pathological	patológico	pathologique	patologico	patológico	Pertaining to pathology	Lspé
N	2284	pathology	patología	pathologie	patologia	patologia	(medicine) The branch of medicine concerned with the study of the nature of disease and its causes, processes, development, and consequences	Lspé
N	16162	pathway	vía / ruta	voie	* pathway / via	via	(biochemistry) A sequence of biochemical compounds and the reactions linking them, that describe a process in metabolism or catabolism	Lspé
N	11378 1	patient	paciente	patient	paziente	paciente	A person or animal who receives treatment from a doctor or other medically educated person	Lspé

N 19319	pattern	modelo	modèle	modello	modelo	Model, example	LST
N 2531	PC	PC	PC	PC	PC	A small computer, built around a microprocessor (...)	LST
N 4521	PCR	PCR	PCR / ACP	PCR	PCR	polymerase chain reaction (biochemistry) a technique in molecular biology for creating multiple copies of DNA from a sample (...)	Lspé
N 3846	peak	pico	pic	picco	pico	The highest value reached by some quantity in a time period	LST
A 1942	pediatric	pediátrico	pédiatrique	pediatrico	pediátrico	Of or pertaining to the pediatrics (...)	Lspé
N 6259	people	gente	gens	gente	gente	A body of human beings considered generally or collectively (...)	LST
N 11671	peptide	péptido	peptide	peptide	péptido / peptídio	A class of organic compounds consisting of various numbers of amino-acids (...)	Lspé
V 2330	perceive	percibir	percevoir	percepire	perceber	To see, to be aware of, to understand	LST
N 13085 7	percent	por ciento	pourcent / pour cent	per cento / per cento	por cento	A part or other object per hundred	LST
N 6215	percentage	porcentaje	pourcentage	percentuale	percentagem	The amount, number or rate of something, regarded as a part of a total of 100 ; a part of a whole	LST
N 2086	perception	percepción	perception	percezione	percepção	The organization, identification, and interpretation of sensory information	Lspé
V 27219	perform	ejecutar / cumplir	exécuter	eseguire / compiere	executar / cumprir	To do something; to execute	LST
N 7034	performance	desempeño	performance	performance	desempenho	The amount of useful work accomplished estimated in terms of time needed, resources used, etc.	LST
N 1620	perfusion	perfusión	perfusion	perfusione	perfusão	(medicine) The introduction of a drug or nutrients through the bloodstream in order to reach an internal organ or tissues	Lspé
N 13357	period	período / periodo	période	periodo	período	A length of time	LST
A 5178	peripheral	periférico	périphérique	periferico	periférico	(neuroanatomy) A part of or located in the peripheral nervous system	Lspé
V 1556	permit	permitir	permettre	permettere	permitir	To allow for, to make something possible	LST
N 5051	person	persona	personne	persona	peessoa	An individual (...)	LST
A 2682	personal	personal	personnel	personale	peessoal	Of or pertaining to a particular person	LST
N 2632	perspective	perspectiva	perspective	prospettiva	perspectiva	The choice of a single angle or point of view from which to sense, categorize, measure or codify experience	LST
N 9489	pH	pH	pH	pH	pH	(chemistry) Measure of the acidity or basicity of an aqueous solution	Lspé
N 2031	phage	fago	phage	fago	fago	(microbiology, virology) A virus that is parasitic on bacteria	Lspé
N 11505	phase	fase	phase	fase	fase	A distinguishable part of a sequence or cycle occurring over time	LST
N 2511	phenomenon	fenómeno	phénomène	fenomeno	fenômeno	An observable fact or occurrence or a kind of observable fact or occurrence	LST
N 7942	phenotype	fenotipo	phénotype	fenotipo	fenótipo	(genetics) The appearance of an organism based on a multifactorial combination of genetic traits and environmental factors (...)	Lspé

A 1800	phenotypic	fenotípico	phénotypique	fenotipico	fenotípico	Of, or relating to a phenotype	Lspé
N 2457	phosphatase	* fosfatasa	phosphatase	fosfatasi	fosfatase / fosfátase	(biochemistry) Any of several enzymes that hydrolyze phosphate esters (...)	Lspé
N 2932	phosphate	fosfato	phosphate	fosfato	fosfato	(chemistry) Any salt or ester of phosphoric acid	Lspé
N 5805	phosphorylation	* fosforilación	phosphorylation	fosforilazione	fosforilação	(biochemistry) The process of transferring a phosphate group from a donor to an acceptor ; often caused by enzymes	Lspé
A 5554	phylogenetic	filogenético	phylogénétique	filetico	filogenético	Of or pertaining to phylogeny ; phylogenetic	Lspé
N 1864	phylogeny	filogenia	phylogénie	filogenia / filogenesi	filogenia	(systematics) The evolutionary history of groups of organisms, such as species or clades	Lspé
A 8905	physical	físico	physique	físico	físico	Having to do with the body	Lspé
N 9452	physician	médico	médecin	medico	médico	A medical doctor trained in human medicine	Lspé
A 3631	physiological	fisiológico	physiologique	fisiologico	fisiológico	Of, or relating to physiology	Lspé
N 1629	physiology	fisiología	physiologie	fisiologia	fisiologia	A branch of biology that deals with the functions and activities of life or of living matter (as organs, tissues, or cells) and of the physical and chemical phenomena involved	Lspé
N 1880	PI	PI	PI	PI	PI	Abbreviation for phosphatidylinositol, a phospholipid component of eukaryotic cell membranes, also abbreviated in PtdIns	Lspé
N 1958	pig	cerdo	porc	porco / maiale	porco	Any of several mammalian species of the genus Sus (...)	Lspé
N 1281	place	lugar / sitio	lieu / endroit / place	luogo / posto	lugar / sítio	1) place : A location or position	LST
N 1281	place	posición	place / position	posizione / posto	posição	2) place : The ordinal relation; the position in the order of proceeding	LST
V 4127	place	colocar	placer	collocare	colocar	To put in a specific location	LST
N 3420	placebo	placebo	placebo	placebo	placebo	(medicine) A dummy medicine containing no active ingredients ; an inert treatment	Lspé
N 12263	plant	planta	plante	pianta	planta	(botany) An organism of the kingdom Plantae (...)	Lspé
N 1646	plaque	placa	plaque	placca	placa	(medicine) Atheroma, an accumulation in artery walls made up of macrophage cells and debris containing lipids, (cholesterol and fatty acids), calcium, and connective tissue	Lspé
N 9782	plasma	plasma	plasma	plasma	plasma	(hematology) A clear component of blood or lymph (...)	Lspé
N 6552	plasmid	* plásmido	plasmide	plasmide / plasmidio	plasmídeo / plasmídio	(cytology) A loop of double-stranded DNA (...)	Lspé
N 2800	Plasmodium falciparum	Plasmodium falciparum	Plasmodium falciparum	Plasmodium falciparum	Plasmodium falciparum	The parasitic protozoan that causes malaria	Lspé
N 5410	plate	placa	plaque / lame	piastra	placa	(chemistry) Any flat piece of material like coated glass or plastic	Lspé
N 3282	platelet	plaqueta	plaquelette	piastrina	plaqueta	(hematology, cytology) A small colorless disk-shaped particle found in the blood of	Lspé

						mammals, which plays an important role in the formation of blood clots	
N 1646	platform	plataforma	plateforme	piattaforma	plataforma	(computing) A particular type of operating system or environment such as a database or other specific software, and/or a particular type of computer or microprocessor (...)	LST
V 7025	play	jugar	jouer	giocare	jugar	To act as the indicated role, especially in a performance	LST
N 3194	plot	diagrama / gráfica	diagramme / graphique	diagramma / grafico	diagrama / gráfico	A graph or diagram drawn by hand or produced by a mechanical or electronic device	LST
N 3448	pneumonia	pulmonía	pneumonie	polmonite	pneumonia	An acute or chronic inflammation of the lungs caused by viruses, bacteria or other microorganisms, or sometimes by physical or chemical irritants	Lspé
N 6978	point	momento	moment	momento	momento	1) point : A particular moment in an event or occurrence (...)	LST
N 6978	point	punto	point	punto	ponto	2) point : A specific location or place, seen as spatial position	LST
V 1765	point	apuntar / señalar	signaler	segnare	apontar	To draw attention to something or indicate a direction	LST
N 5673	policy	política	politique	politica	política	A principle of behavior, conduct etc. thought to be desirable or necessary, especially as formally expressed by a government or other authoritative body	LST
N 5099	polymerase	polimerasa	polymérase	polimerasi	polimerase	(biochemistry) Any of various enzymes that catalyze the formation of polymers of DNA or RNA using an existing strand of RNA or DNA respectively as a template	Lspé
N 7340	polymorphis m	polimorfismo	polymorphisme	polimorfismo	polimorfismo	(genetics) The regular existence of two or more different genotypes within a given species or population ; also, variability of amino acid sequences within a gene's protein	Lspé
N 2365	pool	* pool	pool / masse commune	pool	pool	A supply of resources	LST
A 2598	poor	escaso	faible	scarso	fraco	1) poor : Inadequate, insufficient	LST
A 2598	poor	escaso	faible	scarso	fraco	2) poor : With little or no possessions or money	LST
N 14760	population	población	population	popolazione	população	1) population : (statistics) A group of units (persons, objects or other items) (...)	LST
N 14760	population	población	population	popolazione	população	2) population : The people living within a political or geographical boundary	LST
N 2236	portion	porción	portion	porzione	porção	That which is divided off or separated, as a part from a whole; a separated part of anything	LST
N 10639	position	posición	position	posizione	posição	A place or location	LST
N 1844	positive	positivo	positif	positivo	positivo	A positive result of a test	LST
A 5921	positive	positivo	positif	positivo	positivo	1) positive : Characterized by the presence of features which support a hypothesis	LST
A 5921	positive	positivo	positif	positivo	positivo	2) positive : Characterized by constructiveness or influence for the better	LST
A 5921	positive	positivo	positif	positivo	positivo	3) positive : (chemistry) electropositive	Lspé
V 1836	possess	poseer / tener	posséder / avoir	possedere / avere	possuir / ter	To have (...)	LST
N 5022	possibility	posibilidad	possibilité	possibilità	possibilidade	The quality of being possible	LST
A 17953	possible	posible	possible	possibile	possível	Apparently valid, likely, plausible.	LST



A 999	posterior	posterior	postérieur	posteriore	posterior	1) posterior : Following in order or time	LST
A 999	posterior	posterior	postérieur	posteriore	posterior	2) posterior : (anatomy) Located towards the rear	Lspé
A 1698	postoperative	postoperatorio / posoperatorio	postopératoire	postoperatorio	pós-operatório	Of or pertaining to, or occurring in the period after a surgical operation	Lspé
A 2190	potent	potente	puissant	potente	potente	Being effective, causing body effect	Lspé
N 5173	potential	potencial	potentiel	potenziale	potencial	Currently unrealized ability	LST
A 11097	potential	potencial	potentiel	potenziale	potencial	Existing in possibility	LST
N 5873	power	potencia	puissance	potenza	potência	(statistics) The probability that a statistical test will reject the null hypothesis when the alternative hypothesis is true	LST
A 1916	practical	práctico	pratique	pratico	prático	Based on practice or action	LST
N 5867	practice	práctica	pratique	pratica	prática	1) practice : The ongoing pursuit of a craft or profession, particularly in medicine or the fine arts	Lspé
N 5867	practice	práctica	pratique	pratica	prática	2) practice : A customary action, habit or behavior ; a manner or routine	LST
N 2430	practitioner	médico	praticien	medico	médico	A person who practices a profession or art, especially law or medicine	Lspé
N 2367	precursor	precursor	précurseur	precursore	precursor	(chemistry) One of the compounds that participates in the chemical reaction that produces another compound	Lspé
V 12971	predict	predecir	prédire	predire	predizer	To (...) estimate a future event on the basis of knowledge and reasoning (...)	LST
N 8389	prediction	predicción	prédiction	predizione	predição	A probability estimation based on statistical methods	LST
A 2412	predictive	predictivo	prédicatif	predittivo	preditivo	(medicine) Expressing the expected accuracy of a statistical measure or of a diagnostic test	Lspé
N 3257	predictor	predictor	* prédicteur	* predittore	predictor	An independent variable.	LST
N 2574	preference	preferencia	préférence	preferenza	preferência	The selection of one thing or person over others	LST
N 6176	pregnancy	gravidez / embarazo	gravidité / grossesse	gravidanza	gravidez	The condition of being pregnant	Lspé
A 1830	pregnant	embarazada / encinta	enceinte	incinta	grávida	Carrying developing offspring within the body	Lspé
A 2426	preliminary	preliminar	préliminaire	preliminare	preliminar	In preparation for the main matter	LST
N 3064	preparation	preparación	préparation	preparazione	preparação	1) preparation : The act of preparing or getting ready	LST
N 3064	preparation	preparación	préparation	preparazione	preparação	2) preparation : That which is prepared	Lspé
V 2512	prepare	preparar	préparer	preparare	preparar	1) prepare : To produce or make by combining elements; to synthesize, compound	Lspé
V 2512	prepare	preparar	préparer	preparare	preparar	2) prepare : To make ready for a specific future purpose; to set up; to assemble	LST
V 1931	prescribe	prescribir / recetar	prescrire / ordonner	prescrivere / ordinare	prescrever / receitar	To order (a drug or medical device) for use by a particular patient	Lspé

N 1726	prescription	prescripción / receta	prescription / ordonnance	prescrizione / ricetta	prescrição / receita	(medicine) A written order, as by a physician or nurse practitioner, for the administration of a medicine or other intervention	Lspé
N 16282	presence	presencia	présence	presenza	presença	The fact or condition of being present (...)	LST
V 6717	present	presentar	présenter	presentare	apresentar	1) present : To put forward in order for it to be seen ; to show, exhibit	LST
V 6717	present	presentar	présenter	presentare	apresentar	2) present : To come to the attention of medical staff, especially with a specific symptom	Lspé
A 20133	present	presente	présent	presente	presente	Relating to something a person is referring to in the very context, with a deictic use similar to the demonstrative adjective this	LST
N 1536	presentation	presentación	présentation	presentazione	apresentação	1) presentation : (medicine) The symptoms and other possible indications of disease, trauma, etc., that are exhibited by a patient (...)	Lspé
N 1536	presentation	presentación	présentation	presentazione	apresentação	2) presentation : A lecture or speech given in front of an audience	LST
N 4206	press	prensa	presse	stampa	imprensa	A collective term for the print-based media (both the people and the newspapers)	LST
N 9497	pressure	presión	pression	pressione	pressão	A pressing : a force applied to a surface	Lspé
N 9012	prevalence	prevalencia	prévalence	prevalenza	prevalência	(epidemiology) The total number of cases of a disease in a given statistical population at a given time divided by the number of individuals in that population	Lspé
V 6667	prevent	prevenir / impedir	prévenir / empêcher	prevenire / impedire	prevenir / impedir	To stop; to keep (from happening)	Lspé
N 5407	prevention	prevención	prévention	prevenzione	prevenção	(medicine) Any measure such as information campaigns, vaccination, early diagnosis etc., intended to limit health-related risks	Lspé
A 12017	previous	previo	préalable	previo	prévio	Occurring before something else	LST
A 8779	primary	primario	primaire	primario	primário	1) primary : (medicine) Relating to day-to-day care provided by health professionals such as nurses, general practitioners, dentists etc.	Lspé
A 8779	primary	primario	primaire	primario	primário	2) primary : (medicine) Relating to the place where a disorder or disease started to occur	Lspé
N 1604	primate	primate	primate	primate	primata	(zoology) A mammal of the order Primates, including simians and prosimians	Lspé
A 2244	prime	primero iniciador /	premier	primo	primeiro	First in time, order, or sequence	LST
N 10000	primer	partidor / cebador	amorce	innesco	iniciador	(biology) A single-stranded nucleic acid molecule required for the initiation of replication of a DNA molecule	Lspé
A 1561	principal	principal	principal	principale	principal	Primary, most important	LST
N 2805	principle	principio	principe	principio	princípio	(physics) A rule or law of nature, or the basic idea on how the laws of nature are applied	Lspé
A 1931	prior	previo	préalable	previo	prévio	Occurring before something else	LST
N 1656	priority	prioridad	priorité	priorità	prioridade	An item's relative importance	LST
A 849	private	privado	privé	privato	privado	1) private : Not in governmental office or employment	LST
A 849	private	privado	privé	privato	privado	2) private : Belonging to, concerning, or accessible only to an individual person or specific group	LST
N 6974	probability	probabilidad	probabilité	probabilità	probabilidade	The relative likelihood of an event happening	LST

N	10590	probe	sonda	sonde	sonda	sonda	(biochemistry) Any group of atoms or molecules radioactively labeled in order to study a given molecule or other structure	Lspé
N	12571	problem	problema	problème	problema	problema	A difficulty that has to be resolved or dealt with	LST
N	12234	procedure	procedimiento	procédure	procedura / procedimento	procedimento	A particular method for performing a task	LST
N	8697	process	proceso	processus	processo	processo	1) process : A series of events which produce a result (...)	LST
N	8697	process	proceso	processus	processo	processo	2) process : (anatomy) Successive physiological responses to keep or restore health	Lspé
V	3103	process	procesar / tratar	traiter	processare / trattare	processar / tratar	To retrieve, store, classify, manipulate, transmit etc. (data, signals, etc.), especially using computer techniques	LST
N	3649	processing	procesamiento o / tratamiento	traitement	processamento / trattamento	processamento o / tratamento	The act of retrieving, storing, classifying, manipulating, transmitting etc. data, especially via computer techniques	LST
V	15391	produce	producir	produire	produrre	produzir	To yield, make or manufacture; to generate	LST
N	6452	product	producto	produit	prodotto	produto	1) product : (chemistry) A chemical substance formed as a result of a chemical reaction	Lspé
N	6452	product	producto	produit	prodotto	produto	2) product : A commodity offered for sale	LST
N	12989	production	producción	production	produzione	produção	The act of producing, making or creating something	LST
N	1720	professional	profesional	professionnel	professionista	profissional	A person who belongs to a profession	Lspé
A	1768	professional	profesional	professionnel	professionale	profissional	Of or pertaining to a profession	Lspé
N	9767	profile	perfil	profile	profilo	perfil	The totality of related characteristics (...)	LST
N	2402	progesterone	* progesterona	progestérone	progesterone	progesterona	(biochemistry, steroids) A steroid hormone, secreted by the ovaries, whose function is to prepare the uterus for the implantation of a fertilized ovum and to maintain pregnancy	Lspé
N	2213	prognosis	pronóstico	pronostic	prognosi	prognóstico / pronóstico	(medicine) A forecast of the future course of a disease or disorder, based on medical knowledge	Lspé
A	2471	prognostic	pronóstico	pronostique	prognostico	pronóstico	Of, pertaining to or characterized by prognosis or prediction	Lspé
N	8332	program	programa	programme	programma	programa	1) program : A set of structured activities	LST
N	8332	program	programa	programme	programma	programa	2) program : A software application (...)	LST
N	1601	progress	progreso	progrès	progresso	progresso	Advancement to a higher or more developed state; development, growth	LST
N	4700	progression	progresión	progression	progressione	progressão	Development, increase, evolution.	LST
A	1778	progressive	progresivo	progressif	progressivo	progressivo	Gradually advancing in extent	LST
N	6594	project	proyecto	projet	progetto	projeto	A planned endeavor, usually with a specific goal and accomplished in several steps or stages	LST
N	7264	proliferation	proliferación	prolifération	proliferazione	proliferação	The process by which an organism produces others of its kind (...)	Lspé
A	1646	prolonged	prolongado	prolongé	prolungato	prolongado	Lengthy in duration ; extended	LST
V	2468	promote	promover	promouvoir	promuovere	promover	1) promote : To encourage, urge or incite	LST

V	2468	promote	promover	promouvoir	promuovere	promover	2) promote (chemistry) : To increase the activity of a catalyst by changing its surface structure	Lspé
N	10933	promoter	promotor	promoteur	promotore	promotor	(chemistry) A material that accelerates the action of a catalyst	Lspé
N	7292	property	propiedad	propriété	proprietà	propriedade	An attribute or abstract quality which is characteristic of a class of objects	LST
N	7567	proportion	proporción	proportion	proporzione	proporção	A quantity of something that is part of the whole amount or number	LST
V	6869	propose	proponer	proposer	proporre	propor	To suggest a plan, course of action, etc	LST
A	3287	prospective	prospectivo	prospectif	prospettivo	prospectivo	(medicine, of research) That has a sufficient population size to produce reliable data over a sufficient period	Lspé
N	1751	prostaglandin	prostaglandina	prostaglandine	prostaglandina	prostaglandina	(biochemistry) Any of a group of naturally occurring lipids derived from the C20 acid prostanic acid (...)	Lspé
N	4953	prostate	próstata	prostate	prostata	próstata	The prostate gland	Lspé
N	3728	protease	proteasa	protéase	proteasi	protease	(biochemistry) An enzyme that cuts or cleaves proteins	Lspé
V	2441	protect	proteger	protéger	proteggere	proteger	To keep safe; to defend; to guard; to prevent harm coming to	Lspé
N	3061	protection	protección	protection	protezione	proteção	The process of keeping safe	Lspé
A	2312	protective	protector	protecteur	protettivo	protetivo / protetor	Serving, intended or wishing to protect	Lspé
N	103478	protein	proteína	protéine	proteina	proteína	Any of a numerous large complex naturally-produced molecules composed of one or more long chains of amino acids (...)	Lspé
N	7531	protocol	protocolo	protocole	protocollo	protocolo	(sciences) The original notes of observations made during an experiment : also, the precise method for carrying out or reproducing a given experiment	LST
V	2751	prove	probar	prouver	provare	provar	To demonstrate that something is true or viable; to give proof for	LST
V	28674	provide	proveer / fornecer	pourvoir / fournir	provvedere / fornire	prover / fornecer	To give what is needed or desired (...)	LST
N	2423	provider	proveedor	fournisseur	fornitore	fornecedor	One who (...) provides a service, commodity, or the means for subsistence	LST
A	1660	proximate	próximo	proche	prossimo	próximo	Close or closest ; adjacent	LST
A	2256	psychiatric	psiquiátrico	psychiatrique	psichiatrico	psiquiátrico	Of or relating to psychiatry	Lspé
N	2878	psychiatry	psiquiatria	psychiatrie	psichiatria	psiquiatria	(medicine) The branch of medicine that subjectively diagnoses, treats, and studies mental illness and behavioral conditions	Lspé
A	2014	psychological	psicológico	psychologique	psicologico	psicológico	Of or pertaining to psychology	Lspé
A	3961	public	público	public	pubblico	público	1) public : Able to be seen or known by everyone	LST
A	3961	public	público	public	pubblico	público	2) public : concerning the whole country, community, etc.	LST
N	1788	publication	publicación	publication	pubblicazione	publicação	1) publication : An issue of printed or other matter, offered for sale or distribution	LST

N 1788	publication	publicación	publication	pubblicazione	publicação	2) publication : The act of publishing printed or other matter	LST
V 6742	publish	publicar	publier	pubblicare	publicar	To write in a publication (usually as an academic)	LST
A 6756	pulmonary	pulmonar	pulmonaire	polmonare	pulmonar	Related to lungs	Lspé
V 1815	purchase	comprar	acheter	comprare	comprar	To buy, obtain by payment of a price in money or its equivalent	LST
N 2455	purification	purificación	purification	purificazione	purificação	The act or process of purifying ; the removal of impurities	Lspé
V 4608	purify	purificar	purifier	purificare	purificar	To cleanse (something), or rid (it) of impurities	Lspé
N 3803	purpose	propósito	but	proposito	propósito	An object to be reached ; a target ; an aim ; a goal	LST
A 4413	putative	supuesto	présumé	presunto	suposto	Accepted by supposition rather than as a result of proof	LST
N 1736	pylorus	píloro	pylore	piloro	piloro	(anatomy) The opening in a vertebrate, including humans, at the lower end of the stomach that opens into the duodenum	Lspé
A 2273	qualitative	cualitativo	qualitatif	qualitativo	qualitativo	Based on some quality rather than on some quantity	LST
N 7858	quality	cualidad	qualité	qualità	qualidade	1) quality : The level of excellence	LST
N 7858	quality	cualidad	qualité	qualità	qualidade	2) quality : The degree to which a man-made object or system is free from bugs and flaws (...)	LST
N 1772	quantification	cuantificación	quantification	quantificazione	quantificação	The act of quantifying	LST
V 2714	quantify	cuantificar	quantifier	quantificare	quantificar	To determine the value of (a variable or expression)	LST
A 6102	quantitative	cuantitativo	quantitatif	quantitativo	quantitativo	Of a form of analysis that determines the amount of some element or compound in a sample	LST
N 1825	quantity	cantidad	quantité	quantità	quantidade	A specific measured amount	LST
N 2257	query	búsqueda	requête	richiesta	busca	A set of instructions passed to a database	LST
N 8842	question	cuestión	question	questione	questão	A subject or topic for consideration or investigation	LST
N 8486	questionnaire	cuestionario	questionnaire	questionario	questionário	A form containing a list of questions ; a means of gathering information for a survey	LST
N 3896	rabbit	conejo	lapin	coniglio	coelho	A mammal of the family Leporidae (...)	Lspé
N 2895	radiation	radiación	radiation / rayonnement	radiazione	radiação	Radioactive energy.	Lspé
V 1414	raise	levantar	soulever	sollevare	levantar	1) raise : To bring into being; to produce; to cause to arise (...)	LST
V 1414	raise	aumentar	augmenter	aumentare	aumentar	2) raise : To (...) increase or develop	LST
A 6280	random	aleatorio	aléatoire	aleatorio	randômico	Of or relating to probability distribution	LST
V 8001	randomize	* randomizar	* randomiser	randomizzare / casualizzare	randomizar / acidentalizar	To arrange randomly; to make random	LST
V 4879	range	oscilar / estar comprendido	osciller / être compris	oscillare / essere	oscilar / estar compreendid	(...) To be able to take any of the values in a specified range	LST

N 12691	range	amplitud	amplitude / étendue / intervalle de variation	campo di variazione	amplitude	(statistics) The difference between the largest and smallest observations in the sample	LST
A 4925	rapid	rápido	rapide	rapido	rápido	Very swift or quick	LST
A 2533	rare	raro	rare	raro	raro	Very uncommon, scarce	LST
N 20588	rat	rata	rat	ratto	rato	A medium-sized rodent belonging to the genus Rattus	Lspé
N 32928	rate	tasa	taux	tasso	taxa	The proportional relationship between one amount, value, etc. and another	LST
N 2132	rating	evaluación	évaluation	valutazione	avaliação	An evaluation (...)	LST
N 14488	ratio	ratio / razón	ratio	ragione	razão	A number representing a comparison between two things	LST
A 1564	raw	bruto	brut	bruto	bruto	Uncorrected, without analysis	LST
V 4520	reach	alcanzar / llegar	arriver / atteindre	arrivare / raggiungere	alcançar / atingir / chegar	To extend in dimension, time etc. (...)	LST
N 6340	reaction	reacción	réaction	reazione	reação	1) reaction : An action or statement in response to a stimulus or other event	Lspé
N 6340	reaction	reacción	réaction	reazione	reação	2) reaction : (chemistry) A transformation in which one or more substances is converted into another(...)	Lspé
A 2325	reactive	reactivo	réactif	reattivo	reativo	(chemistry) that readily takes part in reactions	Lspé
N 1783	reactivity	* reactividad	réactivité	reattività	reatividade	(chemistry) relative susceptibility to chemical reaction	Lspé
V 4817	read	leer	lire	leggere	ler	To look at and interpret letters or other information that is written	LST
N 1105	reading	medición	relevé / mesure	misura	medição	1) reading : A value indicated by a measuring device	LST
N 1105	reading	lectura	lecture	lettura	leitura	2) reading : The process of interpreting a symbol, a sign or a measuring device	LST
N 2503	reagent	reactivo	réactif	reagente / reattivo	reagente / reativo	(chemistry) A compound or mixture of compounds used to treat materials, samples, other compounds or reactants in a laboratory or sometimes an industrial setting	Lspé
A 2775	real	real	réel	reale	real	True, genuine, not merely nominal or apparent	LST
A 2108	real-time	en tiempo real	en temps réel	in tempo reale	em tempo real	That responds to events or signals within a predictable time	LST
N 5625	reason	razón	raison	ragione	razão	A cause	LST
V 10090	receive	recibir	recevoir	ricevere	receber	To take, as something that is offered, given, committed, sent, paid, etc.; to accept; to be given something	LST
A 10179	recent	reciente	récent	recente	recente	Having happened a short while ago	LST
N 17995	receptor	receptor	récepteur	recettore / ricettore	receptor	1) receptor : (biochemistry, medicine) A protein on a cell wall that binds with specific molecules (...)	Lspé

N 17995	receptor	receptor	récepteur	recettore / ricettore	receptor	2) receptor : (biology) Any specialized cell or structure that responds to sensory stimuli	Lspé
N 3551	recognition	reconocimien to	reconnaissance	riconosciment o	reconhecimen to	(immunology) The propriety consisting for antibodies to bind to some specific antigens and not to others	Lspé
V 2523	recognize	reconocer	reconnaitre	riconoscere	reconhecer	1) recognize : To acknowledge or consider as something	LST
V 2523	recognize	reconocer	reconnaitre	riconoscere	reconhecer	2) recognize : (immunology) To have the property to bind to specific antigens	Lspé
A 5148	recombinant	recombinante	recombinant	ricombinante	recombinante	Formed by combining existing elements	Lspé
N 3031	recombinatio n	recombinació n	recombinaison	ricombinazion e	recombinação	(genetics) The formation of genetic combinations in offspring that are not present in the parents	Lspé
V 2749	recommend	recomendar	recommander	raccomandare	recomendar	To advise, propose, counsel favorably	LST
N 1366	recommenda tion	recomendació n	recommandatio n	raccomandazio ne	recomendaçã o	1) recommandation : Act of recommending	LST
N 1366	recommenda tion	recomendació n	recommandatio n	raccomandazio ne	recomendaçã o	2) recommandation : That which is recommended	LST
N 4327	record	registro	dossier	cartella	registro / registro	Any instance of a physical medium on which information was put for the purpose of preserving it and making it available for future reference	LST
V 4483	record	registrar	enregistrer	registrare	registrar	To make a record of information	LST
V 2165	recover	recobrase / recuperarse / restablecerse	récupérer / se rétablir	recuperarsi / ricuperarsi / ristabilirsi	recobrar-se / recuperar-se/ restabelecer- se	To get better, regain one's health	Lspé
N 2966	recovery	recuperación / restablecimie nto	récupération / rétablissement	recupero / ricovero / ristabilimento	recuperação / restabelecime nto	A return to normal health	Lspé
V 1286	recruit	reclutar	recruter	reclutare	recrutar	1) recruit : To prompt a protein, leukocyte etc. to intervene in a given region of the body	Lspé
V 1286	recruit	contratar	recruter	assumere	recrutar	2) recruit : To enroll or enlist new members or potential employees (...)	LST
N 2668	recruitment	reclutamiento	recrutement	reclutamento	recrutamento	(biology, ecology) The addition of new recruits to a population	Lspé
N 1715	recurrence	recurrencia	récurrence	ricorrenza	recorrência	A return of symptoms as part of the natural progress of a disease	Lspé
A 1619	recurrent	recurrente	récurrent	ricorrente	recorrente	Recurring time after time	Lspé
A 3364	red	rojo	rouge	rosso	vermelho	Having red as its color	LST
V 15946	reduce	reducir	réduire	ridurre	reduzir	To bring down the size, quantity, quality, value or intensity; to diminish, to lower, to impair	LST

A 3566	reduced	reducido	réduit	ridotto	reduzido	Made smaller or less	LST
N 8871	reduction	reducción	réduction	riduzione	redução	The act, process, or result of reducing	LST
V 4256	refer	referir	référer	riferire	referir	To place in or under by a mental or rational process; to assign to (...)	LST
N 3413	reference	referencia	référence	riferimento	referência	1) reference : A measurement one can compare to	LST
N 3413	reference	referencia	référence	riferimento	referência	2) reference : (academic writing) A short written identification of a previously published work which is used as a source for a text	LST
V 6946	reflect	reflejar	réfléter	riflettere	refletir	To give evidence (...)	LST
V 5524	regard	concernir	regarder / concerner	riguardare / concernere	concernir	To have to do with, to concern	LST
N 1824	regimen	régimen	régime	regime	regime	Any regulation or remedy which is intended to produce beneficial effects by gradual operation	Lspé
N 29622	region	región	région	regione	região	A place in or a part of the body (...)	Lspé
A 2798	regional	regional	régional	regionale	regional	Of, or pertaining to a specific region or district	LST
N 3718	regression	regresión	régression	regressione	regressão	1) regression : (statistics) An equation using specified and associated data for two or more variables (...)	LST
N 3718	regression	regresión	régression	regressione	regressão	2) regression : The diminishing of a cellular mass like a tumor, or of an organ size	Lspé
A 1776	regular	regular	régulier	regolare	regular	Having a constant pattern	LST
V 12320	regulate	regular	réguler / régler	regolare	regular	To adjust to a particular specification or requirement	LST
N 6625	regulation	regulación	régulation	regolazione	regulação	1) regulation : (genetics) Mechanism controlling DNA transcription	Lspé
N 6625	regulation	regulación	régulation	regolazione	regulação	2) regulation : (medicine) Physiological process which consists in maintaining homeostasis	Lspé
N 2778	regulator	regulador	régulateur	regolatore	regulador	(genetics) A gene involved in controlling the expression of one or more other genes	Lspé
A 6893	regulatory	regulador	* régulateur	regolatore	regulador	Of or pertaining to regulation	Lspé
N 1573	rehabilitation	rehabilitación	rééducation / réadaptation	riabilitazione	reabilitação	The process of rehabilitating	Lspé
V 17564	relate	relacionar	relier / lier	legare / collegare	relacionar	To make a connection or correlation from one thing to another	LST
A 5629	related	relacionado	lié	legato	relacionado	Standing in relation or connection	LST
N 4992	relation	relación	relation	relazione	relação	The manner in which two things may be associated	LST
N 13698	relationship	relación	rapport	rapporto	relação	Connection or association (...)	LST
A 10564	relative	relativo	relatif	relativo	relativo	Connected to or depending on something else	LST
N 6832	release	liberación	libération	liberazione	liberação	(biochemistry) Process by which a chemical substance is set free	Lspé
V 3222	release	liberar	libérer	liberare	liberar	(biochemistry) To set free a chemical substance	Lspé
N 1920	relevance	pertenencia /	pertinence	pertinenzza	pertinência	The property or state of being relevant or pertinent	LST



A 5636	relevant	pertinencia pertinente	pertinent	pertinente	pertinente	Directly related, connected, or pertinent to a topic	LST
N 3368	reliability	fiabilidad	fiabilité	affidabilità	fiabilidade / confiabilidad e	The quality of a measurement indicating the degree to which the measure is consistent (...)	LST
A 2518	reliable	fiable	fiable	affidabile	fiável	Suitable or fit to be relied on, worthy of dependance or reliance	LST
V 1596	rely	apoyarse	s'appuyer	appoggiarsi	apoiar-se	To depend (...)	LST
V 10297	remain	permanecer / quedarse	rester / demeurer	restare / rimanere	permanecer / ficar	To continue unchanged in place, form, or condition, or undiminished in quantity; to abide; to stay; to endure; to last	LST
A 3071	remaining	restante	restant	restante	restante	Which remains	LST
N 2230	removal	remoción	prélèvement	estirpazione / asportazione	remoção	The process of moving, or the fact of being removed.	Lspé
V 6062	remove	remover / retirar / extirpar	enlever / retirer / extirper	rimuovere / asportare / estirpare	remover / retirar / extirpar	To move something from one place to another, especially to take away	Lspé
A 5401	renal	renal	rénal	renale	renal	Pertaining to the kidneys	Lspé
N 2666	repair	reparación	réparation	riparazione	reparação	The act of repairing something	Lspé
N 5535	repeat	repetición	répétition	ripetizione	repetição	Patterns of nucleic acids that occur in multiple copies throughout the genome	Lspé
V 3017	repeat	repetir	répéter	ripetere	repetir	To do or say again (and again)	LST
A 1599	repeated	repetido	répété	ripetuto	repetido	Having been done again	LST
V 2256	replace	reemplazar	remplacer	rimpiazzare	substituir	To supply or substitute an equivalent for	LST
N 2357	replacement	reemplazo	remplacement	rimpiazzo / ricambio	substituição	The act of replacing something	Lspé
V 2968	replicate	replicar	répliquer	replicare	replicar	To make a copy (replica) of	Lspé
N 5222	replication	replicación	réplication	replicazione	replicação	(biology) The process of producing replicas of DNA or RNA molecules	Lspé
N 13710	report	informe	rapport	rapporto / relazione	relatório	A piece of information describing (...)	LST
V 31887	report	reportar / informar	rapporter / signaler	riportare / segnalare	relatar	To make a formal statement (...)	LST
N 2427	reporter	reportero	rapporteur	reporter	repórter	(biology) A gene attached by a researcher to a regulatory sequence of another gene of interest (...)	Lspé
V 7171	represent	representar	représenter	rappresentare	representar	1) represent : To exhibit to another mind in language; to show (...)	LST
V 7171	represent	representar	représenter	rappresentare	representar	2) represent : To constitute, to make up, to be a example of	LST

N 2776	representation	representación	représentation	rappresentazione	representação	A figure, image or idea that substitutes reality	LST
N 1756	representative	representante	représentant	rappresentante	representante	One that is taken as typical of its class	LST
A 1767	representative	representativo	représentatif	rappresentativo	representativo	Typical : having the same properties or interest as a larger group	LST
A 2537	reproductive	reproductivo	reproductif	riproduttivo	reprodutivo	Of or relating to reproduction	Lspé
V 19527	require	requerir / necesitar	requérir / nécessiter	richiedere / necessitare	requerer / necessitar	Naturally to demand as indispensable; to need, to call for as necessary	LST
N 3407	requirement	requisito	exigence	requisito	requisito	A necessity or prerequisite ; something required or obligatory	LST
N 26542	research	investigación	recherche	ricerca	pesquisa	Diligent inquiry or examination to seek or revise facts, principles, theories, applications, etc. (...)	LST
N 3343	researcher	investigador	chercheur	ricercatore	pesquisador	One who researches	LST
N 2325	resident	residente	résident	residente	residente	Person, animal or plant living at a location or in an area	Lspé
A 1563	residual	residual	résiduel	residuale	residual	Of, relating to, or remaining as a residue	Lspé
N 8180	residue	residuo	résidu	residuo	resíduo	(chemistry) The substance that remains after evaporation, distillation, filtration or any similar process.	Lspé
N 9142	resistance	resistencia	résistance	resistenza	resistência	The act of resisting or the capacity to resist	Lspé
A 3840	resistant	resistente	résistant	resistente	resistente	That is not affected or overcome by a disease, drug, chemical agent etc.	Lspé
N 1231	resolution	resolución	résolution	risoluzione	resolução	1) resolution : (computing) The degree of fineness with which an image can be recorded or produced	LST
N 1231	resolution	resolución	résolution	risoluzione	resolução	2) resolution : (medicine) In a pathological process, the phase during which pathogens and damaged tissues are removed by macrophages	Lspé
V 1948	resolve	resolver	résoudre	risolvere	resolver	To find a solution to	LST
N 6264	resource	recurso	ressource	risorse	recurso	Something that one uses to achieve an objective, e.g. raw materials or personnel	LST
A 2066	respective	respectivo	respectif	rispettivo	respectivo	Relating to particular persons or things, each to each	LST
A 7242	respiratory	respiratorio	respiratoire	respiratorio	respiratório	Relating to respiration, breathing	Lspé
V 3103	respond	responder	répondre	rispondere	responder	To act in return; to exhibit some action or effect in return to a force or stimulus (...)	Lspé
N 4108	respondent	encuestado	répondant	rispondente	inquirido	Person that participates in research involving questionnaires	LST
N 36862	response	respuesta	réponse	risposta	resposta	A reaction to a stimulus or provocation	Lspé
A 4915	responsible	responsable	responsable	responsabile	responsável	Being a primary cause or agent of some event or action	LST
A 1899	responsive	sensible	sensible	sensibile	sensível	Able to receive external stimuli	Lspé
N 1630	responsiveness	sensibilidad	sensibilité	sensibilità	sensibilidade	The quality or state of being responsive	Lspé

V 2505	restrict	restringir	restreindre	restringere	restringir	To restrain within bounds; to limit; to confine (...)	LST
N 1854	restriction	restricción	restriction	restrizione	restrição	1) restriction : The act of restricting, or the state of being restricted	LST
N 1854	restriction	restricción	restriction	restrizione	restrição	2) restriction : (biology) The mechanism by which a cell degrades foreign DNA material	Lspé
N 55061	result	resultado	résultat	risultato	resultado	That which results ; the conclusion or end to which any course or condition of things leads, or which is obtained by any process or operation (...)	LST
V 13731	result	resultar	résulter	risultare	resultar	To proceed, spring or rise, as a consequence, from facts, arguments, premises, combination of circumstances, consultation, thought or endeavor	LST
A 2269	resulting	resultante	résultant	risultante	resultante	Of something that follows as the result of something else	LST
V 1940	retain	conservar	conserver / garder	conservare	conservar	To keep in possession or use	LST
V 2068	return	volver	revenir	tornare	voltar	To come or go back	LST
V 11612	reveal	revelar	révéler	rivelare	revelar	To uncover; to show and display that which was hidden	LST
V 866	reverse	invertir	inverser	invertire	invertir	1) reverse : To change totally; to alter to the opposite	LST
V 866	reverse	revertir	renverser	invertire	reverter	2) reverse : (chemistry) To change the direction of a reaction such that the products become the reactants and vice versa	Lspé
A 2683	reverse	reverso	inverse	inverso	reverso	(genetics) In which cDNA synthetization is obtained from an RNA template	Lspé
N 6248	review	estudio	étude	studio	estudo	1) review : A survey (...)	LST
N 6248	review	revista	revue	rivista	revista	2) review : A periodical	LST
V 4426	review	examinar	examiner	esaminare	examinar	To survey	LST
A 4868	rheumatoid	reumatoide	rhumatoïde	reumatoide	reumatoide	Presenting analogies with rheumatism	Lspé
N 1668	rhythm	ritmo	rythme	ritmo	ritmo	A regular quantitative change in a variable (notably natural) process	LST
A 1800	ribosomal	ribosomal / ribosómico	ribosomique	ribosomiale	ribossomal / ribossômico	Of or relating to ribosomes	Lspé
A 2967	rich	rico	riche	ricco	rico	Plentiful, abounding	LST
N 1824	right	derecho	droit	diritto	direito	A legal or moral entitlement	LST
A 3380	right	derecho	droit	destro	direito	Designating the right side of the body	Lspé
N 1817	rise	aumento	augmentation	aumento	aumento	The process of (...) moving upwards or becoming greater	LST
N 28414	risk	riesgo	risque	rischio	risco	The likelihood of a negative outcome	Lspé
N 2083	RNA	ARN	ARN	RNA	ARN	(genetics) Ribonucleic acid.	Lspé
A 1774	robust	robusto	robuste	robusto	robusto	Not greatly influenced by errors in assumptions about the distribution of sample errors	LST
N 26957	role	papel	rôle	ruolo	papel	The function or position of something	LST
N 3794	room	habitación / cuarto	pièce	stanza	quarto	A separate part of a building, enclosed by walls, a floor and a ceiling	LST
N 1025	root	raíz	racine	radice	raiz	1) root : (mathematics) Of a number or expression, a number which, when raised to a	LST

N 1025	root	raíz	racine	radice	raiz	specified power, yields the specified number or expression 2) root : The lowest place, position, or part	Lspé
N 1599	round	etapa	étape	tappa	etapa	A series of duties or tasks which must be performed in turn, and then repeated	LST
A 2223	routine	de rutina	de routine	di routine	de rotina	Regular, habitual	Lspé
N 2307	rule	regla	règle	regola	regra	A determinate method prescribed for performing any operation and producing a certain result	LST
V 3302	run	tratar / procesar	traiter	trattare / processare	tratar / processar	(transitive) To execute or carry out a plan, procedure or program	LST
A 2614	rural	rural	rural	rurale	rural	Pertaining to the less-populated, non-urban areas	LST
N 2446	Saccharomyces cerevisiae	Saccharomyces cerevisiae	Saccharomyces cerevisiae	Saccharomyces cerevisiae	Saccharomyces cerevisiae	A taxonomic species within the family Saccharomycetaceae — a yeast (...)	Lspé
N 2622	safety	inocuidad	innocuité	innocuità	inocuidade	The condition or feeling of being safe; security; certainty	Lspé
A 21800	same	mismo	même	stesso	mesmo	Not different as regards self ; identical	LST
N 36397	sample	muestra	échantillon	campione	amostra	A part of anything taken or presented for inspection, or shown as evidence of the quality of the whole (...)	Lspé
V 1940	sample	muestrear	échantillonner	campionare	amostrar	To take or to test a sample or samples (...)	Lspé
N 2438	sampling	muestreo	échantillonnage	campionamento	amostragem	The process or technique of obtaining a representative sample	Lspé
N 2273	satisfaction	satisfacción	satisfaction	soddisfazione	satisfação	A fulfillment of a need or desire	LST
V 1787	say	decir	dire	dire	dizer	To communicate, either verbally or in writing	LST
N 11080	scale	escala	échelle	scala	escala	An ordered numerical sequence used for measurement	LST
V 3354	scan	escanear	scanner / numériser	scannerizzare	escanear	(computing) To create a digital copy of an image using a scanner	LST
N 1809	scenario	escenario	scenario	scenario	cenário	An outline or model of an expected or supposed sequence of events	LST
N 1625	scheme	esquema	schéma	schema	esquema	A chart or diagram of a system or object	LST
N 2346	schizophrenia	esquizofrenia	schizophrénie	schizofrenia	esquizofrenia	(pathology) A psychiatric diagnosis denoting a persistent, often chronic, mental illness variously affecting behavior, thinking, and emotion	Lspé
N 2111	school	facultad	faculté	facoltà	faculdade	1) school : Within a larger educational institution, an organizational unit, such as a department or institute, which is dedicated to a specific subject area	LST
N 2111	school	escuela	école	scuola	escola	2) school : An institution dedicated to teaching and learning ; an educational institution	LST
N 11101	science	ciencia	science	scienza	ciência	The collective discipline of study or learning acquired through the scientific method ; the sum of knowledge gained from such methods and discipline	LST
A 3128	scientific	científico	scientifique	scientifico	científico	Derived from or consistent with the scientific method	LST

N 20011	score	puntuación	pointage	puntaggio	pontuação	The performance of an individual or group on an examination or text (...)	LST
V 3301	score	registrar	enregistrer	registrare	registrar	To achieve (a score) in e.g. a test	LST
N 2100	screen	criba	crible	setaccio	rastreio	(genetics) A technique used to identify genes so as to study gene functions	Lspé
V 1535	screen	cribar	cribler	setacciare	rastrear	1) screen :To search chemical libraries by means of a computational technique in order to identify chemical compounds which would potentially bind to a given biological target such as a protein	Lspé
V 1535	screen	cribar	dépister	schermare	despistar / rastrear despistagem / rastreio / rastreamento	2) screen : To examine patients or treat a sample in order to detect a chemical or a disease, or to assess susceptibility to a disease	Lspé
N 5562	screening	cribado	dépistage	screening		(medicine) A method to identify a disease in a population which is not showing any symptoms of this disease	Lspé
N 8615	search	búsqueda	recherche	ricerca	busca	A attempt to find something	LST
V 3250	search	buscar	chercher	cercare	buscar	To look in (a place) for something	LST
A 2989	second	segundo	second	secondo	segundo	Number-two ; following after the first one with nothing between them (...)	LST
N 12034	second	segundo	seconde	secondo	segundo	The SI unit of time (...)	LST
A 2387	secondary	secundario	secondaire	secondario	secundário	1) secondary : Succeeding next in order to the first; of second place, origin, rank, rank, etc.; not primary	LST
A 2387	secondary	secundario	secondaire	secondario	secundário	2) secondary : Of less than primary importance	LST
A 2387	secondary	secundario	secondaire	secondario	secundário	3) secondary : (medicine) Dependent or consequent upon another disease; or occurring in the second stage of a disease	Lspé
V 2552	secrete	secretar	sécréter	secernere	secretar	(physiology, of organs) To extract a substance from blood (...) for fulfilling of a physiological function	Lspé
N 2529	secretion	secreción	sécrétion	secrezione	secreção	1) secretion : Any substance that is secreted by an organism	Lspé
N 2529	secretion	secreción	sécrétion	secrezione	secreção	2) secretion : The act of secreting a substance, especially from a gland	Lspé
N 2976	section	sección	section	sezione	seção	1) section : (sciences) A thin slice of material prepared as a specimen for research	Lspé
N 2976	section	sección	section	sezione	seção	2) section : (surgery) An incision or the act of making an incision	Lspé
N 2976	section	sección	section	sezione	seção	3) section : A part of a document	LST
N 1953	sector	sector	secteur	settore	setor	A field of economic activity	LST
V 11905	see	ver	voir	vedere	ver	1) see : To perceive or detect with the eyes, or as if by sight	LST
V 11905	see	ver	voir	vedere	ver	2) see : To determine by trial or experiment ; to find out	LST
N 1765	seed	semilla	graine	seme	semente	A fertilized grain, initially encased in a fruit, which may grow into a mature plant	Lspé
V 2702	seek	procurar	chercher	cercare	procurar	To try to reach or come to; to go to; to resort to	LST
V 7145	seem	parecer	sembler / paraitre	sembrare / parere	sembrar / parecer	To appear; to look outwardly; to be perceived as	LST
N 4470	segment	segmento	segment	segmento	segmento	(sciences) A portion	LST

V 9782	select	seleccionar	sélectionner	selezionare	seleccionar	To choose one or more elements of a set, especially a set of options	LST
A 2233	selected	seleccionado	sélectionné	selezionato	seleccionado	That has been selected or chosen	LST
N 4371	selection	selección	sélection	selezione	seleção	1) selection : The process or act of selecting	LST
N 4371	selection	selección	sélection	selezione	seleção	2) selection : Something selected	LST
A 3729	selective	selectivo	sélectif	selettivo	seletivo	Of or pertaining to the process of selection	LST
N 1274	sense	direccionad ad	sens / polarité	senso	sentido	1) sense : (biochemistry) referring to the strand of a nucleic acid that directly specifies the product	Lspé
N 1274	sense	sentido	sens	senso	sentido	2) sense : The meaning, reason, or value of something	LST
A 5238	sensitive	sensitivo	sensitif	sensitivo	sensitivo	Responsive to stimuli	Lspé
N 8391	sensitivity	sensibilidad	sensibilité	sensibilità	sensibilidade	The ability of an organism or organ to respond to external stimuli	Lspé
A 1565	sensory	sensorial / sensorio	sensoriel	sensoriale	sensorial	Of the senses or sensation	Lspé
V 3123	separate	separar	séparer	separare	separar	To divide (a thing) into separate parts	LST
A 3301	separate	separado	séparé	separato	separado	Apart from the rest ; not connected to or attached to anything else	LST
N 2075	separation	separación	séparation	separazione	separação	The act of disuniting two or more things, or the condition of being separated	LST
N 3210	sepsis	sepsis	sepsis	sepsi	sépsis	(pathology) A serious medical condition in which the whole body is inflamed, and a known or suspected infection is present	Lspé
N 63884	sequence	secuencia	séquence	sequenza	sequência	Order in which amino-acid residues, connected by peptide bonds, lie in the chain of peptides and protein	Lspé
V 3258	sequence	secuenciar	séquencer	sequenziare	sequenciar	To determine the order of things, especially of amino acids in a protein, or of bases in a nucleic acid	Lspé
N 5254	series	serie	série	serie	série	A number of things that follow on one after the other or are connected one after the other	LST
N 2168	serine	* serina	sérine	serina	serina	(biochemistry) A nonessential amino acid, CH <sub>2</sub> OH.CH(NH <sub>2</sub> )COOH, found in most animal proteins, especially silk	Lspé
A 1768	serious	serio	sérieux	serio	sério	Important ; critical	Lspé
N 15432	serum	suero	sérum	siero	soro	The clear yellowish fluid obtained upon separating whole blood into solid and liquid components (...)	Lspé
V 3704	serve	servir	servir	servire	servir	To have a given use or purpose; to function for	LST
N 1635	server	servidor	serveur	server	servidor	(computing) A computer dedicated to running programs	LST
N 4890	service	servicio	service	servizio	serviço	1) service : A department in a company, an organization, a government department, etc.	LST
N 4890	service	servicio	service	servizio	serviço	2) service : (economics) The practice of providing a service as economic activity	LST
N 2076	session	sesión	session	sessione	sessão	A period devoted to a particular activity	LST
N 16007	set	serie	série	serie	série	A series of, a group of	LST
V 11351	set	fijar	fixer	fissare	fixar	To determine or settle	LST
N 4475	setting	contexto	contexte	ambito	contexto	The time, place and circumstance in which something (...) is set	LST

A 7870	severe	grave	grave	grave	grave	Very bad or intense	Lspé
N 4265	severeness	gravedad	gravité	gravità	gravidade	The property of being severe	Lspé
N 6356	sex	sexo	sexe	Sesso	sexo	A main division into which an organism is placed according to its reproductive functions or organs	Lspé
A 1333	sexual	sexual	sexuel	sessuale	sexual	1) sexual : Of or relating to gender	Lspé
A 1333	sexual	sexual	sexuel	sessuale	sexual	2) sexual : Of or relating to (...) sexual acts and sexual reproduction	Lspé
N 1764	shape	forma	forme	forma	forma	The appearance of something, especially its outline	LST
V 4996	share	compartir	partager	condividere	compartilhar	To have or use in common	LST
N 2639	shift	cambio	changement	cambio / cambiamento	mudança	An act of shifting; a slight movement or change	LST
N 3302	shock	choque	choc	shock / choc	choque	(...) a life-threatening medical emergency (...)	Lspé
A 3546	short	corto	court	corto	curto	1) short : Having little duration	LST
A 3546	short	corto	court	corto	curto	2) short : Having a small distance from one end or edge to another, either horizontally or vertically	LST
V 74611	show	mostrar	montrer	mostrare	mostrar	To indicate to be true; to demonstrate	LST
N 4767	side	lado	côté	lato	lado	One half (left or right, top or bottom, front or back, etc.) of something or someone	Lspé
N 2896	sign	signo	signe	segno	sinal	(medicine) A property of the body that indicates a disease and, unlike a symptom, is unlikely to be noticed by the patient	Lspé
N 14957	signal	señal	signal	segnale	sinal	A sign made to give notice of some occurrence, command or danger, or to indicate the start of a concerted action	Lspé
N 11297	signaling	señalización	signalisation	segnalazione	sinalização	(biochemistry) The sending of a biochemical signal	Lspé
N 6853	significance	significación	signification	significatività	significância	The extent to which something matters ; importance	LST
A 30064	significant	significativo	significatif	significativo	significativo	Having noticeable effect	LST
A 22789	similar	similar	similaire	simile	similar / semelhante	Having traits or characteristics in common : alike, comparable	LST
N 6455	similarity	semejanza	similarité	similarità	similaridade / semelhança	Closeness of appearance to something else	LST
A 5609	simple	simple	simple	semplice	simples	Uncomplicated	LST
A 1636	simulated	simulado	simulé	simulato	simulado	Made to imitate something else ; artificial	LST
N 2825	simulation	simulación	simulation	simulazione	simulação	Something which simulates a system or environment in order to predict actual behavior	LST
A 17144	single	solo	seul	solo	único	Not accompanied by anything else, one in number	LST
N 26749	site	sitio	site	sito	sítio	Region of a protein, a piece of DNA or RNA where chemical reactions take place	Lspé
N 3748	situation	situación	situation	situazione	situação	Position or status with regard to conditions and circumstances	LST
N 16181	size	tamaño	taille	taglia	tamanho	The dimensions or magnitude of a thing ; how big something is	LST

A 2057	skeletal	esquelético	squelettique	scheletrico	esquelético	Of or relating to the skeleton	Lspé
N 2114	skill	habilidad	habileté	abilità	habilidade	The capacity to do something well ; technique, ability (...)	LST
N 5457	skin	piel	peau	pelle	pele	The outer protective layer of the body of any animal, including of a human	Lspé
N 2832	sleep	sueño	sommeil	sonno	sono	The state of reduced consciousness during which a human or animal rests in a daily rhythm	Lspé
N 3520	slide	portaobjetos	préparation microscopique	vetrino	lâmina	(sciences) A flat, rectangular piece of glass on which a prepared sample may be viewed through a microscope	Lspé
N 1581	slope	pendiente	penne	pendenza	declive	(mathematics) The ratio of the vertical and horizontal distances between two points on a line; zero if the line is horizontal, undefined if it is vertical	LST
A 1969	slow	lento	lent	lento	lento	Not happening in a short time ; spread over a comparatively long time	LST
A 20151	small	pequeño	petit	piccolo	pequeno	Not large or big; insignificant; few in numbers or size	LST
N 1767	smoker	fumador	fumeur	fumatore	fumador / fumante	A person who smokes tobacco habitually	Lspé
N 4169	smoking	tabaquismo	tabagisme	tabagismo	tabagismo	The burning and inhalation of tobacco	Lspé
A 2748	smooth	liso	lisse	liscio	liso	(medicine) Of muscles : involuntary and non striated	Lspé
N 2338	SNP	SNP	SNP	SNP	SNP	Single nucleotide polymorphism. A single base pair of DNA which is polymorphic with respect to a population	Lspé
A 2462	social	social	social	sociale	social	1) social : Relating to society	LST
A 2462	social	social	social	sociale	social	2) social : (Internet) Relating to social media or social networks	LST
A 2462	social	social	social	sociale	social	3) social : (botany, zoology) Cooperating or growing in groups	Lspé
N 3255	society	sociedad	société	società	sociedade	A group of people who meet from time to time to engage in a common interest ; an association or organization	LST
N 3412	sodium	sodio	sodium	sodio	sódio	A soft, waxy silvery reactive metal that is never found unbound in nature, and a chemical element (symbol Na) (...)	Lspé
N 6750	software	software	logiciel	software	software	(computing) Encoded computer instructions, usually modifiable (...)	LST
A 550	solid	sólido	solide	solido	sólido	1) : solid : Large, massive	LST
A 550	solid	sólido	solide	solido	sólido	2) solid : In the state of a solid	Lspé
A 550	solid	continuo	continu	continuo	continuo	3) solid : (Typography) written as a word, without spaces or hyphens	LST
A 2748	soluble	soluble	soluble	solubile	solúvel	Able to be dissolved	Lspé
N 4176	solution	solución	solution	soluzione	solução	1) solution : A homogeneous mixture, which may be liquid, gas or solid, formed by dissolving one or more substances	Lspé
N 4176	solution	solución	solution	soluzione	solução	2) solution : An act, plan or other means, used or proposed, to solve a problem	LST
A 1648	somatic	somático	somatique	somatico	somático	Of, or relating to the wall of the body ; somatopleuric ; parietal	Lspé
N 8987	source	fuelle	source	fonte	fonte	The person, place or thing from which something (information, goods, etc.) comes or is acquired	LST



A 1725	south	del sur	du sud	del sud	do sul	Of or pertaining to the south ; southern	LST
N 3859	space	espacio	espace	spazio	espaço	A (chiefly empty) area or volume with set limits or boundaries	LST
A 3996	spatial	espacial	spatial	spaziale	espacial	Pertaining to space	LST
A 2087	special	especial	spécial	speciale	especial	Distinguished by a unique or unusual quality	LST
N 16013	species	especie	espèce	specie	espécie	(biology, taxonomy) A rank in the classification of organisms, below genus and above subspecies ; a taxon at that rank	Lspé
A 34582	specific	específico	spécifique	specifico	específico	Special, distinctive or unique	LST
N 6417	specificity	especificidad	spécificité	specificità	especificidad e	The state of being specific rather than general	LST
V 1805	specify	especificar	spécifier	specificare	especificar	To state explicitly, or in detail, or as a condition	LST
N 4047	specimen	espécimen	spécimen	campione	espécime	A sample, especially one used for diagnostic analysis	Lspé
N 2119	spectrum	espectro	spectre	spettro	espectro	A range ; a continuous, infinite, one-dimensional set, possibly bounded by extremes	LST
N 2880	sperm	esperma	sperme	sperma	esperma	Semen ; the generative substance of male animals	Lspé
A 2337	spinal	espinal	spinal	spinale	espinal / espinhal	Of or relating to the spine	Lspé
V 2447	splice	empalmar	épisser	tagliare e ricongiungere	cortar e ligar	(genetics) To remove intron sequences from the pre-messenger RNA, and then join together exons	Lspé
N 1616	splice	empalme / ajuste / * splicing	épissage	* splicing / rottura e riunione	* splicing / excisão- reparação /	The process of removing from the pre-messenger RNA intron sequences, and then joining together exons	Lspé
A 2070	spontaneous	espontáneo	spontané	spontaneo	espontâneo	Self-generated, happening without any apparent external cause	Lspé
N 3574	spot	zona	zone	zona	zona	A location or area	LST
N 3018	stability	estabilidad	stabilité	stabilità	estabilidade	The condition of being stable or in equilibrium, and thus resistant to change	Lspé
A 3836	stable	estable	stable	stabile	estável	Relatively unchanging	Lspé
N 2865	staff	personal	personnel	personale	peçoal	The employees of a business	LST
N 11403	stage	etapa	étape	tappa	etapa corar /	A phase	LST
V 7403	stain	colorear	colorer	colorare	colorar / colorir	To treat a microscope specimen with a dye	Lspé
N 2358	staining	tinción / coloración	coloration	colorazione	coloração	An auxiliary technique used in microscopy and consisting in adding a dye to a substrate in order to qualify or quantify the presence of a specific compound	Lspé
N 4040	standard	estándar	standard	standard	standard	A principle or example or measure used for comparison	LST
A 11432	standard	estándar	standard	standard	standard	Falling within an accepted range	LST

A 1794	standardized	estandarizado	standardisé	standardizzato	estandardizado	Designed in a standard manner or according to an official standard	LST
N 2835	start	comienzo / inicio	commencement / début	cominciamento / inizio	começo / início	The beginning of an activity	LST
V 4222	start	comenzar / empezar	commencer	cominciare	começar	To begin, commence, initiate	LST
N 5650	state	estado	état	stato	estado	1) state : A condition ; a set of circumstances applying at a given time	LST
N 5650	state	estado	état	stato	estado	2) state : A political division	LST
V 1625	state	declarar	déclarer	dichiarare	declarar	To declare to be a fact	LST
N 1641	statement	declaración	déclaration	dichiarazione	declaração	A presentation of opinion or position	LST
N 5893	statistic	estadística	statistique	statistica	estatística	A single item in a statistical study	LST
A 13887	statistical	estadístico	statistique	statistico	estatístico	Of or pertaining to statistics	LST
N 11529	status	estatus / estado	statut / état	stato	estado	A person's condition (...)	LST
N 2448	stem	tronco / tallo	tronc	tronco	tronco	(anatomy) A part of an anatomic structure considered without its branches or ramifications	Lspé
N 8314	step	etapa	étape	tappa	etapa	Proceeding ; measure (...)	LST
N 2792	steroid	esteroide	stéroïde	steroidi	esteroide	(biochemistry) A class of organic compounds having a structure of 17 carbons arranged in four rings (...)	Lspé
V 7389	stimulate	estimular	stimuler	stimolare	estimular	(biochemistry) To arouse an organism to functional activity	Lspé
N 6376	stimulation	estimulación	stimulation	stimolazione	estimulação	(biology) Any action or condition that creates a response ; sensory input	Lspé
N 3236	stimulus	estímulo	stimulus	stimolo	estímulo	Something external that elicits or influences a physiological or psychological activity or response	Lspé
V 3101	store	almacenar	stocker / emmagasiner	immagazzinare	estocar / armazenar	(computing) Write into memory or registers	LST
N 14267	strain	cepa / estirpe	souche	ceppo / stipite	cepa / estirpe	(biology) A particular breed or race of animal, microbe etc.	Lspé
N 3548	strand	cadena / hebra	brin	filamento	cadeia / fita	(genetics) A nucleotide chain	Lspé
N 8923	strategy	estrategia	stratégie	strategia	estratégia	A plan of action intended to accomplish a specific goal	LST
N 2556	strength	fuerza	force	forza	força	The intensity of a force or power; potency	Lspé
N 8440	stress	estrés	stress	stress	stress / estresse	(biology) Aggression of an organism resulting in a response in an attempt to restore previous conditions	Lspé
N 4806	stroke	accidente * cerebrovascul	accident vasculaire	accidente cerebrovascola	acidente vascular	(medicine) The loss of brain function arising when the blood supply to the brain is suddenly interrupted	Lspé

		ar / ACV	cérébral / AVC	re / ictus	cerebral / accidente vascular encefálico / AVC / AVE		
A 3540	strong	fuerte	fort	forte	forte	1) strong : (mathematics, logic) Having a wide range of logical consequences; widely applicable	LST
A 3540	strong	fuerte	fort	forte	forte	2) strong : Convincing	LST
A 6419	structural	estructural	structurel	strutturale	estrutural	Of, relating to or having structure	LST
N 23169	structure	estructura	structure	struttura	estrutura	The overall form or organization of something	LST
N 6135	student	estudiante	étudiant	studente	estudante	A person who studies a particular academic subject	LST
N 12966 3	study	estudio	étude	studio	estudo	Any particular branch of learning that is studied ; any object of attentive consideration	LST
V 11003	study	estudiar	étudier	studiare	estudar	To acquire knowledge on a subject	LST
N 2769	subgroup	subgrupo	sous-groupe	sottogruppo	subgrupo	A group within a larger group ; a group whose members are some, but not all, of the members of a larger group	LST
N 17423	subject	sujeto	sujet	soggetto	sujeito	A human, animal or an inanimate object that is being examined, treated, analyzed, etc.	Lspé
V 2241	subject	someter	soumettre	sottoporre	submeter	To cause to undergo a particular experience (...)	LST
A 1986	subject	sujeto	sujet	soggetto	sujeito	Likely to be affected by or to experience something	Lspé
A 5195	subsequent	subsiguiente	subséquent	sussequente	subsequeute	Following in order of place; succeeding	LST
N 4502	subset	subconjunto	sous-ensemble	sottoinsieme	subconjunto	A group of things or people, all of which are in a specified larger group	LST
N 2281	substance	sustancia	substance	sostanza	substância	A form of matter that has constant chemical composition and characteristic properties	Lspé
A 2680	substantial	sustancial	substantiel	sostanziale	substancial	Large in size, quantity or value	LST
N 3860	substitution	sustitución	substitution	sostituzione	substituição	(chemistry) The replacement of an atom, or group of atoms, in a compound, with another	Lspé
N 5152	substrate	sustrato	substrat	substrato	substrato	(biochemistry) The material or substance on which an enzyme acts upon	Lspé
N 1612	subtype	subtipo	sous-type	sottotipo	subtipo	A group of specific things within a larger, more general group	LST
N 6681	subunit	* subunidad	sous-unité	* subunità	subunidade	(biochemistry) Any protein that forms assemblies with others, especially those that form multimers or oligomers	Lspé
N 2383	success	éxito	succès	successo	sucesso / êxito	The achievement of one's aim or goal	LST
A 2628	successful	exitoso	réussi	riuscito	sucedido	Resulting in success, having the desired effect	LST
V 1801	suffer	sufrir	souffrir	soffrire	sofrer	To have a disease or condition	Lspé
A 3683	sufficient	suficiente	suffisant	sufficiente	suficiente	Equal to the end proposed ; adequate to wants, enough	LST

V	31904	suggest	sugerir	suggérer	suggestire	sugerir	To make one suppose; cause one to suppose	LST
A	1886	suitable	apropiado	approprié	appropriato	apropiado	Appropriate to a certain occasion	LST
N	1985	sulfate	sulfato	sulfate	solfato	sulfato	(chemistry) Any salt of sulfuric acid.	Lspé
N	1752	sum	suma	somme	somma	soma	A quantity obtained by addition or aggregation	LST
V	2824	summarize	resumir	résumer	riassumere	resumir / recopilar	To give a recapitulation of the salient facts; to recapitulate or review	LST
N	2678	summary	resumen	résumé	riassunto	resumo	An abstract or a condensed presentation of the substance of a body of material	LST
N	2530	superfamily	* superfamilia	* superfamille	superfamiglia	superfamilia	(molecular biology) A large group of related proteins or other molecules	Lspé
N	2647	supernatant	sobrenadante	surnageant	sopranatante	sobrenadante	The liquid that lies above a sediment or precipitate; supernate	Lspé
V	2392	supplement	* suplementar	supplémenter	* supplementare	suplementar	To provide or make a supplement to something	Lspé
N	3753	support	ayuda	aide	aiuto	ajuda	1) support : Financial or other help	LST
N	3753	support	prueba	preuve	prova	prova	2) : support : Evidence	LST
V	13742	support	convalidar / apoyar	valider / soutenir	convalidare / avvalorare	avaliar / apoiar	To verify; to make good; to substantiate; to establish; to sustain	LST
V	2595	suppress	suprimir	supprimer	sopprimere	suprimir	To stop (...)	Lspé
N	2270	suppression	supresión	suppression	soppressione	supressão	The act or instance of suppressing	Lspé
N	1862	suppressor	supresor	suppresseur	soppressore	suppressor	A gene that suppresses the effect of another through epistasis	Lspé
N	9021	surface	superficie	surface	superficie	superficie	The outside hull of a tangible object	LST
N	1646	surfactant	surfactante / * tensoactivo / tensoactivo	* surfactant / tensioactif / agent de surface / * surfactif	* surfactante / surfattante / tensioattivo	surfactante / tensioativo / tensoativo	(chemistry) A surface active agent, or wetting agent, capable of reducing the surface tension of a liquid (...)	Lspé
N	6691	surgery	cirugía	chirurgie	chirurgia	cirurgia	(medicine) A procedure involving major incisions to remove, repair, or replace a part of the body	Lspé
A	4171	surgical	quirúrgico	chirurgical	chirurgico	cirúrgico	Of or relating to surgery	Lspé
N	2524	surveillance	vigilancia	surveillance	sorveglianza	vigilância	Continuous monitoring of disease occurrence (...)	Lspé
N	9456	survey	encuesta	enquête	inchiesta	inquérito	A questionnaire or similar instrument used for examining the opinions of a group of people	LST
N	9191	survival	supervivencia	survie	sopravvivenza	sobrevivência	The fact or act of surviving ; continued existence or life	Lspé
N	3798	susceptibility	susceptibilidad d	susceptibilité	suscettibilità	susceptibilidad de	The condition of being susceptible ; vulnerability	Lspé

N	10209	symptom	síntoma	symptôme	sintomo	sintoma	(medicine) A perceived change in some function, sensation or appearance of a person that indicates a disease or disorder (...)	Lspé
N	8659	syndrome	síndrome	syndrome	sindrome	síndrome	(pathology) A recognizable pattern of signs, symptoms and/or behaviors	Lspé
A	4208	synovial	sinovial	synovial	sinoviale	sinovial	Of, pertaining to, or producing synovia	Lspé
N	2422	synthase	* sintasa	* synthase	sintasi	sintase	(biochemistry) Any enzyme that catalyzes the synthesis of a biological compound but, unlike synthetases, does not make use of ATP as a source of energy	Lspé
N	7138	synthesis	síntesis	synthèse	sintesi	síntese	(chemistry) The reaction of elements or compounds to form more complex compounds	Lspé
V	2245	synthesize	sintetizar	synthétiser	sintetizzare	sintetizar	(chemistry) To produce a substance by chemical synthesis	Lspé
A	1725	synthetic	sintético	synthétique	sintetico	sintético	Of, or relating to synthesis	Lspé
N	10451	system	sistema	ystème	sistema	sistema	1) system : (medicine) The body organs that contribute to a vegetative function	Lspé
N	10451	system	sistema	ystème	sistema	sistema	2) system : A way of organizing or planning	LST
N	10451	system	sistema	ystème	sistema	sistema	3) system : (computing) The set of hardware and software operating in a computer	LST
A	2192	systematic	sistemático	systematique	sistemativo	sistemático	1) systematic : Carried out using a planned, ordered procedure	LST
A	2192	systematic	sistemático	systematique	sistemativo	sistemático	2) systematic : Regular	LST
A	4233	systemic	sistémico	systemique	sistemico	sistémico	Pertaining to an entire organism	Lspé
A	1585	systolic	sistólico	systolique	sistolico	sistólico	Pertaining to a systole or heart contraction	Lspé
N	3226	T-cell	linfocito T / célula T	lymphocyte T / cellule T	linfocita T	linfócito T / células T	(immunology) A lymphocyte (...)	Lspé
N	45238	table	tabla	tableau	tavola	tabela	A two-dimensional presentation of data	LST
N	2914	tag	etiqueta	étiquette	etichetta	etiqueta	(biology) Any short peptide sequence artificially attached to proteins mostly in order to help purify, solubilize or visualize these proteins	Lspé
V	15237	take	tomar	prendre	prendere	tomar	To get or put something into one's or someone's possession or control	LST
N	11993	target	objetivo	objectif	obiettivo	objetivo / alvo	A goal or objective	LST
V	4481	target	apuntar	cibler	mirare	apontar	To aim something (...)	LST
N	3649	task	tarea	tâche	compito	tarefa	A piece of work done as part of one's duties	LST
N	2097	team	equipo	équipe	squadro	equipe	Any group of people involved in the same activity (...)	LST
A	2953	technical	técnico	technique	tecnico	técnico	Of or pertaining to mechanic arts	LST
N	9524	technique	técnica	technique	tecnica	técnica	A method of achieving something or carrying something out, especially one requiring some skill or knowledge	LST
N	5042	technology	tecnología	technologie	tecnologia	tecnologia	The organization of knowledge for practical purposes	LST
N	8958	temperature	temperatura	température	temperatura	temperatura	A measure of cold or heat, often measurable with a thermometer	Lspé
N	2815	template	molde	matrice	stampo	molde	(molecular biology) A macromolecule which provides a pattern for the synthesis of another molecule	Lspé
A	1000	temporal	temporal	temporel	temporale	temporal	1) temporal : Of or relating to time	LST

A 1000	temporal	temporal	temporal	temporale	temporal	2) temporal : of the temples of the head	Lspé
V 2904	tend	tender	tendre	tendere	tender	To be likely, or probable to do something, or to have a certain characteristic	LST
N 11453	term	término	terme	termine	termo	A word or a phrase, especially one from a specialized area of knowledge	LST
N 2464	terminal	terminal	terminaison	terminale	terminal	(biology) The end ramification (of an axon, etc.) or one of the extremities of a polypeptide	Lspé
N 25507	test	prueba	test	prova	teste / prova	A challenge, trial	LST
V 14348	test	testar	tester	testare	testar	To place a product or piece of equipment under everyday and/or extreme conditions and examine it for its durability, etc	LST
N 4328	testing	prueba	test	prova	teste	The act of conducting a test (...)	LST
N 2250	testis	testículo	testicule	testicolo	testículo	A testicle	Lspé
N 2274	text	texto	texte	testo	texto	A book, tome or other set of writings	LST
V 4501	thank	agradecer	remercier	ringraziare	agradecer	To express gratitude or appreciation toward	LST
A 1739	theoretical	teórico	théorique	teorico	teórico	Of or relating to theory ; abstract	LST
N 3422	theory	teoría	théorie	teoria	teoria	(sciences) A coherent statement or set of ideas that explains observed facts or phenomena, or that sets out the laws and principles of something known or observed (...)	LST
A 4611	therapeutic	terapéutico	thérapeutique	terapeutico	terapêutico	Of, or relating to therapy	Lspé
N 16226	therapy	terapia	thérapie	terapia	terapia	Treatment of disease or disability, physical or mental	Lspé
V 3825	think	pensar	penser	pensare	pensar	To be of the opinion	LST
A 4870	third	tercero	troisième	terzo	terceiro	The ordinal form of the cardinal number three ; coming after the second	LST
N 4806	threshold	umbral	seuil	soglia	limiar	(engineering) The quantitative point at which an action is triggered, especially a lower limit	LST
A 1890	thyroid	tiroides	thyroïde	tiroide	tiroide	Referring to the gland situated in the neck and its blood vessels, etc.	Lspé
N 43295	time	tiempo	temps	tempo	tempo	A measurement of a quantity of time (...)	LST
N 30743	tissue	tejido	tissu	tessuto	tecido	(biology) A group of similar cells that function together to do a specific job	Lspé
N 2142	tolerance	tolerancia	tolérance	tolleranza	tolerância	The ability of the body (or other organism) to resist the action of a poison, to cope with a dangerous drug or to survive infection by an organism	Lspé
N 8009	tool	herramienta	outil	attrezzo	ferramenta	Equipment used in a profession	LST
A 1719	top	superior	supérieur	superiore	superior	Situated on the top of something	LST
N 1574	topology	topología	topologie	topografia	topologia	(medicine) The anatomical structure of part of the body	Lspé
N 5393	total	total	total	totale	total	An amount obtained by addition of smaller amounts	LST
A 16798	total	total	total	totale	total	Entire, relating to the whole of something	LST
N 2436	toxicity	toxicidad	toxicité	tossicità	toxicidade / toxicidade	(toxicology) The quality of being toxic	Lspé
N 4049	tract	tracto / vía	tractus / voie	tratto / via	tracto / via	A series of connected body organs, as in the digestive tract	Lspé

A 2405	traditional	tradicional	traditionnel	tradizionale	tradicional	Of or pertaining to tradition	LST
V 1619	train	entrenar / formar / adistrar entrenamiento	entraîner / rééduquer	formare / addestrare	treinar / formar / adestrar	To teach and form by practice; to educate; to exercise with discipline	Lspé
N 6341	training	o / formación / adiestramiento	entraînement / rééducation	formazione / addestramento	treinamento / formação / adestramento	The activity of imparting and acquiring skills	Lspé
N 3621	trait	rasgo	trait	tratto	traço	An identifying characteristic, habit or trend	Lspé
N 9085	transcript	transcrito	transcrit	trascritto	transcrito	(genetics) A sequence of RNA produced by transcription	Lspé
N 12989	transcription	transcripción	transcription	trascrizione	transcrição	(genetics) The synthesis of RNA under the direction of DNA.	Lspé
A 6060	transcriptional	* transcripcional	* transcriptionnel	* trascrizionale	* transcricional	Of or pertaining to transcription, especially to the transcription of genetic information	Lspé
N 3206	transduction	transducción	transduction	trasduzione	transdução	(biology) The transfer of genetic material from one bacterial cell to another by a bacteriophage or plasmid	Lspé
V 1868	transfect	* transfectar	* transférer	* trasfettare	* transfectar	(genetics) To introduce foreign material into eukaryotic cells	Lspé
N 2859	transfection	* transfección	transfection	transfezione	transfecção	(molecular biology) The introduction of foreign DNA into a eukaryotic cell	Lspé
N 5684	transfer	transferencia	transfert	trasferimento	transferência	(genetics) The conveying of genetic material from one cell to another	Lspé
V 3331	transfer	transferir	transférer	trasferire	transferir	To move or pass from one place, person or thing to another	Lspé
V 4577	transform	transformar	transformer	trasformare	transformar	(genetics) To subject (a cell) to transformation.	Lspé
N 3322	transformation	transformación	transformation	trasformazione	transformação	(genetics) The alteration of a bacterial cell caused by the transfer of DNA from another, especially if pathogenic	Lspé
A 4278	transgenic	transgénico	transgénique	transgenico	transgênico	Of or pertaining to an organism whose genome has been changed	Lspé
A 2099	transient	transitorio	transitoire	transitorio	transitório	Disappearing with time, transitory	LST
N 1437	transition	transición	transition	transizione	transição	1) transition : (genetics) A point mutation in which one base is replaced by another of the same class (purine or pyrimidine)	Lspé
N 1437	transition	transición	transition	transizione	transição	2) transition : The process of change from one form, state, style or place to another	LST
N 3510	translation	traducción	traduction	traduzione	tradução	(genetics) The process whereby a strand of mRNA directs assembly of amino acids into proteins within a ribosome	Lspé
N 1759	translocation	traslocación	translocation	traslocazione	translocação	(genetics) A transfer of a chromosomal segment to a new position, especially on a nonhomologous chromosome; the segment so transferred	Lspé
N 4680	transmission	transmisión	transmission	trasmissione	transmissão	(medicine, biology) The passing of a communicable disease from an infected host individual or group to a conspecific individual or group	Lspé

N 1827	transplantati on	trasplante	transplantation	trapiantazione / trapianto	transplantação o	(medicine) A surgical operation in which an organ is moved from a donor to a recipient; an organ transplant	Lspé
N 3927	transport	transporte	transport	trasporto	transporte	An act of transporting ; conveyance	Lspé
N 1857	transporter	transportador	transporteur	trasportatore	transportador	(biochemistry) A carrier	Lspé
N 1834	trauma	trauma	traumatisme / trauma	trauma	trauma	Any serious injury to the body, often resulting from violence or an accident	Lspé
V 8820	treat	tratar	traiter	trattare	tratar	1) treat : To care for medically or surgically; to apply medical care to	Lspé
V 8820	treat	tratar	traiter	trattare	tratar	2) treat : To subject to a chemical or other action (...)	Lspé
N 45035	treatment	tratamiento	traitement	trattamento	tratamento	Medical care for an illness or injury	Lspé
N 7833	tree	árbol	arbre	albero	árvore	Any structure or construct having branches	Lspé
N 6475	trend	tendencia	tendance	tendenza	tendência	A tendency	LST
N 21034	trial	ensayo	essai	prova	ensaio	(...) a test	Lspé
V 1976	trigger	desencadenar	déclencher	innescare	desencadear	To initiate something	Lspé
A 4702	true	verdadero	vrai	vero	verdadeiro	Concurring with a given set of facts	LST
N 3633	tube	tubo	tube / éprouvette	tubo	tubo	An approximately cylindrical container	Lspé
N 3227	tuberculosis	tuberculosis	tuberculose	tubercolosi	tuberculose	(pathology) An infectious disease of humans and animals caused by a species of mycobacterium, usually Mycobacterium tuberculosis (...)	Lspé
N 34144	tumor	tumor	tumeur	tumore	tumor	An abnormal growth (...)	Lspé
N 37353	type	tipo	type	tipo	tipo	A grouping based on shared characteristics ; a class	LST
A 2654	typical	típico	typique	tipico	típico	Characteristically representing something by form, group, idea or type	LST
N 3094	tyrosine	tirosina	tyrosine	tirosina	tirosina	(biochemistry) A nonessential amino acid C <sub>9</sub> H <sub>11</sub> NO <sub>3</sub> found in most animal proteins, especially casein	Lspé
N 2018	ultrasound	ultrasonido	ultrason	ultrasuono	ultrassom	(physics) Sound with a frequency greater than the upper limit of human hearing ; approximately 20 kilohertz	Lspé
A 1737	unable	incapaz	incapable	incapace	incapaz	Not able, lacking a certain ability	LST
A 1948	unclear	incierto	incertain	incerto	incerto	Ambiguous : liable to more than one interpretation	LST
V 5549	undergo	someterse	être soumis / subir	sottoporsi	submeter-se	To experience; to pass through a phase	LST
V 1611	underlie	subyacer	sous-tendre / être à la base	essere alla base	fundamentar	To serve as a basis of; form the foundation of	LST
A 2327	underlying	subyacente	sous-jacent	sottostante	subjacente	Implicit	LST
V 5173	understand	entender / comprender /	comprendre	capire / intendere /	perceber / entender /	To be aware of the meaning of	LST



		percibir		comprendere	compreender		
N 3233	understanding	comprensión	compréhension	comprensione	compreensão	Mental, sometimes emotional process of comprehension, assimilation of knowledge, which is subjective by its nature	LST
V 1860	undertake	emprender	entreprendre	intraprendere	emprender	To take upon oneself; to start, to embark on (a specific task etc.)	LST
A 4677	unique	único	unique	unico	único	Being the only one of its kind	LST
N 4374	unit	unidad	unité	unità	unidade	1) unit : (sciences) A standard measure of a quantity	LST
N 4374	unit	unidad	unité	unità	unidade	2) unit : An organized group comprising people and/or equipment	LST
N 8701	university	universidad	université	università	universidade	Institution of higher education (...)	LST
A 4808	unknown	desconocido	inconnu	sconosciuto	desconhecido	Not known	LST
A 2042	unlikely	improbable	improbable	improbabile	improvável	Improbable	LST
A 1944	unpublished	inédito	inédit	inedito	inédito	Not published	LST
A 1777	untreated	no tratado	non traité	non trattato	não tratado	Not treated	Lspé
A 3230	upper	superior	supérieur	superiore	superior	At a higher level, rank or position	LST
N 3152	uptake	captación	captage / capture	captazione	captação	The absorption, especially of food or nutrient by an organism	Lspé
A 1034	urban	urbano	urbain	urbano	urbano	1) urban : Related to the city	LST
A 1034	urban	urbano	urbain	urbano	urbano	2) urban : Characteristic of city life	LST
A 1113	urinary	urinario	urinaire	urinario	urinário	1) urinary : (anatomy) Pertaining to urine	Lspé
A 1113	urinary	urinario	urinaire	urinario	urinário	2) urinary : (anatomy) Of or relating to the organs involved in the formation an excretion of urine	Lspé
N 1804	usage	uso	usage	uso	uso	The manner or the amount of using; use	LST
N 25180	use	uso / empleo	usage / emploi	uso / impiego	uso / emprego	The act of using	LST
V 14328	use	usar / utilizar	utiliser	usare / utilizzare	usar / utilizar	To employ; to apply; to utilize	LST
A 5885	useful	útil	utile	utile	útil	Having a practical or beneficial use	LST
N 6209	user	usuario	utilisateur	utente	usuário / utilizador	One who uses or makes use of something (...)	LST
A 1802	uterine	uterino	utérin	uterino	uterino	Of or relating to the uterus	Lspé
N 2106	utility	utilidad	utilité	utilità	utilidade	The state or condition of being useful; usefulness	LST
N 2218	utilization	utilización	utilisation	utilizzazione	utilização	The act of using something	LST
V 3569	utilize	utilizar / usar	utiliser	utilizzare /	utilizar / usar	To make use of; to use	LST

				usare				
N 2764	vaccination	vacunación	vaccination	vaccinazione	vacinação	Inoculation with a vaccine, in order to protect from a particular disease or strain of disease	Lspé	
N 3983	vaccine	vacuna	vaccin	vaccino	vacina	(immunology) A substance given to stimulate the body's production of antibodies and provide immunity against a disease, prepared from the agent that causes the disease, or a synthetic substitute	Lspé	
A 1562	valid	válido	valide	valido	válido	Acceptable, proper or correct	LST	
V 2666	validate	validar	valider	validare	validar	To check or prove the validity of ; verify	LST	
N 3315	validation	validación	validation	validazione	validação	The act of validating something	LST	
N 3813	validity	validez	validité	validità	validade	A quality of a measurement indicating the degree to which the measure reflects the underlying construct, that is, whether it measures what it purports to measure	LST	
N 33604	value	valor	valeur	valore	valor	Numerical quantity measured or assigned or computed	LST	
N 4255	variability	variabilidad	variabilité	variabilità	variabilidade	The degree to which a thing is variable (...)	LST	
N 11562	variable	variable	variable	variabile	variável	(mathematics) A quantity that may assume any one of a set of values	LST	
A 3129	variable	variable	variable	variabile	variável	Marked by diversity or difference	LST	
N 5467	variance	varianza	variance	varianza	variância	(statistics) The second central moment in probability	LST	
N 2234	variant	variante	variant	variante	variante	1) variant : (genetics) A different sequence of a gene	Lspé	
N 2234	variant	variante	variant	variante	variante	2) variant : Something that is slightly different from a type or norm	LST	
N 5619	variation	variación	variation	variazione	variação	1) variation : (genetics) The modification of a hereditary trait	Lspé	
N 5619	variation	variación	variation	variazione	variação	2) variation : (...) A partial change in the form, position, state or qualities of a thing	LST	
N 4494	variety	variedad	variété	varietà	variedade	The quality of being varied ; diversity	LST	
A 10856	various	vario	varié	vario	vário	Having a broad range of different elements	LST	
V 6544	vary	variar	varier	variare	variar	Not to remain constant: to change with time or a similar parameter	LST	
A 6238	vascular	vascular	vasculaire	vascolare	vascular	Of, pertaining to or containing vessels	Lspé	
N 9782	vector	vector	vecteur	vettore	vetor	(epidemiology) A carrier of a disease-causing agent	Lspé	
N 2026	vein	vena	veine	vena	veia	(anatomy) A blood vessel that transports blood from the capillaries back to the heart	Lspé	
N 1781	velocity	velocidad	vélocité / vitesse	velocità	velocidade	Rapidity of motion	LST	
A 2114	venous	venoso	veineux	venoso	venoso	Of or pertaining to veins	Lspé	
N 1136	ventilation	ventilación	ventilation	ventilazione	ventilação	1) ventilation : The bodily process of breathing ; the inhalation of air to provide oxygen, and the exhalation of spent air to remove carbon dioxide	Lspé	
N 1136	ventilation	ventilación	ventilation	ventilazione	ventilação	2) ventilation : (medicine) The mechanical system used to assist breathing	Lspé	
A 2787	ventricular	ventricular	ventriculaire	ventricolare	ventricular	Of or relating to a ventricle or ventriculus	Lspé	
V 2412	verify	verificar / averiguar	vérifier	verificare	verificar / conferir /	To confirm or test the truth or accuracy of something	LST	

						averiguar	
N 3292	version	versión	version	versione	versão	1) version : (computing) A particular revision (of software, firmware, CPU, etc.)	LST
N 3292	version	versión	version	versione	versão	2) version : A specific form or variation of something	LST
N 3233	vessel	vaso	vaisseau	vaso	vaso	(biology) A tube or canal that carries fluid in an animal or plant	Lspé
N 1717	viability	viabilidad	viabilité	vitalità	viabilidade	(...) The ability to live (...)	Lspé
N 2077	view	concepción	conception	concezione	concepção	1) view : A way of understanding something, an opinion, a theory	LST
N 2077	view	vista	vue	vista	vista	2) view : A picture (...)	LST
A 6886	viral	viral	viral	virale	viral	Of or relating to a biological virus	Lspé
N 15101	virus	virus	virus	virus	vírus	(pathology, microbiology, virology) A submicroscopic non-cellular structure consisting of a core of DNA or RNA surrounded by a protein coat, that requires a living host cell to replicate and often causes disease in the host organism	Lspé
N 3596	visit	consulta	consultation	consultazione	consulta	(medicine) A meeting with a doctor at their surgery or the doctor's at one's home	Lspé
A 2933	visual	visual	visuel	visuale	visual	Related to or affecting the vision	Lspé
N 1707	visualization	visualización	visualisation	visualizzazione	visualização	The act of visualizing, or something visualized	LST
V 2082	visualize	visualizar	visualiser	visualizzare	visualizar	To make visible	LST
N 3421	vitamin	vitamina	vitamine	vitamina	vitamina	Any of a specific group of organic compounds essential in small quantities for healthy human growth, metabolism, development and body function (...)	Lspé
N 2161	volt	voltio	volt	volt	volt	In the International System of Units, the derived unit of electrical potential and electromotive force (voltage) (...)	Lspé
N 8477	volume	volumen	volume	volume	volume	A unit of three-dimensional measure of space that comprises a length, a width and a height (...)	LST
N 3764	wall	pared	paroi	parete	parede	(anatomy, zoology, botany) A divisive or containing structure in an organ or cavity	Lspé
V 5655	wash	lavar	laver	lavare	lavar	(chemistry) To pass through or over a liquid for the purpose of purifying it, especially by removing soluble constituents	Lspé
N 8001	water	agua	eau	acqua	água	A chemical found at room temperature and pressure as a clear liquid (...)	Lspé
N 1772	wave	onda	onde	onda	onda	(physics) A moving disturbance in the energy level of a field	Lspé
N 7103	way	manera / modo	manière / façon	maniera / modo	maneira / modo	A method or manner of doing something	LST
A 2820	weak	débil	faible	debole	débil	Not having power to convince	LST
N 2669	web	web	web / toile	web	web	The World Wide Web	LST
N 1606	website	sitio	site	sito	site	A collection of HTML and subordinate documents on the World Wide Web that are typically accessible from the same URL and residing on the same server, and form a coherent, usually interlinked whole	LST
N 11496	week	semana	semaine	settimana	semana	Any period of seven consecutive days	LST

N	13131	weight	peso	poids	peso	peso	(physics) Mass.	LST
N	1881	well	pocillo	puits	pozzetto	poço / vaso / cela	(biology) In a microtiter plate, each of the small equal circular or square sections which serve as test tubes	Lspé
A	4396	white	blanco	blanc	bianco	branco	Of the Caucasian race	Lspé
A	4893	whole	entero	entier	intero	inteiro	Entire	LST
A	5547	wide	largo / amplio	large / ample	largo / ampio	largo / amplo	Large in scope	LST
A	6113	wild-type	de tipo natural / de tipo silvestre	de type naturel / de type sauvage	wild type / di ceppo selvatico	de tipo selvagem	(biology) The typical form of an organism, strain, gene or characteristic as it occurs in nature	Lspé
N	1457	window	ventana	fenêtre	finestra	janela	1) window : (medicine) The time between first infection and detectability	Lspé
N	1457	window	ventana	fenêtre	finestra	janela	2) window : (computing) A rectangular area on a computer terminal or screen (...)	LST
N	24115	woman	mujer	femme	donna	mulher	An adult female human	Lspé
N	2780	word	palabra	mot	parola	palavra	The smallest discrete unit of written language which has a particular meaning, composed of one or more letters or symbols and one or more morphemes	LST
N	6200	work	trabajo	travail	lavoro	trabalho	1) work : A literary, artistic or intellectual production	LST
V	1847	work	trabajar	travailler	lavorare	trabalhar	1) work : To operate in a certain place, area, or specialty	LST
N	6200	work	trabajo	travail	lavoro	trabalho	2) work : Something produced using the specified material or tool	LST
V	1847	work	funcionar	fonctionner	funzionare	funcionar	2) work : To function correctly; to act as intended; to achieve the goal designed for	LST
N	3392	worker	trabajador	travailleur	lavoratore	trabalhador	A person who performs labor for his living (...)	LST
N	5628	world	mundo	monde	mondo	mundo	The Earth	LST
V	3113	write	escribir	écrire	scrivere	escrever	(computing) To record (data) mechanically or electronically	LST
N	29951	year	año	année	anno	ano	The time it takes the Earth to complete one revolution of the Sun (...)	LST
N	7403	yeast	levadura	levure	lievito	levedura	A single-celled fungus of a wide variety of taxonomic families	Lspé
V	3839	yield	producir / rendir	produire / rendre	produrre / rendere	produzir / render	To furnish; to afford; to render; to give forth	LST
A	6931	young	joven	jeune	giovane	jovem	In the early part of life or growth	Lspé
N	1564	zinc	zinc / cinc	zinc	zinco	zinco	(chemistry) A chemical element (symbol Zn) with an atomic number of 30	Lspé

## Annexe 3 : Définitions à valider des experts

Nous avons formulé les définitions suivantes dans le Wiktionary ; elles devront être validées par des experts.

EST	(genetics) Expressed Sequence Tag. A short sub-sequence of a cDNA sequence
library	1) library : A collection of DNA material from a single organism or relative to a single disease 2) library : An institution which holds books and/or other forms of stored information for use by the public or qualified people
tag	(biology) Any short peptide sequence artificially attached to proteins mostly in order to help purify, solubilize or visualize these proteins
core	1) core : (medicine) A tiny sample of organic material obtained by means of a fine-needle biopsy 2) core : (biochemistry) The central part of a protein structure consisting in mostly hydrophobic aminoacids
acquisition	(computing) The process of sampling signals that measure real world physical conditions and converting these signals into digital numeric values that can be manipulated by a computer
prediction	A probability estimation based on statistical methods
invasion	(medicine) The spread of cancer cells, bacterias and such to the organism
label	(biochemistry) To replace specific atoms by their isotope in order to track the presence or movement of this isotope through a reaction, metabolic pathway or cell
enrich	1) enrich : (chemistry) To make to rise the proportion of a given constituent 2) enrich : To add new elements, to complete
buffer	(chemistry) To maintain the acidity of a solution near a chosen value by adding an acid or a base
overlap	(genetics) To have some similar nucleotide sequences
administer	(medicine) To give a drug to a patient, be it orally or by any other means
deliver	(medicine) To administer a drug
frame	1) frame : (genetics : reading frame) A way of dividing nucleotide sequences into a set of consecutive triplets 2) frame : (Internet) An individually scrollable region of a web page
delivery	1) delivery (medicine) Administration of a drug 2) delivery (genetics) Process of introducing foreign DNA into host cells 3) delivery :

	The act of giving birth
PI	Abbreviation for phosphatidylinositol, a phospholipid component of eukaryotic cell membranes, also abbreviated in PtdIns
alignment	(bioinformatics) A way of arranging DNA, RNA or protein sequences in order to identify regions of similarity
intervention	(medicine) Operation
transmission	(medicine, biology) The passing of a communicable disease from an infected host individual or group to a conspecific individual or group
implication	Logical consequence
progression	Development, increase, evolution.
follow-up	(medicine) The revisiting of a patient in ambulatory care
stress	(biology) Aggression of an organism resulting in a response in an attempt to restore previous conditions
recruit	1) recruit : To prompt a protein, leukocyte etc. to intervene in a given region of the body 2) recruit : To enroll or enlist new members or potential employees (...)
line	(genetics) Population of cells derived from a single cell and containing the same genetic makeup
formation	The process during which something comes into being and gains its characteristics
contribution	The taking part, often with the idea that it has lead to (scientific etc.) progress.
arm	1) arm : (anatomy) The extended portion of the upper limb, from the shoulder to the elbow 2) arm : (genetics) One of the two parts of a chromosome 3) arm : A group of patients in a medical trial
gap	1) gap : (genetics) Unsequenced region in a sequence alignment 2) gap : Something missing
altered	Modified, changed
close	Marked, evident
fusion	The result of the hybridization of two genes which originally coded for separate proteins
output	1) output : (computing) Data sent out to the computer, as to output device such as a monitor or printer 2) output : (medicine) The flow rate of body liquids such as blood and urine
probe	(biochemistry) Any group of atoms or molecules radioactively labeled in order to study a given molecule or other structure
automated	Made automatic
combined	Resulting from the addition of several sources, parts, elements, aspects, etc. able to be united together, to converge.
global	Which has to be considered in its entirety
limited	Of numbers, amounts, data : poor, small, felt as insufficient
possible	Apparently valid, likely, plausible.
present	Relating to something a person is referring to in the very context, with a deictic use similar to the demonstrative adjective this

thyroid	Referring to the gland situated in the neck and its blood vessels, etc.
binding	(chemistry) The action or result of making two or more molecules stick together
column	1) column : (chemistry) An object used to separate the different components of a liquid or to purify chemical compounds 2) column : A vertical line of entries in a table, usually read from top to bottom
coverage	(genetics) The average number of reads representing a given nucleotide in the reconstructed sequence
fiber	A long tubular cell found in muscle tissue ; myocyte
regulator	(genetics) A gene involved in controlling the expression of one or more other genes
candidate	A synonym for a candidate gene, i.e., a gene which may play a role in a given disease
grade	(medicine) The degree of malignity of a tumor expressed on a scale
compartment	(anatomy) A region in the body, delimited by a biological membrane
mediator	A chemical substance transmitting information to a targeted cell
prospective	(medicine, of research) That has a sufficient population size to produce reliable data over a sufficient period
replication	(biology) The process of producing replicas of DNA or RNA molecules
inclusion	1) inclusion : An addition or annex to a group, set or total 2) inclusion : (biology) A nuclear or cytoplasmic aggregate of stainable substances
maintenance	(biology) The natural process which keeps an organism alive
box	(genetics) One of two specific regions in a promoter.
transporter	(biochemistry) A carrier
BLAST	(Basic Local Alignment Search Tool) An algorithm for comparing primary biological sequence information.
adhesion	(biochemistry) The binding of a cell to a surface or substrate
compound	A pure chemical substance consisting of two or more different chemical elements that can be separated into simpler substances by chemical reactions
substance	A form of matter that has constant chemical composition and characteristic properties
coding	(genetics) Which has the same base sequence as the RNA transcript produced
conservation	(biology) The fact that a proportion of similar or identical nucleic acid sequences or protein structures have been maintained within given molecules or across species over a period of time
construct	(genetics) A segment of nucleic acid, created artificially, for transplantation into a target cell or tissue
predictor	An independent variable.
strand	(genetics) A nucleotide chain
divergence	The process in which two or more populations accumulate genetic changes (mutations) through time

heterogeneity	The quality of a substance which is not uniform
infusion	(medicine) The administration of liquid substances directly into a vein for medical purposes
intensive	(medicine) Related to the need to manage life-threatening conditions by means of sophisticated life support and monitoring
channel	(biochemistry) A type of membrane protein
colon	(anatomy) The last part of the digestive system
map	Graphical representation of relations existing between several objects, components or themes
well	(biology) In a microtiter plate, each of the small equal circular or square sections which serve as test tubes
monitoring	(medicine) The continuous observation of a patient's health condition for any changes which may occur over time, using a monitor or measuring device of some sort
positive	A positive result of a test
predictive	(medicine) Expressing the expected accuracy of a statistical measure or of a diagnostic test
reverse	(genetics) In which cDNA synthetization is obtained from an RNA template
splice	The process of removing from the pre-messenger RNA intron sequences, and then joining together exons
motif	In a nucleotide or amino-acid sequence, pattern that is widespread and has, or is conjectured to have, a biological significance
smooth	(medicine) Of muscles : involuntary and non striated
staining	An auxiliary technique used in microscopy and consisting in adding a dye to a substrate in order to qualify or quantify the presence of a specific compound
screen	(genetics) A technique used to identify genes so as to study gene functions
confidence	(probability) Likelihood something is reliable
band	(genetics) Each of the lines representing the different components of a molecule and made visible after an electrophoresis in gel of a mixture
lane	(genetics) Each of the columns representing a molecule and made visible after an electrophoresis in gel of a mixture
prevention	(medicine) Any measure such as information campaigns, vaccination, early diagnosis etc., intended to limit health-related risks
burden	(medicine) The total amount of toxins, parasites, cancer cells, plaque or the such present in an organism
instruction	A set of directions provided by a manufacturer for the users of a product or service
copy	(genetics) The result of gene or chromosomal duplication
terminal	(biology) The end ramification (of an axon, etc.) or one of the extremities of a polypeptide
error	1) error : Any alteration in the DNA chemical structure occurring during DNA replication, recombination or repairing 2) error :
affected	(statistics) The difference between a measured or calculated value and a true one Acted on, influenced, or moved, either physically or materially



sequence	Order in which amino-acid residues, connected by peptide bonds, lie in the chain of peptides and protein
site	Region of a protein, a piece of DNA or RNA where chemical reactions take place
set	A series of, a group of
event	(medicine) An episode of severe health conditions
individual	An element belonging to a population
release	(biochemistry) Process by which a chemical substance is set free
migration	(biochemistry) The movement of cells in particular directions to specific locations
superfamily	(molecular biology) A large group of related proteins or other molecules
control	To verify or check
release	(biochemistry) To set free a chemical substance
localize	To determine where something takes place or is to be found
accompany	To be found at the same time
expose	To submit to an active (mostly dangerous) substance like an allergen, ozone, nicotine, solvent, or to any other stress, in order to test the reaction, resistance, etc.
encode	Constitute the code necessary for the biosynthesis of a protein by means of a matrix so as to transcript DNA material
splice	(genetics) To remove intron sequences from the pre-messenger RNA, and then join together exons
represent	1) represent : To exhibit to another mind in language; to show (...) 2) represent : To constitute, to make up, to be an example of
initiation	1) initiation : (chemistry) The first step of transcription or of transduction 2) initiation : The act of initiating, or the process of being initiated or introduced
exchange	1) exchange : (biochemistry) The transfer of substances or elements like gas, amino-acids, ions etc. sometimes through a surface like a membrane 2) exchange : The action consisting in giving or borrowing something, often with an idea of reciprocity
window	1) window : (medicine) The time between first infection and detectability 2) window : (computing) A rectangular area on a computer terminal or screen (...)
application	1) application : A potential use, often found by scientific research 2) application : The act of applying or laying on, in a literal sense ; as the application of emollients to a diseased limb 3) application : (computing) A computer program or the set of software that the end user perceives as a single entity as a tool for a well-defined purpose
resolution	1) resolution : (computing) The degree of fineness with which an image can be recorded or produced 2) resolution : (medicine) In a pathological process, the phase during which pathogens and damaged tissues are removed by macrophages
system	1) system : (medicine) The body organs that contribute to a vegetative function 2) system : A way of organizing or planning 3) system : (computing) The set of hardware and software operating in a computer

light 1) light : Easily interrupted by stimulation 2) light : Gentle; having little force or momentum

open 1) open : (computing) Made public, usable without a free license 2) open : Not settled or adjusted ; not decided or determined (...) 3) open : (medicine) Resulting from an incision, puncture or any other process by which the skin no longer protects an internal part of the body

screen 1) screen :To search chemical libraries by means of a computational technique in order to identify chemical compounds which would potentially bind to a given biological target such as a protein 2) screen : To examine patients or treat a sample in order to detect a chemical or a disease, or to assess susceptibility to a disease

blot (biochemistry) A method of transferring proteins, DNA or RNA, onto a carrier

marker 1) marker : (medicine) A substance used as an indicator of a biological state, most commonly disease 2) marker : (genetics) A gene or DNA sequence with a known location on a chromosome that can be used to identify individuals or species

## Annexe 4 : Tableau des allosuffixes

Les deux dernières colonnes indiquent dans quelle classe grammaticale les suffixes sont utilisés.

### Allosuffixes des cinq langues

espagnol	français	italien	portugais	anglais		
-able, -ible, -uble	-able, -ible, -uble	-bile	-ável, -ível, -úvel	-able, -ible, -uble	A	
-acto	-act	-atto	-ato, -acto	-act, -ate	N	
-aje	-age	-aggio, -agine	-agem	-age	N	
-al	-al, -el	-ale	-al	-al	A	
-ano	-ain, -en	-ano	-ano	-an	A	
-ancia, -anza, -encia	-ance, -ence	-anza, -enza	-ança, -ência, -ença	-ance, -ence, -ency	N	
-ante, -ente	-ant, -ent	-ante	-ante, -ente	-ant, -ing	N	A
-arar	-arar	-arare	-arar	-are, -arate	V	
-ario	-aire	-ario	-ário	-ary, -ar	A	
-asa, -asis	-ase	-asi	-ase	-ase, -asis	N	
-ato	-at, -ate	-ato	-ato	-ate, -at	N	
-cibir, -cebir	-cevoir	-cepire, -cevere	-ceber	-ceive	V	
-ceso	-cès	-cesso	-cesso	-cess	N	
-cluir	-clure	-cludere	-cluir	-clude	V	
-cribir	-crire	-crivere	-crever	-cribe	V	
-cula, -culo	-cule	-cola, -colo	-culo, -cula	-cule, -cle	N	

-ducir	-duire	-durre	-duzir	-duce	V	
-ecto	-ect	-etto	-eto, -eito, -ecto	-ect	A	N
-ectar, -etar	-éter, -ecter	-ettare, -etare, -ettere	-etar, -ectar	-ect, -ete, -et	V	
-elo	-èle	-ello	-elo	-el	A	
-ender	-endre	-endere	-ender	-end	V	
-erso	-erse	-erso	-erso	-erse	A	
-esis	-èse	-esi	-ese	-esis	N	
-ferir	-férer	-ferire	-ferir	-fer	V	
-frir, -frecer	-ffrir	-ffrire	-frer, -ferecer	-ffer	V	
-forma	-forme	-forma	-forma	-form	N	
-formar	-former	-formare	-formar	-form	V	
-ficar	-fier	-ficare	-ficar	-fy	V	
-geno	-gène	-geno, -gene	-geno, -gênio	-gen	A	N
-genia	-génie	-genia	-genia	-geny	N	
-grama	-gramme	-gramma	-grama	-gram	N	
-gráfico	-graphique	-grafico	-gráfico	-graphic, -graphical	A	
-ico	-ique	-ico	-ico	-ic, -ical	A	
-icar	-iquer	-icare	-icar	-icate, -y	V	
-ica	-ique	-ica	-ica	-ics	N	
-ido	-ide	-ido	-ido, -ídio	-id	N	
-imo	-imum	-imo	-imo	-imum	N	
-imal	-imal	-imale	-imo	-imal	A	
-ino	-in	-ino	-ino	-ine	A	
-ina	-in, -ine	-ina, -ino	-ina	-in, -ine	N	
-inar	-iner	-inare	-inar	-inate, -ine	V	
-ismo	-isme	-ismo	-ismo	-ism	N	
-ista	-iste	-ista	-ista	-ist	N	
-istrar	-istrer	-istrare	-istrar	-ister	V	
-ito	-ite	-ite, -ita	-ite, -ita	-it, -ite	N	
-itis	-ite	-ite	-ite	-itis, -ite	N	

-ivo	-if	-ivo	-ivo	-ive	A
-iva	-ive	-iva	-iva	-ive	N
-izar, -idar	-iser, -ider	-izzare, -dare	-izar	-ize, -ide, -idate	V
-lar	-lier, -laire	-lare	-lar	-lar	A
-logo	-logue	-logo	-logo	-logue, -log	N
-logía	-logie	-logia	-logia	-logy	N
-miento,					
-mento	-ment	-mento	-mento	-ment	N
-metro	-mètre	-metro	-metro	-meter	N
-meter	-mettre	-mettere	-mitir	-mit	V
-oque	-oc	-occo	-oque	-ock	N
-ocito	-ocyte	-ocito	-ócito	-ocyte	N
-oide	-oïde	-oide	-oide	-oid	A
-ol	-ol, -ole, -ôle	-olo, -ollo, -ole, -uolo	-ol	-ol, -ole	N
-oma, -omo	-ome	-oma	-oma, -omo	-ome	N
-ón	-on	-one, -on	-on, -ão, -ônio	-on	N
-ona	-one	-ona	-ona	-one	N
-or	-eur	-ore	-or	-or	A N
-oria	-oire	-oria	-ória	-ory	N
-osis	-ose	-osi	-ose	-osis, -ose	N
-oso	-eux	-oso	-oso	-ous	A
-fago	-phage	-fago	-fago	-phage	N
-ponder	-pondre	-pondere	-ponder	-pond	V
-poner	-poser	-porre	-por	-pose	V
-ria	-rie	-ria	-ria	-ry, -ria	N
-scrito	-scrit	-scritto	-scrito	-script	N
-sión	-sion	-sione	-são	-sion	N
-tar	-ter	-tare	-tar	-t, -tate	V
-tir	-ter, -tir	-tere	-tir	-t, -te	V
-tener	-tenir	-tenere	-ter	-tain	V

-ción	-tion	-zione	-ção, -tão	-tion	N
-torio	-toire	-torio	-tório	-tory	A
-tud	-tude	-tudine	-tude	-tude	N
-dad	-té	-tà	-dade	-ty	N
-tipo	-type	-tipo	-tipo	-type	N
-ular	-uler	-olare, -ulare	-ular	-ulate	V
-ura	-ure	-ura	-ura	-ure	N
-us, -o	-us	-us, -o	-us, -o	-us	N
-var	-ver	-vare	-var	-ve, -vate	V
-tribuir	-tribuer	-tribuire	-tribuir	-tribute	V

Allosuffixes des quatre langues romanes

-cionar	-tionner	-zionare	-cionar	V	
-cuerpo	-corps	-corpo	-corpo	N	
-ero	-ier	-ero	-eiro	A	
-il	-ile	-ile	-il	A	
-uerte	-ort	-orte	-orte	A	N

## Annexe 5 : Liste du lexique spécifique de la santé

Cette liste rassemble tous les mots qui expriment l'une des 1017 acceptions de la santé du corpus scientifique ScienText. Ils sont donnés dans l'ordre des cinq sous-catégories que nous précisons dans la section Sous-catégories du lexique de spécialité de la santé, p.153. Le fichier complet des termes fréquents de la santé s'intitule `lexiquespecificquedelasANTE.ods`, il se trouvera sur le portail hal dans les documents de l'auteur de cette thèse.

### **Anatomie générale**

body, physical, physiology, physiological, function, region, topology, organ, tract, system, envelope, wall, somatic, compartment, epithelium, epithelial, follicle, dominant, island, stem, side, lateral, median, distal, longitudinal, left, right, anterior, posterior, basal, root, internal, external, skeletal, bone, height, head, joint, cartilage, synovial, head, temporal, alveolar, oral, neck, cervical, chest, abdominal, cord, spinal, arm, hand, hip, knee, skin, tissue, muscle, contraction, smooth, fiber, heart, cardiac, ventricular, systolic, atrial, myocardial, coronary, circulation, circulate, cardiovascular, vessel, vascular, vein, venous, artery, arterial, endothelial, flow, blood, lymph, neutrophil, lymphocyte, T-cell, monocyte, macrophage, marrow, angiogenesis, platelet, plasma, serum, ventilation, respiratory, air, oxygen, airway, lung, pulmonary, alveolar, exchange, brain, cerebral, cortex, perception, sensory, field, receptor, central, peripheral, nerve, nervous, neural, neuron, neuronal, motor, dendritic, eye, chamber, visual, sleep, nutrition, food, glucose, lipid, oil, fat, fatty, vitamin, uptake, digest, gastric, gastrointestinal, pylorus, liver, hepatic, pancreatic, intestinal, colon, colorectal, kidney, renal, bladder, urinary, urinary, sexual, reproductive, sex, gender, sexual, woman, female, female, uterine,

endometrial, man, male, male, testis, sperm, endogenous, secretion, secrete, gland, secretion, thyroid, breast, mammary, milk, ovary, ovarian, estrogen, progesterone, hormone, release, release, insulin, prostate, androgen, steroid, prostaglandin, dopamine, cAMP, glutamate, immunity, immune, affinity, defense, protect, protection, protective, barrier, resistance, resistant, chemokine, antibody, recognition, interleukin, cytokine, recognize

### **Anatomie comparée**

development, developmental, differentiation, differentiate, induction, oocyte, egg, pregnancy, pregnant, embryo, embryonic, fetal, labor, contraction, bear, delivery, birth, native, family, parent, mother, maternal, neonatal, infant, child, childhood, young, adult, adult, maturation, mature, age, aged, old, elderly, decline, resident, home, phylogeny, phylogenetic, evolution, evolutionary, biological, biology, nature, natural, species, compete, homologous, homologue, homology, divergence, linkage, wild-type, conservation, polymorphism, lineage, ortholog, noise, variation, mutant, mutant, mutation, null, transition, principle, morphology, morphologic, family, tree, branch, strain, organism, systemic, eukaryote, eukaryotic, female, male, animal, animal, mammal, mammalian, human, human, black, white, primate, monkey, pig, dog, bovine, murine, rabbit, rat, mouse, fish, insect, drosophila, mosquito, egg, yeast, Saccharomyces, cerevisiae, Caenorhabditis elegans, colony, social, plant, seed, leaf, line, division, cleavage, duplication

### **Pathologie humaine : anatomopathologie générale**

activity, condition, health, healthy, normal, intact, abnormality, pathology, pathological, pathogenesis, injury, fracture, damage, trauma, emergency, shock, pain, suffer, local, stress, failure, dysfunction, disorder, disease, illness, illness, ill, prognosis, prognostic, onset, spontaneous, early, advanced, functional, mild, chronic, primary, secondary, recurrent, recurrence, acute, severe, severeness, serious, critical, intensive, complication, syndrome, symptom, sign, temperature, fever, elevated, elevation, depression, cold, fatigue, affect, affected, alteration, alter, altered, impair, impairment, lose,



loss, deficiency, deficient, disability, assist, assistance, ventilation, inflammation, inflammatory, sepsis, hepatitis, arthritis, asthma, pneumonia, rheumatoid, osteoarthritis, node, fibrosis, cancer, tumor, mass, melanoma, malignant, grade, carcinoma, adenocarcinoma, metastasis, metastatic, proliferation, invasion, invasive, leukemia, oncogene, pressure, hypertension, stroke, infarction, MI, infarction, arrest, ischemia, ischemic, hypoxia, plaque, atherosclerosis, cholesterol, diabetes, diabetic, output, heparin, lesion, transplantation, match, donor, maintenance, regulation, stable, stability, stimulate, stimulation, stimulus, light, reaction, sensitive, sensitivity, trigger, respond, response, responsive, responsiveness, susceptibility, subject, autoimmune, immunodeficiency, allergy, expose, exposure, recruit, recruitment, immunization, immune, vaccination, vaccine, autoantibody, immunoglobulin, epitope, antigen, toxicity, cytotoxic, burden, clearance, eliminate, discharge, tolerance, process, recover, recovery, rehabilitation, resolution, regression, survival, live, life, apoptosis, apoptotic, lysate, necrosis, die, death, viability, psychiatric, psychiatry, psychological, cognitive, mental, dependence, drug, alcohol, consumption, ethanol, smoker, smoking, cigarette, depression, schizophrenia, anxiety, memory, obesity, feed, epidemiological, epidemiology, pathogen, parasite, virus, viral, phage, bacterium, bacterial, bacillus, Escherichia coli, vector, contact, transmission, infection, infect, infected, infectious, infectious, acquire, window, early, tuberculosis, malaria, Plasmodium falciparum, sampling, arm, subject, predictive, in vivo, prospective, prevalence, incidence, morbidity, mortality

### **Pathologie humaine : biologie médicale**

substance, fluid, fluid, liquid, solid, heat, energy, base, acid, acid, pH, electron, positive, oxide, particle, mmol, gas, oxygen, nitrogen, iron, ion, zinc, sodium, calcium, carbon, molecule, molecular, conformation, bind, binding, ligand, cluster, chain, complex, block, compound, phosphate, sulfate, nitric, extract, extract, extraction, column, reaction, reactive, reactivity, label, metabolism, metabolic, metabolite, oxidation, oxidative, antioxidant, agent, homology, affinity, adhesion, generate, generation, derive, product, chemical, precursor, analogue, substitution, isolation, isolate, isolated, cleavage, residue, residual, degradation, treat, digest, activation, activator, activate, purification, purify, fix, fixed, synthesis, synthesize, synthetic, histological, biopsy, sample, sample, core, specimen, fragment, section, microscope, microscopy, nm, slide, glass, analysis, preparation, prepare,

plate, well, trial, tube, in vitro, grow, culture, culture, cultured, incubate, incubation, matrix, solution, soluble, surfactant, medium, concentration, dilute, dilution, buffer, buffer, supplement, mix, mixture, centrifugation, supernatant, interface, layer, indicator, water, wash, filter, immunohistochemical, detection, detect, detectable, marker, assay, assay, protein, core, site, domain, loop, isoform, superfamily, cyclin, histone, interferon, lipoprotein, cytochrome, albumin, peptide, channel, cell, cellular, cytoplasm, cytoplasmic, membrane, PI, mitochondrion, mitochondrial, intracellular, extracellular, matrix, collagen, glycoprotein, actin, integrin, fibroblast, biosynthesis, biochemical, pathway, receptor, agonist, antagonist, template, hybridization, hybridize, hybrid, inhibition, inhibitor, inhibit, inhibitive, reverse, catalytic, enhance, enhanced, enrich, reagent, promote, promoter, substrate, enzyme, fluorescence, fluorescent, effector, protease, kinase, phosphatase, dehydrogenase, polymerase, synthase, nucleus, nuclear, nucleic, chip, array, microarray, electrophoresis, gel, inclusion, staining, stain, volt, dye, subunit, amino, acid, serine, cysteine, tyrosine

### **Thérapeutique**

professional, professional, practitioner, general, medicine, medical, practice, doctor, clinician, nurse, nursing, intern, physician, pediatric, care, patient, visit, follow-up, routine, primary, examination, examine, present, presentation, diagnose, diagnosis, diagnostic, history, episode, event, therapeutic, therapy, treat, treatment, untreated, compliance, consent, prescription, prescribe, administration, administer, delivery, deliver, medication, drug, medicine, intake, dose, load, antibiotic, aspirin, placebo, application, extract, exercise, strength, force, dynamics, move, movement, extension, train, training, diet, dietary, regimen, potent, beneficial, benefit, effect, adverse, risk, concern, negative, prevent, prevention, screening, screen, safety, hospital, hospitalization, clinic, clinical, admission, discharge, intervention, operation, operate, monitoring, monitor, surveillance, surgery, surgical, postoperative, open, section, removal, remove, draw, replacement, infusion, inject, injection, perfusion, catheter, intravenous, chemotherapy, light, laser, radiation, field, magnetic, wave, ultrasound, phenotype, phenotypic, trait, genomics, genome, genomic, chromosome, chromosomal, arm, locus, genetics, genotype, gene, genetic, allele, haplotype, chromatin, DNA, cDNA, EST, RNA, plasmid, terminal,

sense, downstream, complementary, SNP, mapping, map, frame, strand, nucleotide, oligonucleotide, codon, motif, repeat, annotation, annotate, overlap, alignment, align, sequence, sequence, gap, BLAST, variant, band, lane, ribosomal, mRNA, boundary, junction, candidate, inactivation, transformation, transform, convert, translocation, transgenic, conjugate, transfer, transfer, transduction, transfection, transfect, insert, insertion, introduce, introduction, complement, host, delivery, construct, recombination, recombinant, free, migration, carry, transport, transporter, carrier, mediator, mediate, suppression, suppress, suppressor, block, conserve, reverse, restriction, library, screen, screen, imaging, phosphorylation, marker, determinant, deletion, signaling, signal, probe, assembly, splice, splice, translation, expression, express, encode, fusion, methylation, overexpression, coverage, transcription, transcriptional, transcript, operon, box, copy, clone, clone, monoclonal, replication, replicate, error, repair, PCR, primer, amplification, amplify, initiation, antisense, intron, exon, code, coding, regulation, regulator, regulatory, reporter, tag, blot

## **Annexe 6 : liste du lexique scientifique transdisciplinaire**

Cette liste rassemble tous les mots qui expriment l'une des 1191 acceptions du LST du corpus scientifique ScienText. Ils sont donnés dans l'ordre des sept sous-catégories que nous précisons dans la section Sous-catégories du LST, p.156. Le fichier complet des termes fréquents de la santé s'intitule lexiqueTRANSDISCIPLINAIRE.ods, il se trouvera sur le portail hal dans les documents de l'auteur de cette thèse.

### **La connaissance**

mean, language, word, term, define, definition, knowledge, know, learn, perceive, knowledge, understand, understanding, insight, concept, fact, idea, represent, education, educational, school, student, history, university, academic, college, institute, school, laboratory, lab, inform, information, communication, publication, publication, publish, unpublished, access, open, private, public, common, library, write, text, content, solid, author, edit, work, manuscript, article, paper, literature, press, review, journal, interview, paper, source, reference, document, mention, illustrate, introduction, conclusion, example, express, declare, say, record, record, report, report, note, draft, preliminary, summarize, summary, section, display, display, visualization, visualize, see, representation, image, resolution, view, map, table, matrix, fig, figure, plot, scheme, column, code, PC, computer, code, program, application, package, system, software, version, language, list, acquisition, acquire, input, output, store, capture, kilobyte, memory, window, scan, web, website, database, query, hit, internet, network, interface, access, server, platform, page, frame, link, file, format, free, communication, social, present, presentation

### **Les dynamiques sociales**

society, social, government, office, community, organization, association, foundation, institution, institute, department, service, committee, assembly, exchange, group, member, belong, partner, people, external, background, setting, context, state, environment, environmental, situation, labor, work, task, pool, manual, manufacturer, public, private, sector, service, worker, staff, team, colleague, manage, management, room, office, facility, center, economic, cost, invest, interest, financial, income, dollar, purchase, rich, poor, support, grant, product, kit, world, international, geographical, country, state, national, regional, district, city, urban, rural, island, isolation, African, American, British, Canadian, European, north, south, population, demographic, household, home, live, ethic, ethical, current, behavior, behavioral, policy, law, right, thank, person, individual, respective, personal, status, traditional, consensus, conventional, usage, agreement, alternative, interesting, interest, new, novel, direction, opportunity, progress, advance, light, application, optimal, optimize, advantage, gain, potential, potential, contribute, contribution, help, help, improve, improved, improvement, useful, utility, good, facilitate, easy, simple, basic, effort, challenge, difficult, difficulty, difficulty, complex, complexity, adequate, appropriate, bad, meet, satisfaction, achieve, success, successful, work, get, obtain, allow, permit, enable, provide, provider, give, receive, accept, fit, suitable, sufficient, relevance, relevant, importance, important, significance, significant, key, principal, major, main, central, priority, crucial, essential, critical, secondary, minor, exist, existence, existing, actual, real, present, presence, occur, availability, available, unknown, need, need, lack, lack, miss, absence, absent, gap, ask, require, requirement, necessary, unable, possibility, possible, can, ability, able, capable, capacity, bad, poor, restriction, recognize, condition, barrier, failure, fail, negative, positive

### **Les orientations de la recherche**

science, scientific, field, phenomenon, investigation, research, investigate, investigator, researcher, expert, search, search, study, study, look, seek, see, examine, explore, observe, observation, attention, focus, focus, consider, consideration, consideration, attempt, attempt, issue, question, regard, concern, case, address, problem, resolve, objective, aim, aim, goal, purpose, target, target, solution, recommend, recommendation, recommendation,

hypothesis, hypothesize, base, base, basis, putative, expect, unlikely, likely, probability, think, believe, predict, feel, assume, assumption, admit, suggest, point, attitude, positive, theory, theoretical, review, review, framework, view, perspective, statement, state, underlie, underlying, law, offer, propose, choice, option, preference, select, selected, selection, selection, selective, method, methodology, procedure, mode, approach, decision, orientation, alternative, choose, demonstrate, strategy, project, conceive, create, design, design, build, system, systematic, organization, structure, structural, coordinate, coordination

### **La description des caractéristiques**

characterization, characterize, description, describe, refer, call, be, name, profile, characteristic, characteristic, identification, identify, classification, classify, type, subtype, class, kind, category, assign, specify, general, special, specific, specificity, particular, only, independent, rare, unique, representative, representative, typical, quality, property, feature, aspect, dimension, mixed, criterion, factor, determinant, detail, detailed, have, possess, own, free, appear, apparent, seem, appearance, form, shape, block, edge, solid, linear, composition, compose, contain, content, comprise, consist, constitute, form, represent, make, include, account, part, partial, portion, segment, element, individual, component, module, fraction, set, cluster, cluster, subgroup, subset, unit, pair, couple, accompany, full, total, whole, entire, global, overall, complete, full-length, comprehensive, single, double, multiple, majority, half, color, blue, red, green, brown

### **La démarche scientifique**

program, measure, model, model, ontology, pattern, experiment, experimental, empirical, test, test, testing, instruction, carry, follow, conduct, run, protocol, rule, guideline, process, processing, process, way, manner, mean, technique, technical, technology, skill, practice, practical, material, collect, collection, list, survey, questionnaire, item, score, score, elicit, recruit, respondent, participant, participate, participation, resource, tool, instrument, device, logistic, mechanical, mechanism, simulated, simulation, automated, act, action, behavior, direction, take, undertake, employ, use,

use, user, utilization, utilize, serve, function, implement, implementation, subject, drive, perform, do, apply, divide, incorporate, integrate, integration, feed, enrich, inclusion, enter, entry, combination, combine, combined, separate, separate, separation, exclude, exclusion, complete, leave, avoid, place, direct, construct, construction, work, course, scenario, series, cycle, stage, step, point, phase, session, round, preparation, prepare, initiation, initiate, start, begin, start, place, initial, prime, original, first, prior, previous, continue, follow, next, following, subsequent, posterior, second, secondary, third, late, last, final, end, time, temporal, duration, length, short, long, period, interval, past, recent, future, transient, prolonged, long-term, delay, occurrence, distance, frequent, frequency, common, rhythm, continuous, constant, regular, systematic, light, slow, velocity, rapid, fast, direct, real-time, disruption, experience, undergo, transition, formation, change, change, change, experience, evolve, develop, become, come, go, shift, replace, cross, retain, balance, keep, hold, remain, remaining, maintain, return, reverse, trend, tend

### **La mesure**

localization, localization, localize, locate, place, space, spatial, position, location, point, spot, area, top, upper, center, central, local, proximate, dimension, distance, length, short, long, surface, diameter, size, wide, volume, weight, mass, density, intensity, magnitude, measurement, measure, measure, determination, determine, set, count, reading, read, assess, assessment, estimate, estimation, estimation, estimate, evaluate, evaluation, rating, reliability, reliable, unit, km, cm, mm, ml, kg, mg, decade, year, annual, month, week, daily, day, hour, minute, min, second, ms, reading, amount, value, count, number, level, reach, extend, add, addition, additional, further, total, sum, divide, abundance, numerous, high, broad, great, large, substantial, considerable, extensive, dominant, maximal, maximum, peak, progression, progressive, extension, extend, expansion, expand, accumulation, accumulate, growth, grow, raise, rise, increase, increase, increased, moderate, threshold, limitation, limit, limit, limited, restrict, constraint, reduce, reduced, reduction, decrease, decrease, decline, contraction, lower, fall, low, small, minimal, minimize, minimum, analysis, analytic, analyze, calculation, calculate, computational, compute, quantification, quantify, quantity, quantitative, distribution, axis, parallel, curve, helix, slope, adjacent, linear, equation, graph, order, root, absolute, alpha, beta, gamma, log, percentage, percent, coefficient, average, average,

median, mean, mean, null, degree, scale, datum, dataset, statistic, statistical, population, cohort, longitudinal, control, range, range, spectrum, intermediate, extent, cover, index, proportion, rate, ratio, gradient, variable, parameter, multivariate, variance, differential, regression, association, deviation, algorithm, gamma, prediction, predictor, odds, random, randomize, distribution, distribute, normal, abnormal, abnormality, exception, likelihood, confidence, power, robust, strong, bias, error, raw

### **L'interprétation**

comparison, compare, comparable, comparative, baseline, reference, standard, standard, standardized, correlate, correlate, correlation, connect, associate, link, link, relate, related, relation, relationship, true, exact, false, defect, noise, correspond, corresponding, agree, match, match, pair, same, identity, identical, equal, equivalent, similar, similarity, share, other, contrast, distinguish, differ, difference, different, differential, distinct, variability, variable, variation, vary, variant, version, variety, various, diverse, diversity, heterogeneity, normalization, normalize, modification, modify, modified, adaptation, adapt, dynamic, regulate, adjustment, adjust, fit, modulation, modulate, correction, correct, correct, quality, qualitative, repeat, repeated, control, controlled, verify, check, feedback, ensure, explain, explanation, interpret, interpretation, discussion, discuss, conclude, discussion, comment, consequence, cause, cause, attribute, reason, account, origin, sense, outcome, result, positive, produce, production, yield, promote, generate, raise, inducible, induce, lead, result, resulting, open, arise, infer, consecutive, implicate, implication, imply, involve, involvement, role, play, show, indicate, indication, exhibit, highlight, reveal, responsible, influence, influence, impact, interact, interaction, rely, depend, dependent, dependence, relative, cross, indirect, active, act, performance, effective, effectiveness, efficacy, efficiency, efficient, validation, validate, valid, validity, support, support, accuracy, accurate, confirm, approve, prove, establish, evidence, reflect, evident, close, obvious, certain, clear, strong, weak, consistent, unclear, discovery, discover, finding, find



## Annexe 7 : l'interlexique : 1038 séries de mots analogues en anglais, espagnol, français, italien et portugais.

Nous n'écrivons pas dans cette annexe la définition du mot. Nous renvoyons pour cela à l'Annexe 2 : Les 2000 lemmes anglais du corpus ScienText traduits, p.298, mais le document complet, INTERLEXIQUE1038EN\_ROM.ods se trouvera sur le portail hal. L'astérisque (\*) avant un mot signifie qu'il n'est pas défini dans le dictionnaire généraliste consulté pour la langue en question.

A abdominal	abdominal	abdominal	addominale	abdominal	Lspé
A abnormal	anormal	anormal	anomalo	anormal	LST
A absent	ausente	absent	assente	ausente	LST
A absolute	absoluto	absolu	assoluto	absoluto	LST
N abundance	abundancia	abondance	abbondanza	abundância	LST
	académico /			académico /	
A academic	universitario	académique / universitaire	accademico / universitario	universitário	LST
V accept	aceptar	accepter	accettare	aceitar	LST
V accumulate	acumularse	s'accumuler	accumularsi	acumular-se	LST
N acid	ácido	acide	acido	ácido	Lspé
A acid	ácido	acide	acido	ácido	Lspé
V acquire	adquirir	acquérir	acquisire	adquirir	Lspé
V acquire	adquirir	acquérir	acquisire	adquirir	LST
N actin	actina	actine	actina	actina	Lspé
V activate	activar	activer	attivare	ativar	Lspé
A active	activo	actif	attivo	ativo	LST
N activity	actividad	activité	attività	atividade	Lspé
V adapt	adaptarse	s'adapter	adattarsi	adaptar-se	LST

A additional	adicional	additionnel	addizionale	adicional	LST
N adenocarcinoma	adenocarcinoma	adénocarcinome	adenocarcinoma	adenocarcinoma	Lspé
N adhesion	adhesión	adhésion	adesione	adesão	Lspé
A adjacent	adyacente	adjacent	adiacente	adjacente	LST
N adjustment	ajuste	ajustement	aggiustamento	ajuste	LST
V administer	administrar	administrer	amministrare	administrar	Lspé
N administration	administración	administration	somministrazione	administração	Lspé
N admission	admisión / entrada	admission / entrée	ammissione / entrata	admissão / entrada	Lspé
N adult	adulto	adulte	adulto	adulto	Lspé
A adult	adulto	adulte	adulto	adulto	Lspé
N affinity	afinidad	affinité	affinità	afinidade	Lspé
N affinity	afinidad	affinité	affinità	afinidade	Lspé
A African	africano	africain	africano	africano	LST
N agent	agente	agent	agente	agente	Lspé
N agonist	agonista	agoniste	agonista	agonista	Lspé
N albumin	albúmina	albumine	albumina	albumina	Lspé
N alcohol	alcohol	alcool	alcol / alcole	álcool	Lspé
N algorithm	algoritmo	algorithme	algoritmo	algoritmo	LST
V align	alinear	aligner	allineare	alinhar	Lspé
N alignment	alineamiento	alignement	allineamento	alinhamento	Lspé
N allele	alelo	allèle	allele	alelo	Lspé
N allergy	alergia	allergie	allergia	alergia	Lspé
N alpha	alfa	alpha	alfa	alfa	LST
V alter	alterar	altérer	alterare	alterar	Lspé
N alteration	alteración	altération	alterazione	alteração	Lspé
N alternative	alternativa	alternative	alternativa	alternativa	LST
A alternative	alternativo	alternatif	alternativo	alternativo	LST
A alveolar	alveolar	alvéolaire	alveolare	alveolar	Lspé
A alveolar	alveolar	alvéolaire	alveolare	alveolar	Lspé
N American	americano	Américain	americano	americano	LST

N amplification	amplificación	amplification	amplificazione	amplificação	Lspé
V amplify	amplificar	amplifier	amplificare	amplificar	Lspé
N analogue	análogo	analogue	analogo	análogo	Lspé
A analytic	analítico	analytique	analitico	analítico	LST
N androgen	andrógeno	androgène	androgeno	andrógeno	Lspé
N angiogenesis	angiogénesis	angiogénèse	angiogenesi	angiogênese	Lspé
N animal	animal	animal	animale	animal	Lspé
A animal	animal	animal	animale	animal	Lspé
V annotate	anotar	annoter	annotare	anotar	Lspé
A annual	anual	annuel	annuale	anual	LST
N antagonist	antagonista	antagoniste	antagonista	antagonista	Lspé
A anterior	anterior	antérieur	anteriore	anterior	Lspé
N antibiotic	antibiótico	antibiotique	antibiotico	antibiótico	Lspé
N antigen	antígeno	antigène	antigene	antígeno	Lspé
N antioxidant	antioxidante	antioxydant	antiossidante	antioxidante	Lspé
				* anti-sentido / sentido	
A antisense	* antisentido	* antisens	* antisenso	negativo	Lspé
N apoptosis	apoptosis	* apoptose	apoptosi	apoptose	Lspé
A apoptotic	* apoptótico	* apoptotique	* apoptotico	apoptótico	Lspé
A apparent	aparente	apparent	apparente	aparente	LST
A appropriate	apropiado	approprié	appropriato	apropiado	LST
A arterial	arterial	artériel	arteriale / arterioso	arterial	Lspé
N arthritis	artritis	arthrite	artrite	artrite	Lspé
N aspect	aspecto	aspect	aspetto	aspecto	LST
N aspirin	aspirina	aspirine	aspirina	aspirina	Lspé
V assist	asistir	assister	assistere	assistir	Lspé
N assistance	asistencia	assistance	assistenza	assistência	Lspé
V associate	asociar	associer	associare	asociar	LST
N association	asociación	association	associazione	associação	LST
N association	asociación	association	associazione	associação	LST

N asthma	asma	asthme	asma	asma	Lspé
N atherosclerosis	aterosclerosis	athérosclérose	aterosclerosi	aterosclerose	Lspé
N author	autor	auteur	autore	autor	LST
N autoantibody	* autoanticuerpo	* autoanticorps	autoanticorpo	* autoanticorpo	Lspé
A autoimmune	* autoinmune	* auto-immun	autoimmune	autoimune	Lspé
A automated	automatizado	automatisé	automatizzato	automatizado	LST
N bacillus	bacilo	bacille	bacillo	bacilo	Lspé
A bacterial	bacteriano	bactérien	batterico	bacteriano	Lspé
N band	banda	bande	banda	banda	Lspé
N barrier	barrera	barrière	barriera	barreira	Lspé
N barrier	barrera	barrière	barriera	barreira	LST
A basal	basal	basal	basale / prossimale	basal	Lspé
N base	base	base	base	base	Lspé
N base	base	base	base	base	LST
V base	basar / fundar	baser / fonder	basare / fondare	basear / fundar	LST
A basic	básico	basique	basico	básico	LST
N basis	base	base	base	base	LST
N beta	beta	bêta	beta	beta	LST
A biological	biológico	biologique	biologico	biológico	Lspé
N biology	biología	biologie	biologia	biologia	Lspé
N biopsy	biopsia	biopsie	biopsia	biopsia	Lspé
N biosynthesis	biosíntesis	biosynthèse	biosintesi	biossíntese	Lspé
N BLAST	BLAST	BLAST	BLAST	BLAST	Lspé
N block	bloque	bloc	blocco	bloco	Lspé
N block	bloque	bloc	blocco	bloco	LST
A bovine	bovino	bovin	bovino	bovino	Lspé
Caenorhabditis					
N elegans	Caenorhabditis elegans	Caenorhabditis elegans	Caenorhabditis elegans	Caenorhabditis elegans	Lspé
V calculate	calcular	calculer	calcolare	calcular	LST
N calculation	calculación / cálculo	calcul	calcolo	calcução / cálculo	LST

N cAMP	AMPc / cAMP	AMPc	cAMP	cAMP	Lspé
N cancer	cáncer	cancer	cancro	câncer	Lspé
N candidate	candidato	candidat	candidato	candidato	Lspé
A capable	capaz	capable	capace	capaz	LST
N capacity	capacidad	capacité	capacità	capacidade	LST
V capture	capturar	capturer	catturare	capturar	LST
N carbon	carbono	carbone	carbonio	carbono	Lspé
N carcinoma	carcinoma	carcinome	carcinoma	carcinoma	Lspé
A cardiac	cardíaco	cardiaque	cardiaco	cardíaco	Lspé
A cardiovascular	cardiovascular	cardiovasculaire	cardiovascolare	cardiovascular	Lspé
N cartilage	cartílago	cartilage	cartilagine	cartilagem	Lspé
A catalytic	catalítico	catalytique	catalitico	catalítico	Lspé
N category	categoría	catégorie	categoria	categoria	LST
N catheter	catéter	catéter	catetere	cateter	Lspé
N cause	causa	cause	causa	causa	LST
V cause	causar	causer	causare	causar	LST
N cell	célula	cellule	cellula	célula	Lspé
A cellular	celular	cellulaire	cellulare	celular	Lspé
N center	centro	centre	centro	centro	LST
N center	centro	centre	centro	centro	LST
A central	central	central	centrale	central	LST
A central	central	central	centrale	central	LST
A central	central	central	centrale	central	Lspé
N centrifugation	centrifugación	centrifugation	centrifugazione	centrifugação	Lspé
A cerebral	cerebral	cérébral	cerebrale	cerebral	Lspé
A cervical	cervical	cervical	cervicale	cervical	Lspé
N characteristic	característica	caractéristique	caratteristica	característica	LST
A characteristic	característico	caractéristique	caratteristico	característico	LST
V characterize	caracterizar	caractériser	caratterizzare	caracterizar	LST
N cholesterol	colesterol	cholestérol	colesterolo	colesterol	Lspé

N chromatin	cromatina	chromatine	cromatina	cromatina cromossomo / * cromossoma /	Lspé
N chromosome	cromosoma	chromosome	cromosoma	cromosoma	Lspé
A chronic	crónico	chronique	cronico	crônico	Lspé
N cigarette	cigarro	cigarette	sigaretta	cigarro	Lspé
V circulate	circular	circuler	circolare	circular	Lspé
N circulation	circulación	circulation	circolazione	circulação	Lspé
N class	clase	classe	classe	clase	LST
N classification	clasificación	classification	classificazione	classificação	LST
V classify	clasificar	classifier / classer	classificare	classificar	LST
N cleavage	* clivaje	clivage	clivaggio	clivagem	Lspé
N clinic	clínica	clinique	clinica	clínica	Lspé
A clinical	clínico	clinique	clinico	clínico	Lspé
N clinician	clínico	clinicien	clinico	clínico	Lspé
N clone	clon	clone	clone	clone	Lspé
V clone	clonar	cloner	clonare	clonar	Lspé
N cm	cm	cm	cm	cm	LST
V code	codificar	coder	codificare	codificar	Lspé
N codon	codón	codon	codone	códon	Lspé
N coefficient	coeficiente	coefficient	coeficiente	coeficiente	LST
A cognitive	cognitivo	cognitif	cognitivo	cognitivo	Lspé
N cohort	cohorte	cohorte	coorte	coorte	LST
N collagen	colágeno	collagène	collageno	colágeno	Lspé
N colleague	colega	collègue	collega	colega	LST
N colon	colon	côlon	colon	cólon / colo	Lspé
A colorectal	colorectal	colorectal	* coloretale	colorretal	Lspé
N combination	combinación	combinaison	combinazione	combinação	LST
V combine	combinar	combiner	combinare	combinar	LST
A combined	combinado	combiné	combinato	combinado	LST

N comment	comentario	commentaire	commento	comentário	LST
N committee	comité	comité	comitato	comitê	LST
N community	comunidad	communauté	comunità	comunidade	LST
A comparable	comparable	comparable	comparabile	comparável	LST
A comparative	comparativo	comparatif	comparativo	comparativo	LST
V compare	comparar	comparer	comparare	comparar	LST
N comparison	comparación	comparaison	comparazione	comparação	LST
N compartment	compartimento	compartiment / loge	compartimento	compartimento	Lspé
A complementary	complementario	complémentaire	complementare	complementar	Lspé
V complete	completar	compléter	completare	completar	LST
A complete	completo	complet	completo	completo	LST
N complex	complejo	complexe	complesso	complexo	Lspé
A complex	complejo	complexe	complesso	complexo	LST
N complexity	complejidad	complexité	complessità	complexidade	LST
N complication	complicación	complication	complicazione	complicação	Lspé
N component	componente	composant	componente	componente	LST
V compose	componer	composer	comporre	compor	LST
N composition	composición	composition	composizione	composição	LST
V comprise	comprender	comprendre	comprendere	compreender	LST
A computational	computacional	computationnel	computazionale	computacional	LST
V conceive	concebir	concevoir	concepire	conceber	LST
N concentration	concentración	concentration	concentrazione	concentração	Lspé
N concept	concepto	concept	concetto	conceito	LST
V concern	concernir / pertenecer	concerner	concernere	concernir / pertencer	LST
V conclude	concluir	conclure	concludere	concluir	LST
N conclusion	conclusión	conclusion	conclusione	conclusão	LST
	conducir / efectuar /	conduire / effectuer / réaliser /		conduzir / efetuar /	
V conduct	realizar	mener	condurre / effettuare / svolgere	realizar	LST
V confirm	confirmar	confirmer	confermare	confirmar	LST
N conformation	conformación	conformation	conformazione	conformação	Lspé

V conjugate	conjugar	conjuguer	coniugare	conjugar	Lspé
V connect	conectar	connecter	connettere	conetar	LST
A consecutive	consecutivo	consécutif	conseguente	consecutivo	LST
N consensus	consenso	consensus	consenso	consenso	LST
N consent	consenso	consentement	consenso	consentimento	Lspé
N consequence	consecuencia	conséquence	conseguenza	consequência	LST
N conservation	conservación	conservation	conservazione	conservação	Lspé
V consider	considerar	considérer	considerare	considerar	LST
A considerable	considerable	considérable	considerabile / considerevole	considerável	LST
N consideration	consideración	considération	considerazione	consideração	LST
N consideration	consideración	considération	considerazione	consideração	LST
V consist	consistir	consister	consistere	consistir	LST
A consistent	consistente	consistant	consistente	consistente	LST
A constant	constante	constant	costante	constante	LST
V constitute	constituir	constituer	costituire	constituire	LST
N constraint	constricción	contrainte	costrizione	constricção	LST
N consumption	consumo	consommation	consumo	consumo	Lspé
N contact	contacto	contact	contatto	contato	Lspé
N content	contenido	contenu	contenuto	conteúdo	LST
N content	contenido	contenu	contenuto	conteúdo	LST
N context	contexto	contexte	contesto	contexto	LST
V continue	continuar	continuer	continuare	continuar	LST
A continuous	continuo	continuel	continuo	contínuo	LST
N contrast	contraste	contraste	contrasto	contraste	LST
N contribution	contribución	contribution	contribuzione	contribuição	LST
N control	control	contrôle	controllo	controle	LST
V control	controlar	contrôler	controllare	controlar	LST
A conventional	convencional	conventionnel	convenzionale	convencional	LST
V convert	convertir	convertir	convertire	converter	Lspé
V coordinate	coordinar	coordonner	coordinare	coordenar	LST



	cuern / cordn /				
N cord	médula	corde / cordon / moelle	corda / cordone / midollo	corda / cordão / medula	Lspé
A coronary	coronario	coronaire	coronario	coronário	Lspé
A correct	correcto	correct	corretto	correto	LST
N correlate	correlato	corrélat	correlato	correlato	LST
N correlation	correlación	corrélacion	correlazione	correlação	LST
V correspond	corresponder	correspondre	corrispondere	corresponder	LST
A corresponding	correspondiente	correspondant	corrispondente	correspondente	LST
N cortex	córtex / corteza	cortex	corteccia	córtex / córtice	Lspé
N cost	coste / costo	cout	costo	custo	LST
A critical	crítico	critique	critico	crítico	LST
A critical	crítico	critique	critico	crítico	Lspé
A crucial	crucial	crucial	cruciale	crucial	LST
N culture	cultivo	culture	coltura	cultura	Lspé
N curve	curva	courbe	curva	curva	LST
N cyclin	* ciclina	* cycline	ciclina	ciclina	Lspé
N cytochrome	citocromo	cytochrome	citocromo	citocromo	Lspé
N cytoplasm	citoplasma	cytoplasme	citoplasma	citoplasma	Lspé
		cytoplasmique /			
A cytoplasmic	citoplasmático	cytoplasmatique	citoplasmatico	citoplasmático	Lspé
A cytotoxic	* citotóxico	cytotoxique	citotossico	citotóxico	Lspé
N decade	década / decenio	décade / décennie	decade / decennio	década / decênio	LST
N decision	decisión	décision	decisione	decisão	LST
V declare	declarar	déclarer	dichiarare	declarar	LST
N decline	declive	déclin	declino	declínio	LST
N decline	declive	déclin	declino	declínio	Lspé
A deficient	deficiente	déficient	deficiente	deficiente	Lspé
N definition	definición	définition	definizione	definição	LST
N degradation	degradación	dégradation	degradazione	degradação	Lspé
V demonstrate	demonstrar	démontrer	dimostrare	demonstrar	LST

A dendritic	dendrítico	dendritique	dendritico	dendrítico	Lspé
N density	densidad	densité	densità	densidade	LST
N department	departamento	département	dipartimento	departamento	LST
V depend	dependeer	dépendre	dipendere	dependeer	LST
N dependence	dependencia	dépendance	dipendenza	dependência	LST
N dependence	dependencia	dépendance	dipendenza	dependência	Lspé
A dependent	dependiente	dépendant	dipendente	dependente	LST
N depression	depresión	dépression	depressione	depressão	Lspé
N depression	depresión	dépression	depressione	depressão	Lspé
V derive	derivar	dériver	derivare	derivar	Lspé
V describe	describir	décrire	descrivere	descrever	LST
N detail	detalle / pormenor	détail	dettaglio	detalhe / pormenor	LST
N determinant	determinante	déterminant	determinante	determinante	Lspé
N determination	determinación	détermination	determinazione	determinação	LST
V determine	determinar	déterminer	determinare	determinar	LST
N diabetes	diabetes	diabète	diabete	diabetes	Lspé
A diabetic	diabético	diabétique	diabetico	diabético	Lspé
V diagnose	diagnosticar	diagnostiquer	diagnosticare	diagnosticar	Lspé
N diagnosis	diagnóstico	diagnostic	diagnosi	diagnóstico	Lspé
A diagnostic	diagnóstico	diagnostique	diagnostico	diagnóstico	Lspé
N diameter	diámetro	diamètre	diametro	diâmetro	LST
N diet	dieta	diète	dieta	dieta	Lspé
V differ	diferir	différer	differire	diferir	LST
N difference	diferencia	différence	differenza	diferença	LST
A different	diferente	différent	differente	diferente	LST
A differential	diferencial	différentiel	differenziale	diferencial	LST
A differential	diferencial	différentiel	differenziale	diferencial	LST
N difficulty	dificultad	difficulté	difficoltà	dificuldade	LST
N difficulty	dificultad	difficulté	difficoltà	dificuldade	LST
N dilution	dilución	dilution	diluizione	diluição	Lspé

N dimension	dimensión	dimension	dimensione	dimensão	LST
N dimension	dimensión	dimension	dimensione	dimensão	LST
A direct	directo	direct	diretto	direto	LST
V discuss	discutir	discuter	discutere	discutir	LST
N discussion	discusión	discussion	discussione	discussão	LST
N disorder	desorden / trastorno	désordre / trouble	disordine / disturbo	desordem / transtorno	Lspé
A distal	distal	distal	distale	distal	Lspé
N distance	distancia	distance	distanza	distância	LST
N distance	distancia	distance	distanza	distância	LST
A distinct	distinto	distinct	distinto	distinto	LST
V distinguish	distinguir	distinguer	distinguere	distinguir	LST
V distribute	distribuir / repartir	distribuer / répartir	distribuire / ripartire	distribuir / repartir	LST
N distribution	distribución	distribution	distribuzione	distribuição	LST
N distribution	distribución	distribution	distribuzione	distribuição	LST
N district	distrito	district	distretto	distrito	LST
N divergence	divergencia	divergence	divergenza	divergência	Lspé
A diverse	diverso	divers	diverso	diverso	LST
N diversity	diversidad	diversité	diversità	diversidade	LST
V divide	dividir	diviser	dividere	dividir	LST
V divide	dividir	diviser	dividere	dividir	LST
N division	división	division	divisione	divisão	Lspé
V document	documentar	documenter	documentare	documentar	LST
A dominant	dominante	dominant	dominante	dominante	LST
A dominant	dominante	dominant	dominante	dominante	Lspé
N dopamine	dopamina	dopamine	dopamina	dopamina	Lspé
N dose	dosis	dose	dose	dose	Lspé
N drug	droga	drogue	droga	droga	Lspé
N duplication	duplicación	duplication	duplicazione	duplicação	Lspé
A dynamic	dinámico	dynamique	dinamico	dinâmico	LST
N dynamics	dinámica	dynamique	dinamica	dinâmica	Lspé

A economic	económico	économique	economico	econômico	LST
V edit	editar	éditer	editare	editar	LST
N effect	efecto	effet	effetto	efeito	Lspé
N effector	efector	effecteur	effettore	efetor	Lspé
N efficiency	eficiencia	efficience	efficienza	eficiência	LST
A efficient	eficiente	efficient	efficiente	eficiente	LST
N electron	electrón	électron	elettrone	elétron	Lspé
N element	elemento	élément	elemento	elemento	LST
V elicit	elicitar / provocar	éliciter / provoquer	elicitare / provocare	elicitar / provocar	LST
V eliminate	eliminar	éliminer	eliminare	eliminar	Lspé
A empirical	empírico	empirique	empirico	empírico	LST
A endometrial	endometrial	* endométrial	endometriale	endometrial	Lspé
A endothelial	endotelial	endothélial	endoteliale	endotelial	Lspé
V enter	entrar	entrer	entrare	entrar	LST
N entry	entrada	entrée	entrata	entrada	LST
A epidemiological	epidemiológico	épidémiologique	epidemiologico	epidemiológico	Lspé
N epidemiology	epidemiología	épidémiologie	epidemiologia	epidemiologia	Lspé
N episode	episodio	épisode	episodio	episódio	Lspé
A epithelial	epitelial	épithélial	epiteliale	epitelial	Lspé
N epitope	* epítopo	* épitope	epitopo	epítopo	Lspé
A equivalent	equivalente	équivalent	equivalente	equivalente	LST
N error	error	erreur	errore	erro	Lspé
N error	error	erreur	errore	erro	LST
N Escherichia coli	Escherichia coli	Escherichia coli	Escherichia coli	Escherichia coli	Lspé
A essential	esencial	essentiel	essenziale	essencial	LST
N EST	EST	EST	EST	EST	Lspé
N estimate	estimación	estimation	stima	estimação	LST
V estimate	estimar	estimer	stimare	estimar	LST
N estimation	estimación	estimation	stima	estimação	LST
N estimation	estimación	estimation	stima	estimação	LST

N estrogen	estrógeno	œstrogène	estrogeno	estrogênio / estrógeno	Lspé
N ethanol	etanol	éthanol	etanolo	etanol	Lspé
A eukaryotic	eucariótico	* eucaryote	eucariotico	eucariótico	Lspé
A European	européo	européen	européo	européu	LST
A evident	evidente	évident	evidente	evidente	LST
A exact	exacto	exact	esatto	exato	LST
V examine	examinar	examiner	esaminare	examinar	Lspé
V examine	examinar	examiner	esaminare	examinar	LST
V exclude	excluir	exclure	escludere	excluir	LST
N exclusion	exclusión	exclusion	esclusione	exclusão	LST
V exist	existir	exister	esistere	existir	LST
N existence	existencia	existence	esistenza	existência	LST
N exon	* exón	* exon	esone	éxon / exão	Lspé
N expansion	expansión	expansion	espanzione	expansão	LST
N experience	experiencia	expérience	esperienza	experiência	Lspé
N experiment	experimento	expérience	esperimento	experimento	LST
A experimental	experimental	expérimental	sperimentale	experimental	LST
N expert	experto / perito	expert	esperto / perito	expert / experto / perito	LST
V explore	explorar	explorer	esplorare	explorar	LST
V expose	exponer	exposer	esporre	expor	Lspé
N expression	expresión	expression	espressione	expressão	Lspé
N extension	extensión	extension	estensione	extensão	Lspé
A external	externo	externe	esterno	externo	Lspé
A external	externo	externe	esterno	externo	LST
A extracellular	* extracelular	extracellulaire	extracellulare	extracelular	Lspé
N extract	extracto	extrait	estratto	extrato	Lspé
N extract	extracto	extrait	estratto	extrato	Lspé
V facilitate	facilitar	faciliter	facilitare	facilitar	LST
N factor	factor	facteur	fattore	fator	LST
A fetal	fetal	foetal	fetale	fetal	Lspé

N fiber	fibra	fibre	fibra	fibra	Lspé
N fibroblast	fibroblasto	fibroblaste	fibroblasto	fibroblasto	Lspé
N fibrosis	fibrosis	fibrose	fibrosi	fibrose	Lspé
N fig	Fig.	fig.	fig.	fig.	LST
N figure	figura	figure	figura	figura	LST
N file	fichero / archivo	fichier	file	ficheiro / arquivo	LST
N filter	filtro	filtre	filtro	filtro	Lspé
A final	final	final	finale	final	LST
A financial	financiero	financier	finanziario	financeiro	LST
N fluid	fluido	fluide	fluido	fluido	Lspé
A fluid	fluido	fluide	fluido	fluido	Lspé
N fluorescence	fluorescencia	fluorescence	fluorescenza	fluorescência	Lspé
A fluorescent	fluorescente	fluorescent	fluorescente	fluorescente	Lspé
N follicle	folículo	follicule	follicolo	folículo	Lspé
N form	forma	forme	forma	forma	LST
V form	formar	former	formare	formar	LST
N format	formato	format	format	formato	LST
N fracture	fractura	fracture	frattura	fratura	Lspé
N fragment	fragmento	fragment	frammento	fragmento	Lspé
N frequency	frecuencia	fréquence	frequenza	frequência	LST
A frequent	frecuente	fréquent	frequente	frequente	LST
V function	funcionar	fonctionner	funzionare	funcionar	LST
A functional	funcional	fonctionnel	funzionante	funcional	Lspé
A future	futuro	futur	futuro	futuro	LST
N gamma	gama	gamma	gamma	gama	LST
N gamma	gama	gamma	gamma	gama	LST
N gas	gas	gaz	gas	gás	Lspé
A gastric	gástrico	gastrique	gastrico	gástrico	Lspé
A gastrointestinal	gastrointestinal	gastrointestinal	gastrointestinale	gastrointestinal	Lspé
N gel	gel	gel	gel	gel	Lspé

A general	general	général	generale	geral	Lspé
A general	general	général	generale	geral	LST
A genetic	genético	génétique	genetico	genético	Lspé
N genome	genoma	génome	genoma	genoma	Lspé
A genomic	genómico	génomique	genomico	genômico	Lspé
N genomics	genómica	génomique	genomica	* genômica	Lspé
N genotype	genotipo	génotype	genotipo	genótipo	Lspé
N gland	glándula	glande	ghiandola	glândula	Lspé
A global	global	global	globale	global	LST
			glucoso / glicosio / glucosio /		
N glucose	glucosa	glucose	glicoso	glucose / glicose	Lspé
N glutamate	glutamato	glutamate	glutamato / glutammato	glutamato	Lspé
N glycoprotein	glicoproteína	glycoprotéine	glicoproteina	glicoproteína	Lspé
N grade	grado	grade	grado	grau	Lspé
N gradient	gradiente	gradient	gradiente	gradiente	LST
N group	grupo	groupe	gruppo	grupo	LST
N haplotype	* haplotipo	haplotype	* aplotipo	haplótipo	Lspé
N heparin	heparina	héparine	eparina	heparina	Lspé
A hepatic	hepático	hépatique	epatico	hepático	Lspé
N hepatitis	hepatitis	hépatite	epatite	hepatite	Lspé
A histological	histológico	histologique	istologico	histológico	Lspé
N histone	histona	histone	istone	histona	Lspé
N history	historia	histoire	storia	história	Lspé
N history	historia	histoire	storia	história	LST
A homologous	homólogo	homologue	omologo	homólogo	Lspé
N homologue	homólogo	homologue	omologo	homólogo	Lspé
N homology	homología	homologie	omologia	homologia	Lspé
N homology	homología	homologie	omologia	homologia	Lspé
N hormone	hormona	hormone	ormone	hormônio	Lspé
N human	humano	humain	umano	humano	Lspé

A human	humano	humain	umano	humano	Lspé
A hybrid	híbrido	hybride	ibrido	híbrido	Lspé
N hypertension	hipertensión	hypertension	ipertensione	hipertensão	Lspé
N hypothesis	hipótesis	hypothèse	ipotesi	hipótese	LST
N idea	idea	idée	idea	ideia	LST
A identical	idéntico	identique	identico	idêntico	LST
N identification	identificación	identification	identificazione	identificação	LST
V identify	identificar	identifier	identificare	identificar	LST
N identity	identidad	identité	identità	identidade	LST
V illustrate	ilustrar	illustrer	illustrare	ilustrar	LST
N image	imagen	image	immagine	imagem	LST
A immune	inmune	immun	immune	imune	Lspé
N immunity	inmunidad	immunité	immunità	imunidade	Lspé
N immunodeficiency	inmunodeficiencia	immunodéficience	immunodeficienza	imunodeficiência	Lspé
N immunoglobulin	inmunoglobulina	immunoglobuline	immunoglobulina	imunoglobulina	Lspé
N impact	impacto	impact	impatto	impacto	LST
V implement	implementar	implémenter / mettre en œuvre implémentation / mise en	implementare	implementar	LST
N implementation	implementación	œuvre	implementazione	implementação	LST
V implicate	implicar	impliquer	implicare	implicar	LST
N implication	implicación	implication	implicazione	implicação	LST
V imply	implicar	impliquer	implicare	implicar	LST
N importance	importancia	importance	importanza	importância	LST
A important	importante	important	importante	importante	LST
A in vitro	In vitro	in vitro	in vitro	in vitro	Lspé
A in vivo	in vivo	in vivo	in vivo	in vivo	Lspé
N incidence	incidencia	incidence	incidenza	incidência	Lspé
V include	incluir	inclure	includere	incluir	LST
N inclusion	inclusión	inclusion	inclusione	inclusão	LST
N inclusion	inclusión	inclusion	inclusione	inclusão	Lspé



V incorporate	incorporar	incorporer	incorporare	incorporar	LST
V incubate	incubar	incuber	incubare	incubar	Lspé
N incubation	incubación	incubation	incubazione	incubação	Lspé
A independent	independiente	indépendant	indipendente	independente	LST
V indicate	indicar	indiquer	indicare	indicar	LST
N indication	indicación	indication	indicazione	indicação	LST
N indicator	indicador	indicateur	indicatore	indicador	Lspé
A indirect	indirecto	indirect	indiretto	indireto	LST
N individual	individuo	individu	individuo	indivíduo	LST
A individual	individual	individuel	individuale	individual	LST
V induce	inducir	induire	indurre	induzir	LST
N infarction	infarto	infarctus	infarto	infarto / enfarte	Lspé
V infect	infectar	infecter	infettare	infectar	Lspé
V infer	inferir	inférer	inferire	inferir	LST
A inflammatory	inflamatorio	inflammatoire	infiammatorio	inflamatório	Lspé
N influence	influencia	influence	influenza	influência	LST
V influence	influir	influencer / influer	influenzare	influenciar / influir	LST
V inform	informar	informer	informare	informar	LST
N information	información	information	informazione	informação	LST
V inhibit	inhibir	inhiber	inibire	inibir	Lspé
N inhibitor	inhibidor	inhibiteur	inibitore	inibidor	Lspé
A initial	inicial	initial	iniziale	inicial	LST
V initiate	iniciar	initier / amorcer	iniziare	iniciar	LST
V inject	inyectar	injecter	iniettare	injetar	Lspé
N insect	insecto	insecte	insetto	inseto	Lspé
N institute	instituto	institut	istituto	instituto	LST
N institute	instituto	institut	istituto	instituto	LST
N institution	institución	institution	istituzione	instituição	LST
N instrument	instrumento	instrument	strumento	instrumento	LST
N insulin	insulina	insuline	insulina	insulina	Lspé

A intact	intacto	intact	intatto	intato	Lspé
N integration	integración	intégration	integrazione	integração	LST
N integrin	* integrina	* intégrine	integrina	integrina	Lspé
N intensity	intensidad	intensité	intensità	intensidade	LST
A intensive	intensivo	intensif	intensivo	intensivo	Lspé
	interactuar /			* interaccionar /	
V interact	interaccionar	interagir	interagire	interagir	LST
N interface	interfaz	interface	interfaccia	interface	Lspé
N interface	interfaz	interface	interfaccia	interface	LST
N interferon	interferón	interféron	interferone	interferon	Lspé
N interleukin	interleucina	interleukine	interleuchina	interleucina	Lspé
A intermediate	intermedio	intermédiaire	intermedio	intermediário	LST
N intern	interno	interne	interno	interno	Lspé
A internal	interno	interne	interno	interno	Lspé
A international	internacional	international	internazionale	internacional	LST
N internet	internet	internet	internet	internet	<b>LST</b>
V interpret	interpretar	interpréter	interpretare	interpretar	LST
N interpretation	interpretación	interprétation	interpretazione	interpretação	LST
N interval	intervalo	intervalle	intervallo	intervalo	LST
N intervention	intervención	intervention	intervento	intervenção	Lspé
A intestinal	intestinal	intestinal	intestinale	intestinal	Lspé
A intracellular	intracelular	intracellulaire	intracellulare	intracelular	Lspé
	intravenoso /			intravenoso /	
A intravenous	endovenoso	intraveineux / * endoveineux	intravenoso / endovenoso	endovenoso	Lspé
V introduce	introducir	introduire	introdurre	introduzir	Lspé
N intron	* intrón	intron	introne	* íntron	Lspé
N invasion	invasión	invasion	invasione	invasão	Lspé
A invasive	invasivo	invasif	invasivo	invasivo	Lspé
V invest	invertir	investir	investire	investir	LST
V investigate	investigar	investiguer / étudier	investigare	investigar	LST

N investigation	investigación	investigation	investigazione	investigação	LST
N investigator	investigador	investigateur	investigatore	investigador	LST
N ion	ion/ión	ion	ione	ionte/íon/iônio	Lspé
N ischemia	isquemia	ischémie	ischemia	isquemia	Lspé
N isoform	* isoforma	* isoforme	* isoforma	* isoforma	Lspé
N item	ítem	item	item	item	LST
N kg	kg	kg	kg	kg	LST
N kilobyte	kilobyte	* kilooctet	kilobit	kilobit / quilobit	LST
N kinase	cinasa	kinase	cinasi / chinasi	cinase / quinase	Lspé
N kit	kit	kit	kit	kit	LST
N km	km	km	km	km	LST
N lab	Lab.	labo	lab	lab.	LST
N laboratory	laboratorio	laboratoire	laboratorio	laboratório	LST
N language	lenguaje	langage	linguaggio	linguagem	LST
N language	lengua	langue	lingua	língua	LST
N laser	láser	laser	laser	laser	Lspé
A lateral	lateral	latéral	laterale	lateral	Lspé
N limit	límite	limite	limite	limite	LST
V limit	limitar	limiter	limitare	limitar	LST
N limitation	limitación	limitation	limitazione	limitação	LST
A limited	limitado	limité	limitato	limitado	LST
N line	línea	lignée	linea	linhagem	Lspé
A linear	lineal	linéaire	lineare	linear	LST
A linear	lineal	linéaire	lineare	linear	LST
N lipid	lípidio	lipide	lipide	lipídio / lípide	Lspé
N lipoprotein	lipoproteína	lipoprotéine	lipoproteina	lipoproteína	Lspé
A liquid	líquido	liquide	liquido	líquido	Lspé
N list	lista	liste	lista	lista	LST
N literature	literatura	littérature	letteratura	literatura	LST
A local	local	local	locale	local	LST

A local	local	local	locale	local	Lspé
N localization	localización	localisation	localizzazione	localização	LST
N localization	localización	localisation	localizzazione	localização	LST
V localize	localizar	localiser	localizzare	localizar	LST
V locate	localizar	localiser	localizzare	localizar	LST
N locus	* locus	locus	locus	locus	Lspé
N log	log	log	log	log	LST
A logistic	logístico	logistique	logistico	logístico	LST
A longitudinal	longitudinal	longitudinal	longitudinale	longitudinal	LST
A longitudinal	longitudinal	longitudinal	longitudinale	longitudinal	Lspé
N macrophage	* macrófago	macrophage	macrofago	macrófago	Lspé
A magnetic	magnético	magnétique	magnetico	magnético	Lspé
A major	mayor	majeur	maggiore	maior	LST
A malignant	maligno	malin	maligno	maligno	Lspé
A manual	manual	manuel	manuale	manual	LST
N manuscript	manuscrito	manuscrit	manoscritto	manuscrito	LST
N mass	masa	masse	massa	massa	LST
N mass	masa	masse	massa	massa	Lspé
N material	material	matériel	materiale	material	LST
A maternal	maternal	maternel	materno	maternal	Lspé
N matrix	matriz	matrice	matrice	matriz	Lspé
N matrix	matriz	matrice	matrice	matriz	Lspé
N matrix	matriz	matrice	matrice	matriz	LST
N maturation	maduración	maturation	maturazione	maturação	Lspé
A maximum	máximo	maximum	massimo	máximo	LST
A mechanical	mecánico	mécanique	meccanico	mecânico	LST
N mechanism	mecanismo	mécanisme	meccanismo	mecanismo	LST
A median	mediano	médian	mediano	mediano	LST
A median	mediano	médian	mediano	mediano	Lspé
V mediate	mediar	médiatiser	mediare	mediar	Lspé

N mediator	mediador	médiateur	mediatore	mediador	Lspé
N medication	medicación	médication	medicazione	medicação	Lspé
N melanoma	melanoma	mélanome	melanoma	melanoma	Lspé
N member	miembro	membre	membro	membro	LST
N membrane	membrana	membrane	membrana	membrana	Lspé
N memory	memoria	mémoire	memoria	memória	Lspé
N memory	memoria	mémoire	memoria	memória	LST
A mental	mental	mental	mentale	mental	Lspé
V mention	mencionar	mentionner	menzionare	mencionar	LST
A metabolic	metabólico	métabolique	metabolico	metabólico	Lspé
N metabolism	metabolismo	métabolisme	metabolismo	metabolismo	Lspé
N metabolite	metabolito	métabolite	metabolito / metabolita	metabólito	Lspé
N metastasis	metástasis	métastase	metastasi	metástase	Lspé
A metastatic	metastásico	métastatique	metastatico	metastático	Lspé
N methodology	metodología	méthodologie	metodologia	metodologia	LST
N mg	mg * microarreglo / *	mg	mg	mg microarreglo / *microséries	LST
N microarray	micromatriz	* microréseau / * micromatrice	microarray / micromatrice	*microséries	Lspé
N microscope	microscopio microscopia /	microscope	microscopio	microscópio	Lspé
N microscopy	microscopía	microscopie	microscopia	microscopia	Lspé
N min	min	min./mn/m	min/m	min	LST
V minimize	minimizar	minimiser	minimizzare	minimizar	LST
A minimum	mínimo	minimum	minimo	mínimo	LST
A minor	menor	mineur	minore	menor	LST
N minute	minuto	minute	minuto	minuto	LST
A mitochondrial	* mitocondrial	mitochondrial	mitocondriale	mitocondrial	Lspé
N mitochondrion	mitocondria	mitochondrie	mitocondrio	mitocôndria	Lspé
N ml	ml	ml	ml/mL	ml	LST
N mm	mm	mm	mm	mm	LST

N mmol	mmol	mmol	mmol	mmol	Lspé
N mode	modo	mode	modo	modo	LST
N model	modelo	modèle	modello	modelo	LST
V model	modelizar	modéliser	modellare	modelar	LST
N modification	modificación	modification	modificazione	modificação	LST
A modified	* modificado	modifié	modificato	modificado	LST
V modify	modificar	modifier	modificare	modificar	LST
V modulate	modular	moduler	modulare	modular	LST
N modulation	modulación	modulation	modulazione	modulação	LST
A molecular	molecular	moléculaire	molecolare	molecular	Lspé
N molecule	molécula	molécule	molecola	molécula	Lspé
V monitor	monitorizar	* monitorer / surveiller	monitorare	monitorar / monitorizar	Lspé
		monitoring / monitoraggio /		monitoração /	
N monitoring	* monitoreo	surveillance	monitoraggio	monitorização	Lspé
A monoclonal	monoclonal	monoclonal	monoclonale	monoclonal	Lspé
A morphologic	morfológico	morphologique	morfologico	morfológico	Lspé
N morphology	morfología	morphologie	morfologia	morfologia	Lspé
N mortality	mortalidad	mortalité	mortalità	mortalidade	Lspé
N motif	motivo	motif	motivo	motivo	Lspé
A motor	motor	moteur	motore / motorio	motor	Lspé
N movement	movimiento	mouvement	movimento	movimento	Lspé
N ms	ms	ms	ms	ms	LST
A multiple	múltiple	multiple	multiplo	múltiplo	LST
	* multivariado / *			multivariado /	
A multivariate	multivariable	multivarié / multivariable	* multivariato	multivariável	LST
A murine	* murino	murin	murino	murino / murídeo	Lspé
N muscle	músculo	muscle	muscolo	músculo	Lspé
N mutant	mutante	mutant	mutante	mutante	Lspé
A mutant	mutante	mutant	mutante	mutante	Lspé
A national	nacional	national	nazionale	nacional	LST

A native	nativo	natif	nativo	nativo	Lspé
A natural	natural	naturel	naturale	natural	Lspé
N nature	naturaleza	nature	natura	natureza	Lspé
A necessary	necesario	nécessaire	necessario	necessário	LST
N necrosis	necrosis	nécrose	necrosi	necrose	Lspé
A negative	negativo	négatif	negativo	negativo	Lspé
A negative	negativo	négatif	negativo	negativo	LST
A neonatal	neonatal	néonatal	neonatale	neonatal	Lspé
N nerve	nervio	nerf	nervo	nervo	Lspé
A nervous	nervioso	nerveux	nervoso	nervoso	Lspé
A neural	neural	neural	neurale	neural	Lspé
N neuron	neurona	neurone	neurone	neurônio / neurone	Lspé
A neuronal	neuronal	neuronal	neuronale	neuronal	Lspé
A nitric	nítrico	nitrique	nitrico	nítrico	Lspé
N nm	nm	nm	nm	nm	Lspé
A normal	normal	normal	normale	normal	Lspé
A normal	normal	normal	normale	normal	LST
N normalization	normalización	normalisation	normalizzazione	normalização	LST
V normalize	normalizar	normaliser	normalizzare	normalizar / normatizar	LST
A nuclear	nuclear	nucléaire	nucleare	nuclear	Lspé
A nucleic	nucleico	nucléique	nucleico	nucleico	Lspé
N nucleotide	nucleótido	nucléotide	nucleotide	nucleotídeo	Lspé
A numerous	numeroso	nombreux	numeroso	numeroso	LST
N obesity	obesidad	obésité	obesità	obesidade	Lspé
N objective	objetivo	objectif	obiettivo	objetivo	LST
N observation	observación	observation	osservazione	observação	LST
V observe	observar	observer	osservare	observar	LST
V offer	ofrecer	offrir	offrire	oferecer	LST
N oligonucleotide	* oligonucleótido	oligonucléotide	oligonucleotide	oligonucleotídio / oligonucleótido	Lspé

N oncogene	oncogén	oncogène	oncogene	oncógene	Lspé
N ontology	ontología	ontologie	ontologia	ontologia	LST
N operon	* operón	opéron	operone	* operação	Lspé
N opportunity	oportunidad / ocasión	opportunité / occasion	opportunità / occasione	oportunidade / ocasião	LST
V optimize	optimizar	optimiser	ottimizzare	otimizar	LST
A oral	oral	oral	orale	oral	Lspé
N order	orden	ordre	ordine	ordem	LST
N organism	organismo	organisme	organismo	organismo	Lspé
N organization	organización	organisation	organizzazione	organização	LST
N organization	organización	organisation	organizzazione	organização	LST
N orientation	orientación	orientation	orientamento	orientação	LST
A original	original	original	originale	original	LST
N ortholog	ortólogo	* orthologue	* ortologo	* ortólogo	Lspé
				osteoartrite /	
N osteoarthritis	osteoartritis	ostéoartrite	osteoartrite / osteoartrosi	osteoartrose	Lspé
N overexpression	* sobreexpresión	* surexpression	sovraespressione	sobreexpressão	Lspé
A oxidative	* oxidativo	oxydatif	ossidativo	oxidativo	Lspé
A pancreatic	pancreático	pancréatique	pancreatico	pancreático	Lspé
A parallel	paralelo	parallèle	parallelo	paralelo	LST
N parameter	parámetro	paramètre	parametro	parâmetro	LST
N parasite	parásito	parasite	parassita	parasita	Lspé
N part	parte	partie	parte	parte	LST
A partial	parcial	partiel	parziale	parcial	LST
N participant	participante	participant	partecipante	participante	LST
V participate	participar	participer	partecipare	participar	LST
N participation	participación	participation	partecipazione	participação	LST
N particle	partícula	particule	particola	partícula	Lspé
A particular	particular	particulier	particolare	particular	LST
A past	pasado	passé	passato	passado	LST
N pathogenesis	patogenia	pathogénie	patogenesi	patogenia	Lspé



A pathological	patológico	pathologique	patologico	patológico	Lspé
N pathology	patología	pathologie	patologia	patologia	Lspé
N patient	paciente	patient	paziente	paciente	Lspé
N PC	PC	PC	PC	PC	LST
N PCR	PCR	PCR/ACP	PCR	PCR	Lspé
A pediatric	pediátrico	pédiatrique	pediatrico	pediátrico	Lspé
N peptide	péptido	peptide	peptide	péptido / peptídio	Lspé
V perceive	percibir	percevoir	percepire	perceber	LST
N percent	por ciento	pourcent / pour cent	per cento / per cento	por cento	LST
N percentage	porcentaje	pourcentage	percentuale	percentagem	LST
N perception	percepción	perception	percezione	percepção	Lspé
N perfusion	perfusión	perfusion	perfusione	perfusão	Lspé
N period	período / periodo	période	periodo	período	LST
V permit	permitir	permettre	permettere	permitir	LST
A personal	personal	personnel	personale	peçoal	LST
N perspective	perspectiva	perspective	prospettiva	perspectiva	LST
N pH	pH	pH	pH	pH	Lspé
N phage	fago	phage	fago	fago	Lspé
N phase	fase	phase	fase	fase	LST
A phenotypic	fenotípico	phénotypique	fenotipico	fenotípico	Lspé
N PI	PI	PI	PI	PI	Lspé
N placebo	placebo	placebo	placebo	placebo	Lspé
N plant	planta	plante	pianta	planta	Lspé
N plasma	plasma	plasma	plasma	plasma	Lspé
N plasmid	* plásmido	plasmide	plasmide / plasmidio	plasmídeo / plasmídio	Lspé
Plasmodium					
N falciparum	Plasmodium falciparum	Plasmodium falciparum	Plasmodium falciparum	Plasmodium falciparum	Lspé
N platform	plataforma	plateforme	piattaforma	plataforma	LST
N polymerase	polimerasa	polymérase	polimerasi	polimerase	Lspé
N polymorphism	polimorfismo	polymorphisme	polimorfismo	polimorfismo	Lspé

N pool	* pool	pool / masse commune	pool	pool	LST
N positive	positivo	positif	positivo	positivo	LST
A positive	positivo	positif	positivo	positivo	LST
A positive	positivo	positif	positivo	positivo	LST
A positive	positivo	positif	positivo	positivo	Lspé
N possibility	posibilidad	possibilité	possibilità	possibilidade	LST
A posterior	posterior	postérieur	posteriore	posterior	LST
A posterior	posterior	postérieur	posteriore	posterior	Lspé
N potential	potencial	potentiel	potenziale	potencial	LST
A potential	potencial	potentiel	potenziale	potencial	LST
A practical	práctico	pratique	pratico	prático	LST
N precursor	precursor	précurseur	precursore	precursor	Lspé
A predictive	predictivo	prédicatif	predittivo	preditivo	Lspé
N predictor	predictor	* prédicteur	* predittore	predictor	LST
N preference	preferencia	préférence	preferenza	preferência	LST
A preliminary	preliminar	préliminaire	preliminare	preliminar	LST
N preparation	preparación	préparation	preparazione	preparação	LST
N preparation	preparación	préparation	preparazione	preparação	Lspé
V prepare	preparar	préparer	preparare	preparar	Lspé
V prepare	preparar	préparer	preparare	preparar	LST
V prescribe	prescribir / recetar	prescrire / ordonner	prescrivere / ordinare	prescrever / receitar	Lspé
N presence	presencia	présence	presenza	presença	LST
V present	presentar	présenter	presentare	apresentar	LST
V present	presentar	présenter	presentare	apresentar	Lspé
A present	presente	présent	presente	presente	LST
N presentation	presentación	présentation	presentazione	apresentação	Lspé
N presentation	presentación	présentation	presentazione	apresentação	LST
N prevalence	prevalencia	prévalence	prevalenza	prevalência	Lspé
V prevent	prevenir / impedir	prévenir / empêcher	prevenire / impedire	prevenir / impedir	Lspé
N prevention	prevención	prévention	prevenzione	prevenção	Lspé

A primary	primario	primaire	primario	primário	Lspé
A primary	primario	primaire	primario	primário	Lspé
N primate	primate	primate	primate	primata	Lspé
A prime	primero	premier	primo	primeiro	LST
A principal	principal	principal	principale	principal	LST
N priority	prioridad	priorité	priorità	prioridade	LST
N probability	probabilidad	probabilité	probabilità	probabilidade	LST
N problem	problema	problème	problema	problema	LST
N procedure	procedimiento	procédure	procedura / procedimento	procedimento	LST
N process	proceso	processus	processo	processo	LST
N process	proceso	processus	processo	processo	Lspé
V produce	producir	produire	produrre	produzir	LST
N product	producto	produit	prodotto	produto	Lspé
N product	producto	produit	prodotto	produto	LST
N professional	profesional	professionnel	professionista	profissional	Lspé
A professional	profesional	professionnel	professionale	profissional	Lspé
N profile	perfil	profile	profilo	perfil	LST
N progesterone	* progesterona	progestérone	progesterone	progesterona	Lspé
A prognostic	pronóstico	pronostique	prognostico	pronóstico	Lspé
N program	programa	programme	programma	programa	LST
N program	programa	programme	programma	programa	LST
N progress	progreso	progrès	progresso	progresso	LST
N progression	progresión	progression	progressione	progressão	LST
A progressive	progresivo	progressif	progressivo	progressivo	LST
N project	proyecto	projet	progetto	projeto	LST
N proliferation	proliferación	prolifération	proliferazione	proliferação	Lspé
V promote	promover	promouvoir	promuovere	promover	LST
V promote	promover	promouvoir	promuovere	promover	Lspé
N promoter	promotor	promoteur	promotore	promotor	Lspé
N property	propiedad	propriété	proprietà	propriedade	LST

N	proportion	proporción	proportion	proporzione	proporção	LST
V	propose	proponer	proposer	proporre	propor	LST
A	prospective	prospectivo	prospectif	prospettivo	prospectivo	Lspé
N	prostaglandin	* prostaglandina	prostaglandine	prostaglandina	prostaglandina	Lspé
N	prostate	próstata	prostate	prostata	próstata	Lspé
N	protease	proteasa	protéase	proteasi	protease	Lspé
V	protect	proteger	protéger	proteggere	proteger	Lspé
A	protective	protector	protecteur	protettivo	protetivo / protetor	Lspé
N	protein	proteína	protéine	proteina	proteína	Lspé
N	protocol	protocolo	protocole	protocollo	protocolo	LST
A	psychological	psicológico	psychologique	psicologico	psicológico	Lspé
A	public	público	public	pubblico	público	LST
A	public	público	public	pubblico	público	LST
N	publication	publicación	publication	pubblicazione	publicação	LST
N	publication	publicación	publication	pubblicazione	publicação	LST
A	pulmonary	pulmonar	pulmonaire	polmonare	pulmonar	Lspé
N	purification	purificación	purification	purificazione	purificação	Lspé
V	purify	purificar	purifier	purificare	purificar	Lspé
A	qualitative	cualitativo	qualitatif	qualitativo	qualitativo	LST
N	quality	cualidad	qualité	qualità	qualidade	LST
N	quality	cualidad	qualité	qualità	qualidade	LST
N	quantification	cuantificación	quantification	quantificazione	quantificação	LST
V	quantify	cuantificar	quantifier	quantificare	quantificar	LST
A	quantitative	cuantitativo	quantitatif	quantitativo	quantitativo	LST
N	quantity	cantidad	quantité	quantità	quantidade	LST
N	question	cuestión	question	questione	questão	LST
V	randomize	* randomizar	* randomiser	randomizzare / casualizzare	acidentalizar	LST
A	rapid	rápido	rapide	rapido	rápido	LST
A	rare	raro	rare	raro	raro	LST

N rat	rata	rat	ratto	rato	Lspé
A reactive	reactivo	réactif	reattivo	reativo	Lspé
N reactivity	* reactividad	réactivité	reattività	reatividade	Lspé
A real	real	réel	reale	real	LST
V receive	recibir	recevoir	ricevere	receber	LST
A recent	reciente	récent	recente	recente	LST
N receptor	receptor	récepteur	recettore / ricettore	receptor	Lspé
N receptor	receptor	récepteur	recettore / ricettore	receptor	Lspé
A recombinant	recombinante	recombinant	ricombinante	recombinante	Lspé
N recombination	recombinación	recombinaison	ricombinazione	recombinação	Lspé
V recruit	reclutar	recruter	reclutare	recrutar	Lspé
N recruitment	reclutamiento	recrutement	reclutamento	recrutamento	Lspé
N recurrence	recurrencia	récurrence	ricorrenza	recorrência	Lspé
A recurrent	recurrente	récurrent	ricorrente	recorrente	Lspé
V reduce	reducir	réduire	ridurre	reduzir	LST
V reflect	reflejar	réfléter	riflettere	refletir	LST
N regimen	régimen	régime	regime	regime	Lspé
N region	región	région	regione	região	Lspé
A regional	regional	régional	regionale	regional	LST
N regression	regresión	régression	regressione	regressão	LST
N regression	regresión	régression	regressione	regressão	Lspé
A regular	regular	régulier	regolare	regular	LST
N regulation	regulación	régulation	regolazione	regulação	Lspé
N regulation	regulación	régulation	regolazione	regulação	Lspé
N regulator	regulador	régulateur	regolatore	regulador	Lspé
A regulatory	regulador	* réglementaire	regolatore	regulador	Lspé
A relative	relativo	relatif	relativo	relativo	LST
A renal	renal	rénal	renale	renal	Lspé
V replicate	replicar	répliquer	replicare	replicar	Lspé
N replication	replicación	réplication	replicazione	replicação	Lspé

N reporter	reportero	rapporteur	reporter	repórter	Lspé
V represent	representar	représenter	rappresentare	representar	LST
V represent	representar	représenter	rappresentare	representar	LST
N representation	representación	représentation	rappresentazione	representação	LST
N representative	representante	représentant	rappresentante	representante	LST
A representative	representativo	représentatif	rappresentativo	representativo	LST
A reproductive	reproductivo	reproductif	riproduttivo	reprodutivo	Lspé
N resident	residente	résident	residente	residente	Lspé
A residual	residual	résiduel	residuale	residual	Lspé
N residue	residuo	résidu	residuo	resíduo	Lspé
N resistance	resistencia	résistance	resistenza	resistência	Lspé
A resistant	resistente	résistant	resistente	resistente	Lspé
N resolution	resolución	résolution	risoluzione	resolução	LST
N resolution	resolución	résolution	risoluzione	resolução	Lspé
A respective	respectivo	respectif	rispettivo	respectivo	LST
A respiratory	respiratorio	respiratoire	respiratorio	respiratório	Lspé
V respond	responder	répondre	rispondere	responder	Lspé
A responsible	responsable	responsable	responsabile	responsável	LST
V restrict	restringir	restreindre	restringere	restringir	LST
N result	resultado	résultat	risultato	resultado	LST
V result	resultar	résulter	risultare	resultar	LST
A reverse	reverso	inverse	inverso	reverso	Lspé
A rheumatoid	reumatoide	rhumatoïde	reumatoide	reumatoide	Lspé
A ribosomal	ribosomal / ribosómico	ribosomique	ribosomiale	ribossomal / ribossômico	Lspé
N risk	riesgo	risque	rischio	risco	Lspé
A robust	robusto	robuste	robusto	robusto	LST
A rural	rural	rural	rurale	rural	LST
Saccharomyces	Saccharomyces			Saccharomyces	
N cerevisiae	cerevisiae	Saccharomyces cerevisiae	Saccharomyces cerevisiae	cerevisiae	Lspé

V scan	escanear	scanner / numériser	scannerizzare	escanear	LST
N scenario	escenario	scenario	scenario	cenário	LST
A scientific	científico	scientifique	scientifico	científico	LST
N sector	sector	secteur	settore	setor	LST
N segment	segmento	segment	segmento	segmento	LST
A selective	selectivo	sélectif	selettivo	seletivo	LST
A sensitive	sensitivo	sensitif	sensitivo	sensitivo	Lspé
N sensitivity	sensibilidad	sensibilité	sensibilità	sensibilidade	Lspé
A sensory	sensorial / sensorio	sensoriel	sensoriale	sensorial	Lspé
V separate	separar	séparer	separare	separar	LST
N separation	separación	séparation	separazione	separação	LST
N sepsis	septis	sepsis	sepsi	sépsis	Lspé
N sequence	secuencia	séquence	sequenza	sequência	Lspé
V sequence	secuenciar	séquencer	sequenziare	sequenciar	Lspé
V serve	servir	servir	servire	servir	LST
N server	servidor	serveur	server	servidor	LST
N service	servicio	service	servizio	serviço	LST
N service	servicio	service	servizio	serviço	LST
N session	sesión	session	sessione	sessão	LST
A sexual	sexual	sexuel	sessuale	sexual	Lspé
A sexual	sexual	sexuel	sessuale	sexual	Lspé
N shock	choque	choc	shock / choc	choque	Lspé
N sign	signo	signe	segno	senal	Lspé
N significance	significación	signification	significatività	significância	LST
A significant	significativo	significatif	significativo	significativo	LST
A similar	similar	similaire	simile	similar / semelhante	LST
A simple	simple	simple	semplice	simples	LST
A simulated	simulado	simulé	simulato	simulado	LST
N simulation	simulación	simulation	simulazione	simulação	LST
N SNP	SNP	SNP	SNP	SNP	Lspé

A social	social	social	sociale	social	LST
A social	social	social	sociale	social	LST
A social	social	social	sociale	social	Lspé
N society	sociedad	société	società	sociedade	LST
A solid	sólido	solide	solido	sólido	LST
A solid	sólido	solide	solido	sólido	Lspé
A somatic	somático	somatique	somatico	somático	Lspé
A spatial	espacial	spatial	spaziale	espacial	LST
A special	especial	spécial	speciale	especial	LST
A specific	específico	spécifique	specifico	específico	LST
N specificity	especificidad	spécificité	specificità	especificidade	LST
V specify	especificar	spécifier	specificare	especificar	LST
N spectrum	espectro	spectre	spettro	espectro	LST
N sperm	esperma	sperme	sperma	esperma	Lspé
A spinal	espinal	spinal	spinale	espinal / espinhal	Lspé
A spontaneous	espontáneo	spontané	spontaneo	espontâneo	Lspé
N stability	estabilidad	stabilité	stabilità	estabilidade	Lspé
N standard	estándar	standard	standard	standard	LST
A standard	estándar	standard	standard	standard	LST
A standardized	estandarizado	standardisé	standardizzato	estandarizado	LST
N statistic	estadística	statistique	statistica	estatística	LST
N steroid	esteroide	stéroïde	steroido	esteroide	Lspé
V stimulate	estimular	stimuler	stimolare	estimular	Lspé
N stimulation	estimulación	stimulation	stimolazione	estimulação	Lspé
N strategy	estrategia	stratégie	strategia	estratégia	LST
A structural	estructural	structurel	strutturale	estrutural	LST
N structure	estructura	structure	struttura	estrutura	LST
N student	estudiante	étudiant	studente	estudante	LST
A subsequent	subsiguiente	subséquent	sussequente	subsequente	LST
N substance	sustancia	substance	sostanza	substância	Lspé



A substantial	sustancial	substantiel	sostanziale	substancial	LST
N substitution	sustitución	substitution	sostituzione	substituição	Lspé
N substrate	sustrato	substrat	substrato	substrato	Lspé
N subtype	subtipo	sous-type	sottotipo	subtipo	LST
N subunit	* subunidad	sous-unité	* subunità	subunidade	Lspé
V suffer	sufrir	souffrir	soffrire	sofrer	Lspé
N sulfate	sulfato	sulfate	solfato	sulfato	Lspé
N sum	suma	somme	somma	soma	LST
N superfamily	* superfamilia	* superfamille	superfamiglia	superfamília	Lspé
V supplement	* suplementar	supplémenter	* supplementare	suplementar	Lspé
N suppression	supresión	suppression	soppressione	supressão	Lspé
N suppressor	supresor	suppresseur	soppressore	supressor	Lspé
N surface	superficie	surface	superficie	superfície	LST
	surfactante / *				
	tensoactivo /	* surfactant / tensioactif / agent	* surfactante / surfattante /	surfactante /	
N surfactant	tensioactivo	de surface / * surfactif	tensioattivo	tensioativo / tensoativo	Lspé
N susceptibility	susceptibilidad	susceptibilité	suscettibilità	susceptibilidade	Lspé
N syndrome	síndrome	syndrome	síndrome	síndrome	Lspé
A synovial	sinovial	synovial	sinoviale	sinovial	Lspé
N synthase	* sintasa	* synthase	sintasi	sintase	Lspé
V synthesize	sintetizar	synthétiser	sintetizzare	sintetizar	Lspé
N system	sistema	ystème	sistema	sistema	Lspé
N system	sistema	ystème	sistema	sistema	LST
N system	sistema	ystème	sistema	sistema	LST
A systematic	sistemático	systematique	sistemático	sistemático	LST
A systematic	sistemático	systematique	sistemático	sistemático	LST
A systemic	sistémico	systemique	sistemico	sistêmico	Lspé
A systolic	sistólico	systolique	sistolico	sistólico	Lspé
A technical	técnico	technique	tecnico	técnico	LST
N technology	tecnología	technologie	tecnologia	tecnologia	LST

N	temperature	temperatura	température	temperatura	temperatura	Lspé
A	temporal	temporal	temporel	temporale	temporal	LST
A	temporal	temporal	temporal	temporale	temporal	Lspé
V	tend	tender	tendre	tendere	tender	LST
N	term	término	terme	termine	termo	LST
N	terminal	terminal	terminaison	terminale	terminal	Lspé
V	test	testar	tester	testare	testar	LST
N	text	texto	texte	testo	texto	LST
A	therapeutic	terapéutico	thérapeutique	terapeutico	terapêutico	Lspé
N	therapy	terapia	thérapie	terapia	terapia	Lspé
A	thyroid	tiroides	thyroïde	tiroide	tiroide	Lspé
N	tolerance	tolerancia	tolérance	tolleranza	tolerância	Lspé
N	total	total	total	totale	total	LST
A	total	total	total	totale	total	LST
N	tract	tracto / vía	tractus / voie	tratto / via	tracto / via	Lspé
A	traditional	tradicional	traditionnel	tradizionale	tradicional	LST
N	transcript	transcrito	transcrit	trascritto	transcrito	Lspé
V	transfect	* transfectar	* transfecter	* trasfettare	* transfectar	Lspé
N	transfection	* transfección	transfection	transfezione	transfecção	Lspé
N	transfer	transferencia	transfert	trasferimento	transferência	Lspé
V	transfer	transferir	transférer	trasferire	transferir	Lspé
V	transform	transformar	transformer	trasformare	transformar	Lspé
N	transformation	transformación	transformation	trasformazione	transformação	Lspé
A	transgenic	transgénico	transgénique	transgenico	transgênico	Lspé
A	transient	transitorio	transitoire	transitorio	transitório	LST
N	transition	transición	transition	transizione	transição	Lspé
N	transition	transición	transition	transizione	transição	LST
N	translocation	traslocación	translocation	traslocazione	translocação	Lspé
N	transmission	transmisión	transmission	trasmissione	transmissão	Lspé
N	transport	transporte	transport	trasporto	transporte	Lspé

N transporter	transportador	transporteur	trasportatore	transportador	Lspé
N trauma	trauma	traumatisme / trauma	trauma	trauma	Lspé
V treat	tratar	traiter	trattare	tratar	Lspé
V treat	tratar	traiter	trattare	tratar	Lspé
N treatment	tratamiento	traitement	trattamento	tratamento	Lspé
N trend	tendencia	tendance	tendenza	tendência	LST
N tube	tubo	tube / éprouvette	tubo	tubo	Lspé
N tuberculosis	tuberculosis	tuberculose	tubercolosi	tuberculose	Lspé
N tumor	tumor	tumeur	tumore	tumor	Lspé
N type	tipo	type	tipo	tipo	LST
N tyrosine	tirosina	tyrosine	tirosina	tirosina	Lspé
N ultrasound	ultrasonido	ultrason	ultrasuono	ultrassom	Lspé
N unit	unidad	unité	unità	unidade	LST
N unit	unidad	unité	unità	unidade	LST
N university	universidad	université	università	universidade	LST
A urban	urbano	urbain	urbano	urbano	LST
A urban	urbano	urbain	urbano	urbano	LST
A urinary	urinario	urinaire	urinario	urinário	Lspé
A urinary	urinario	urinaire	urinario	urinário	Lspé
A uterine	uterino	utérin	uterino	uterino	Lspé
N utility	utilidad	utilité	utilità	utilidade	LST
N utilization	utilización	utilisation	utilizzazione	utilização	LST
V utilize	utilizar / usar	utiliser	utilizzare / usare	utilizar / usar	LST
A valid	válido	valide	valido	válido	LST
V validate	validar	valider	validare	validar	LST
N validation	validación	validation	validazione	validação	LST
N validity	validez	validité	validità	validade	LST
N variability	variabilidad	variabilité	variabilità	variabilidade	LST
N variance	varianza	variance	varianza	variância	LST
N variant	variante	variant	variante	variante	Lspé

N variant	variante	variant	variante	variante	LST
N variety	variedad	variété	varietà	variedade	LST
V vary	variar	varier	variare	variar	LST
A vascular	vascular	vasculaire	vascolare	vascular	Lspé
N vector	vector	vecteur	vettore	vetor	Lspé
N vein	vena	veine	vena	veia	Lspé
N velocity	velocidad	vélocité / vitesse	velocità	velocidade	LST
A venous	venoso	veineux	venoso	venoso	Lspé
N ventilation	ventilación	ventilation	ventilazione	ventilação	Lspé
N ventilation	ventilación	ventilation	ventilazione	ventilação	Lspé
A ventricular	ventricular	ventriculaire	ventricolare	ventricular	Lspé
				verificar / conferir /	
V verify	verificar / averiguar	vérifier	verificare	averiguar	LST
N version	versión	version	versione	versão	LST
N version	versión	version	versione	versão	LST
A viral	viral	viral	virale	viral	Lspé
N virus	virus	virus	virus	vírus	Lspé
A visual	visual	visuel	visuale	visual	Lspé
N visualization	visualización	visualisation	visualizzazione	visualização	LST
V visualize	visualizar	visualiser	visualizzare	visualizar	LST
N vitamin	vitamina	vitamine	vitamina	vitamina	Lspé
N volt	voltio	volt	volt	volt	Lspé
N volume	volumen	volume	volume	volume	LST
N web	web	web/toile	web	web	LST
N zinc	zinc / cinc	zinc	zinco	zinco	Lspé