



HAL
open science

Représentation et construction d'une identité européenne à travers la presse française. Une étude en linguistique de corpus au sein d'un projet de l'Union Européenne.

Delphine Giuliani

► **To cite this version:**

Delphine Giuliani. Représentation et construction d'une identité européenne à travers la presse française. Une étude en linguistique de corpus au sein d'un projet de l'Union Européenne.. Sciences de l'Homme et Société. Université de Bretagne Sud, 2009. Français. NNT: . tel-00634154

HAL Id: tel-00634154

<https://theses.hal.science/tel-00634154>

Submitted on 20 Oct 2011

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.

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

**Représentation et construction d'une identité européenne à
travers la presse française.**

**Une étude en linguistique de corpus au sein d'un projet de
l'Union Européenne.**

THESE DE DOCTORAT

Discipline : linguistique.

Spécialité : linguistique de corpus

Présentée et soutenue publiquement par

Delphine GIULIANI

le 8 décembre 2009, devant le jury ci-dessous

Présidente de jury : Mireille BILGER

Rapporteurs BILGER Mireille, *professeur* · Université de Perpignan

LE CUNFF Catherine, *maître de conférence HDR* · Université de Paris Nanterre

Examineurs LOZACH'MEUR Ghislaine, *maître de conférence* · Université de Bretagne Ouest

GALLERON Ioana, *maître de conférence* · Université de Bretagne Sud

Directeur de thèse : M. Geoffrey C. WILLIAMS

UNIVERSITE DE BRETAGNE SUD
UFR DE LETTRES ET SCIENCES HUMAINES

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

**Représentation et construction d'une identité
européenne à travers la presse française.
Une étude en linguistique de corpus au sein d'un projet
de l'Union Européenne.**

THESE DE DOCTORAT

Discipline : linguistique.

Spécialité : linguistique de corpus

Présentée et soutenue publiquement par

Delphine GIULIANI

le 8 décembre 2009, devant le jury ci-dessous

Présidente de jury : Mireille BILGER

Rapporteurs BILGER Mireille, *professeur* · Université de Perpignan

LE CUNFF Catherine, *maître de conférence HDR* · Université de Paris Nanterre

Examineurs LOZACH'MEUR Ghislaine, *maître de conférence* · Université de Bretagne Ouest

GALLERON Ioana, *maître de conférence* · Université de Bretagne Sud

Directeur de thèse : M. Geoffrey C. WILLIAMS

Remerciements.

Je tiens à remercier tous ceux qui m'ont apporté leur aide, tant matérielle que psychologique, tout au long de ce travail. Nommer toutes les personnes qui m'ont soutenue lors de ces trois années me serait impossible, je m'excuse auprès de tous ceux que j'oublie.

Mon directeur de thèse a été présent chaque jour de cette thèse pour répondre à mes questions, me donner ses avis, ses conseils et ses encouragements. Il m'a également permis de découvrir la discipline de linguistique de corpus. C'est grâce à lui que cette thèse a pu aboutir et je l'en remercie. Mes collègues de l'équipe LiCoRN (laboratoire HCTI), ont toujours été présents pour m'aider lors de requêtes techniques, ou toute forme de question, et la construction du corpus sur lequel ces travaux se basent est avant tout un travail d'équipe, je les en remercie. Les membres du projet IntUne ont également toujours répondu à mes interrogations quelles qu'elles soient avec pertinence et célérité, ils m'ont aussi permis d'avancer sur certains points de ma recherche lors des réunions, ateliers et colloques du projet, je les en remercie tous.

Je tiens à exprimer toute ma gratitude envers les personnels administratifs de l'école doctorale de l'Université de Bretagne Sud qui m'ont apporté leur assistance et leurs conseils.

Je remercie également Chloé Raimbau qui a accepté de relire le texte en un temps record, et qui m'a apporté de judicieuses remarques.

Je voudrais plus particulièrement remercier mes famille et belle famille, ainsi que mon mari qui m'ont soutenu du mieux qu'ils le pouvaient tout au long de ce projet, avec intelligence et amour.

Enfin, je suis seule responsable des défauts et erreurs qui pourraient subsister dans ce travail.

Table des matières

<i>0- Introduction.....</i>	<i>1</i>
<i> Qu'est-ce qu'IntUne ?</i>	<i>2</i>
<i> Pourquoi un groupe média, quel intérêt à l'analyse des médias ?</i>	<i>3</i>
<i> Le groupe média, un premier projet de création de corpus.</i>	<i>6</i>
<i> Contenu de la thèse.....</i>	<i>8</i>
<i>1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?</i>	<i>12</i>
<i> I.0- Introduction.</i>	<i>12</i>
<i> I.1/ Linguistique de corpus et linguistique générale.</i>	<i>13</i>
<i> I.1.1- Qu'est-ce qu'un corpus ?.....</i>	<i>13</i>
<i> I.1.2- La relation entre linguistique et corpus.....</i>	<i>16</i>
<i> I.1.2.1- Saussure.</i>	<i>16</i>
<i> I.1.2.1.1- Le Cours de linguistique générale.</i>	<i>17</i>
<i> I.1.2.1.2- Les écrits de la main de Saussure.</i>	<i>18</i>
<i> I.1.2.2- Une linguistique de la langue.</i>	<i>20</i>
<i> I.1.2.2.1- Le structuralisme.</i>	<i>20</i>
<i> I.1.2.2.2- Le courant logico-grammatical.....</i>	<i>22</i>
<i> I.1.2.3- Une linguistique de la parole.....</i>	<i>26</i>
<i> I.1.2.3.1- Les sources du contextualisme britannique.</i>	<i>26</i>
<i> I.1.2.3.2- L'école contextualiste britannique.....</i>	<i>28</i>
<i> I.1.2.3.3- La naissance d'une linguistique de corpus moderne.</i>	<i>31</i>
<i> I.1.3- La linguistique de corpus, une discipline aux multiples facettes.....</i>	<i>32</i>
<i> I.1.3.1- La filiation firthienne : John Sinclair d'Edimbourg à Birmingham.....</i>	<i>33</i>
<i> I.1.3.1.1- Les projets OSTI et COBUILD, le corpus en lexicographie appliquée.....</i>	<i>33</i>
<i> I.1.3.1.2- Corpus et linguistique générale.</i>	<i>34</i>
<i> I.1.3.2- La filiation Quirk-Leech.....</i>	<i>37</i>
<i> I.1.3.2.1- L'école de Lancaster, l'apport de Geoffrey Leech.</i>	<i>37</i>
<i> I.1.3.2.2- Le British National Corpus.</i>	<i>38</i>
<i> I.2- Le corpus, une entité à géométrie variable.....</i>	<i>42</i>
<i> I.2.1- Les corpus électroniques : multiplicité de formes et d'approches.</i>	<i>43</i>
<i> I.2.1.1- Différents types de corpus.</i>	<i>43</i>

I.2.1.1.1- Les corpus parallèles.....	43
I.2.1.1.2- Les corpus comparables.	44
I.2.1.2 – Différents types d'approches.	46
I.2.1.2.1- L'approche qualitative.....	46
I.2.1.2.2 – L'approche quantitative.....	47
I.2.1.2.3- La démarche « corpus-based ».....	48
I.2.1.2.4- L'approche « corpus-driven ».....	49
I.2.2- Les limites des corpus.....	50
I.2.2.1- Un problème de définition.	50
I.2.2.2- Le problème de la représentativité.	51
I.2.3- Différentes applications.	52
I.2.3.1- Différentes applications.	52
I.2.3.1.1- Dans le domaine de la traduction, et de son enseignement.	53
I.2.3.1.2- Dans le domaine de l'apprentissage des langues étrangères.	54
I.3- Conclusion.....	55
2- Cadre théorique, méthodologie employée.	57
II.0- Introduction.....	57
II.1- Le mot, une unité problématique.	58
II.1.1- Aristote, la théorie des conditions nécessaires et suffisantes.	60
II.1.2- L'application de ce modèle dans la linguistique de la seconde moitié du vingtième siècle.....	62
II.1.3- Les limites du modèle des CNS dans le domaine de la sémantique et de la lexicographie.....	64
II.2- La théorie des prototypes.	66
II.2.1- Wittgenstein et l'air de famille.	67
II.2.2- Le prototype en psychologie, l'apport d'Eleanor Rosch.....	70
II.2.3- La proposition d'une synthèse de la théorie du prototype.....	75
II.2.4- Les failles de la théorie du prototype en sémantique.	79
II.3- Les prototypes selon Patrick Hanks.....	81
II.3.1- Les problèmes rencontrés en lexicographie.	82
II.3.2- Signification et sens potentiel.	84
II.3.3- Un exemple de prototype, le cas de « bank ».	86
II.4- Application du cadre théorique dans la thèse, création d'un prototype lexicographique.	87

II.4.1-	<i>Le prototype pour mieux appréhender la représentation de l'Europe.....</i>	<i>87</i>
II.4.2-	<i>Un prototype de base.</i>	<i>88</i>
II.4.2.1-	<i>Méthode d'élaboration.</i>	<i>88</i>
II.4.2.2-	<i>Un prototype lexicographique pour le mot « Europe ».</i>	<i>89</i>
II.5-	<i>Conclusion.....</i>	<i>90</i>
3-	<i>Technique de construction et d'analyse du corpus IntUne.....</i>	<i>92</i>
III.0-	<i>Introduction.....</i>	<i>92</i>
III.1-	<i>L'élaboration du corpus IntUne.....</i>	<i>92</i>
III.1.1-	<i>Le corpus pilote.</i>	<i>93</i>
III.1.2-	<i>L'élaboration du corpus IntUne.....</i>	<i>94</i>
III.1.2.1-	<i>Le corpus de presse écrite.</i>	<i>94</i>
III.1.2.2-	<i>Le corpus télévisé.....</i>	<i>97</i>
III.1.2.3-	<i>Structuration du corpus IntUne.....</i>	<i>99</i>
III.1.3-	<i>Le corpus IntUne en quelques chiffres.</i>	<i>102</i>
III.2-	<i>Les analyses menées sur le corpus.....</i>	<i>105</i>
III.3-	<i>Conclusion.....</i>	<i>108</i>
4-	<i>« Europe ». Quelle représentation dans la presse française ?</i>	<i>110</i>
IV.0-	<i>Introduction.....</i>	<i>110</i>
IV.1-	<i>L'Europe dans le corpus IntUne : vue globale.</i>	<i>113</i>
IV.1.1-	<i>L'Europe en chiffres dans le corpus.</i>	<i>113</i>
IV.1.2-	<i>Quelle est la compagnie du concept Europe ?</i>	<i>117</i>
IV.1.2.1-	<i>Le mot « Europe ».....</i>	<i>117</i>
IV.1.2.2-	<i>Le lemme « européen ».</i>	<i>123</i>
IV.2-	<i>Actualisation du prototype sémantique de « Europe » dans la presse française en 2007 et 2009.</i>	<i>127</i>
IV.2.1-	<i>L'Europe, une masse terrestre plus ou moins précisément délimitée.</i>	<i>128</i>
IV.2.1.1-	<i>Une masse terrestre que l'on peut appréhender concrètement.</i>	<i>128</i>
IV.2.1.2-	<i>Une masse terrestre aux contours flous et problématiques.....</i>	<i>129</i>
IV.2.1.3-	<i>Le problème de l'énergie.</i>	<i>130</i>
IV.2.1.4-	<i>Quelle évolution en 2009 ?.....</i>	<i>131</i>
IV.2.2-	<i>L'Europe, une entité dont les pays sont regroupés dans des blocs.</i>	<i>132</i>
IV.2.2.1-	<i>L'Europe occidentale.</i>	<i>132</i>
IV.2.2.1.1-	<i>L'Europe occidentale en 2007.</i>	<i>132</i>

IV.2.2.1.2- L'Europe occidentale en 2009.	134
IV.2.2.2- L'Europe de l'Ouest.....	136
IV.2.2.2.1- L'Europe de l'Ouest en 2007.	136
IV.2.2.2.2- L'Europe de l'Ouest en 2009.....	137
IV.2.2.3- L'Europe continentale.....	138
IV.2.2.3.1- L'Europe continentale en 2007.....	138
IV.2.2.3.2- L'Europe continentale en 2009.	139
IV.2.2.4- L'Europe centrale.....	140
IV.2.2.4.1- L'Europe centrale en 2007.	140
IV.2.2.4.2- L'Europe centrale en 2009.....	142
IV.2.2.5- L'Europe de l'Est.....	143
IV.2.2.5.1- L'Europe de l'Est en 2007.....	144
IV.2.2.5.2- L'Europe de l'Est en 2009.....	145
IV.2.2.6- L'Europe orientale.....	146
IV.2.2.7- L'Europe du Nord.....	147
IV.2.2.8- L'Europe du Sud.	148
IV.2.2.9- Une Europe unifiée ?.....	149
IV.2.3- Une entité économique.....	150
IV.2.3.1- De bons résultats.....	150
IV.2.3.2- Des difficultés.....	151
IV.2.3.3- Une zone qui innove et forme.....	152
IV.2.3.4- Une économie agricole importante.....	153
IV.2.3.5- Une entité négociant d'une voix au niveau mondial.....	153
IV.2.3.6- Pour conclure.....	154
IV.2.4- Une entité partiellement composée d'une Union politique : l'Union Européenne.	154
IV.2.4.1- L'Union européenne confrontée à de nombreuses difficultés.....	155
IV.2.4.2- Une union dotée d'une image favorable.....	157
IV.2.4.3- Une entité politique qui se base sur une culture commune.	158
IV.2.4.3.1- Une vie artistique rassurante, entre homogénéité stabilisatrice et diversité enrichissante.	158
IV.2.4.3.2- Une Histoire commune ancienne, alternant entre violence et coopération.	159
IV.2.4.3.3- Une entité marquée par les religions.....	159
IV.2.4.3.4- Des influences linguistiques communes.....	160
IV.2.4.4- Une entité politique à géométrie variable.	161
IV.2.4.5- Pour conclure.	162

IV.2.5.1- Une idéalisation présente à travers différentes constructions linguistiques.	163
IV.2.5.2- L'idéalisation à travers le thème de l'engagement environnemental.	164
IV.2.5.3- Un sujet de l'élection présidentielle française.	165
IV.2.6- L'Europe politique doit défendre ses frontières.	166
IV.2.6.1- Une entité au sein de tensions internationales.	167
IV.2.6.2- Le problème de l'immigration.	169
IV.2.6.3- La menace terroriste.	170
IV.2.6.4- Protéger les frontières de problèmes sanitaires.	171
IV.2.7- Une entité composée de plusieurs pays.	172
IV.2.7.1- Des comportements différents au niveau économique.	172
IV.2.7.2- Des comportements différents dans d'autres situations.	175
IV.2.7.3- Une entité composée de sociétés.	178
IV.2.8- Une entité sportive.	179
IV.2.8.1- La France dans l'Europe sportive.	179
IV.2.8.2- Quels pays dans les compétitions sportives européennes ?	181
IV.2.8.3- L'Europe sportive en règle générale.	184
IV.3- Conclusion.	184
5- Deux cas particuliers : l'Italie et la Pologne.	187
V.0- Introduction.	187
V.1- Qu'est-ce qu'un pays ou une nation dans la presse française ?	289
V.1.1- Élaboration du prototype de « pays » dans la presse française.	290
V.1.1.1- Qu'est-ce qu'un « pays » selon les dictionnaires ?	290
V.1.1.2- « Pays » dans le corpus IntUne.	293
V.1.2- Élaboration du prototype de « nation » dans la presse française.	299
V.1.2.1- Qu'est-ce qu'une « nation » dans différents dictionnaires ?	299
V.1.2.2- « Nation » dans le corpus IntUne.	200
V.1.3- Conclusion.	205
V.2- La représentation de l'Italie dans la presse française en 2007 et 2009.	207
V.2.1- Vue globale et élaboration d'un prototype.	207
V.2.2- Actualisation des sens potentiels du prototype dans la presse.	211
V.2.2.1- L'Italie comme étendue géographique.	211
V.2.2.1.1- En 2007.	211
V.2.2.1.2- En 2009.	212

V.2.2.2- <i>L'Italie comme une étendue culturelle.</i>	213
V.2.2.2.1- <i>En 2007.</i>	213
V.2.2.2.2- <i>En 2009.</i>	214
V.2.2.3- <i>L'Italie comme un espace historique.</i>	215
V.2.2.3.1- <i>En 2007.</i>	215
V.2.2.3.2- <i>En 2009.</i>	215
V.2.2.4- <i>« Italie » comme étendue économique.</i>	216
V.2.2.4.1- <i>En 2007.</i>	216
V.2.2.4.2- <i>En 2009.</i>	217
V.2.2.5- <i>L'Italie en tant qu'espace politique.</i>	219
V.2.2.5.1- <i>En 2007.</i>	219
V.2.2.5.2- <i>En 2009.</i>	220
V.2.2.6- <i>L'Italie comme un lieu où des personnes vivent.</i>	222
V.2.2.6.1- <i>En 2007.</i>	222
V.2.2.6.2- <i>En 2009.</i>	222
V.2.2.7- <i>L'Italie comme un pays sportif.</i>	224
V.2.2.7.1- <i>En 2007.</i>	224
V.2.2.7.2- <i>En 2009.</i>	225
V.2.2.8- <i>Conclusion.</i>	226
V.3- <i>La représentation de la Pologne dans la presse française en 2007 et en</i>	
<i>2009.</i>	227
V.3.1- <i>Vue globale et élaboration d'un prototype.</i>	227
V.3.2- <i>Analyse des lignes de concordances.</i>	231
V.3.2.1- <i>La Pologne comme une étendue géographique.</i>	231
V.3.2.2- <i>« Pologne » comme une étendue culturelle.</i>	232
V.3.2.3- <i>La Pologne comme un espace historique.</i>	232
V.3.2.4- <i>La Pologne comme une étendue économique.</i>	233
V.3.2.5- <i>La Pologne comme une étendue politique.</i>	234
V.3.2.6- <i>La Pologne comme un lieu où des personnes vivent.</i>	238
V.3.2.7- <i>Un pays sportif.</i>	239
V.3.3- <i>Conclusion.</i>	240
V.4- <i>Conclusion.</i>	241
6- <i>La télévision française : quelle représentation d'un événement au</i>	
<i>retentissement européen ?</i>	243
VI.0- <i>Introduction.</i>	243

VI.1- Cadre méthodologique.....	244
VI.1.1- La relation de communication télévisuelle.	245
VI.1.1.1- L'émetteur.	245
VI.1.1.2- Le récepteur.	247
VI.1.2- Les éléments non linguistiques du message.....	247
VI.1.2.1- Les éléments sonores.	248
VI.1.2.2- La Vidéo.....	249
VI.1.3- Les éléments linguistiques.	252
VI.1.3.1- Les traces de l'émetteur dans l'énonciation.....	252
VI.1.3.2- Les marques spatio-temporelles de l'énonciation.....	253
VI.2- Analyse comparative des traitements de France 3 et TF1 pour un même sujet.	255
VI.2.1- Les titres.	256
VI.2.1.1- France 3.....	256
VI.2.1.2- TF1.....	260
VI.2.1.3- Comparaison.....	263
VI.2.2- Le premier reportage.	264
VI.2.2.1- FR3.....	264
VI.2.2.1.1- L'annonce par la présentatrice.....	264
VI.2.2.1.2- Le reportage.	265
VI.2.2.2- TF1.	267
VI.2.2.2.1- Le lancement du sujet.....	267
VI.2.2.2.2- Le reportage.	268
VI.2.2.3- Comparaison.	270
VI.3- Mise en perspective à l'aide du corpus IntUne.	270
VI.3.1- Les reportages de TF1 sur l'Italie en 2007.	271
VI.3.2- La représentation de l'événement dans le corpus de presse écrite.	276
VI.4- Conclusion.....	278
7- Conclusion.....	280
Annexe 1.Organigramme du projet IntUne.....	287
Annexe 2. Fonctionnement des quatre groupes du projet IntUne.....	287
Annexe 3. Extrait du corpus presse brut.....	288
Annexe 4- Extrait du corpus télévisé brut.	290

<i>Annexe 5- Captures d'écran du logiciel transcriber.....</i>	<i>293</i>
<i>Annexe 6- Feuille de style xslt utilisée pour le corpus télévisé.</i>	<i>294</i>
<i>Annexe 7- Carte physique de l'Europe.</i>	<i>297</i>
<i>Annexe 8- Collocations de "Europe" dans le corpus IntUne 2007 et 2009.</i>	<i>298</i>
<i>Annexe 9- La France dans l'Europe sportive.</i>	<i>302</i>
<i>Annexe 10- Les différents pays de l'Europe sportive.</i>	<i>303</i>
<i>Annexe 11- Contexte de « du pays ».....</i>	<i>304</i>
<i>Annexe 12 – Tableau récapitulatif du reportage de France 3.....</i>	<i>305</i>
<i>Annexe 13- Tableau récapitulatif du reportage de TF1.....</i>	<i>314</i>
<i>Annexe 14 : Textes des reportages télévisés de TF1 concernant l'Italie en 2007.....</i>	<i>327</i>
<i>Bibliographie.....</i>	<i>334</i>

0- Introduction.

« Et de l'union des libertés dans la fraternité des peuples, naîtra la sympathie des âmes, germe de cet immense avenir où commencera pour le genre humain la vie universelle et que l'on appellera la paix de l'Europe » (Victor Hugo, Choses Vues)

« L'Europe doit respirer avec ses deux poumons, celui de l'est et celui de l'ouest. » (Jean-Paul II)

L'Europe est un mot dont la simplicité de prononciation cache une réalité complexe. En perpétuelle évolution, disposant de racines multiples, couvrant des sens divers d'un point de vue synchronique et diachronique, cette entité ne se laisse appréhender qu'avec difficultés. Étudier les représentations et l'émergence ou la construction d'une identité européenne peut recouvrir une multitude de chemins différents. L'étude linguistique d'un corpus de presse en est un que nous proposons d'emprunter dans cette thèse.

La linguistique de corpus est un paradigme qui connaît depuis les vingt dernières années un succès croissant (Williams 2006). Cette locution voit son usage se répandre dans les colloques et les travaux de recherches de disciplines multiples et variées : histoire, sociologie, psychologie, linguistique, didactique sont autant de domaines où de nombreux chercheurs pratiquent des études en linguistique de corpus (Baker 2009, 2008). Mais qu'est-ce que signifie cette locution ? S'agit-il d'une nouvelle discipline, d'un nouveau paradigme, d'une simple méthodologie, d'une branche nouvelle de la linguistique ? Selon le pays dans

lequel on se situe, la langue que l'on utilise – la locution d'origine étant anglophone : *corpus linguistics* - , la tradition linguistique dans laquelle on s'inscrit et les buts dans lesquels on mène une recherche, la réponse diffèrera grandement.

La linguistique de corpus connaît un développement croissant, celui-ci est en partie du aux multiples applications que cette discipline permet. Nous détaillerons ainsi notamment l'utilisation qui peut être faite des découvertes des linguistes de corpus dans des domaines comme la traduction ou l'enseignement d'une langue étrangère. Les corpus peuvent en effet être des bases pour la réalisation de dictionnaires pour des apprenants avancés, permettant une représentation plus fine de la lexie utilisée dans telle ou telle langue. En outre, ils permettent aux apprenants de voir comment les mots sont utilisés authentiquement et en contexte. Cette thèse s'inscrit à la fois dans cette discipline en plein essor aux nombreuses facettes et applications qu'est la linguistique de corpus, et dans le cadre d'un projet européen de recherche sur l'émergence et la représentation d'une identité et d'une citoyenneté européenne, IntUne.

Qu'est-ce qu'IntUne ?

IntUne, *Integrated and United ? A quest in an ever closer Europe*, est un projet intégré du PCRD6 de l'Union Européenne¹ qui implique 32 partenaires dans 15 pays différents. Il s'agit d'étudier les évolutions concernant la façon dont la citoyenneté européenne se construit, notamment du fait du procédé d'élargissement ou encore de la présence de plus en plus importante de l'Union Européenne dans la vie des citoyens européens. Cette étude évolutive a lieu au cours d'une période de quatre ans allant de 2005 à 2009. Ce projet est dirigé de l'université de Sienne en collaboration avec différents centres de recherche en sociologie et en sciences politiques comme le Centre de Recherches Politiques de la Sorbonne ou l'Institut Français des relations internationales.

Pour mener cette recherche, il était question à l'origine du projet d'étudier les opinions des masses et des élites de façon classique à travers des sondages ou des entretiens. Néanmoins la Commission européenne ainsi que les porteurs du projet ont noté la nécessité d'analyser également les médias. Des spécialistes de l'étude des médias et linguistes de corpus ont alors été contactés en Italie, ils ont ensuite sollicité d'autres spécialistes de leur

1 <http://www.intune.it/>

discipline dans quatre pays pour analyser le discours des médias.

Le projet s'organise ainsi en différents groupes étudiant les élites (chercheurs en contact avec les dirigeants élus ou hauts fonctionnaires), les masses (l'opinion publique), les médias (linguistes de corpus et analystes du discours) et les experts (spécialistes des questions de fiscalité, de défense, etc.). Un organigramme en est présenté en annexes (cf. annexe 1). Ces groupes travaillent de façon croisée, s'intéressant tous à trois questions principales qui sont les suivantes : Comment l'organisation politique influence-t-elle la notion de citoyenneté et comment différentes identités peuvent-elles coexister au sein d'un système complexe ? Quelles obligations les citoyens européens développent-ils à l'égard de leur gouvernance ? Qu'attendent les citoyens européens de l'Union Européenne ? (Williams, 2009²). Ces quatre groupes permettent également d'instaurer un échange de spécialistes issus de différentes disciplines. La façon dont leur travail, dans quatre domaines différents, mène à une meilleure compréhension des questions principales posées dans ce projet peut se résumer par le schéma présent dans l'annexe 2. Le projet IntUne vise également à donner une base de connaissances sur la façon dont la citoyenneté européenne s'est développée à un niveau scientifique, et à améliorer le niveau de recherche européen via l'emploi de techniques nouvelles ainsi que la collecte de nouvelles données. En termes d'applications, il s'agit de produire un savoir utilisable pour les décideurs politiques, les acteurs clefs, mais également les masses.

Pourquoi un groupe média, quel intérêt à l'analyse des médias ?

Le projet IntUne souhaite étudier l'émergence de l'identité supranationale des citoyens de l'Union Européenne. Or, de nombreux sociologues et politologues se sont fréquemment posés la question de l'influence des médias sur les opinions publiques. Langelier, chercheur en sociologie, dresse une rétrospective de la réflexion dans ce domaine qui a été menée au cours du vingtième siècle jusqu'à aujourd'hui (Langelier 2006). Il note que cette réflexion est notamment née de la seconde guerre mondiale, après l'usage de la propagande par le régime nazi (Langelier 2006:5). A la suite de cette observation, divers chercheurs se sont posés la question de savoir dans quelle mesure les médias pouvaient influencer l'opinion publique. Une première théorie béhavioriste a été développée dans les années 1950 selon laquelle il aurait suffi de créer un stimuli informationnel pour obtenir le

2 L'identité européenne, un concept à géométrie variable, issu d'un colloque à Poitiers.

résultat souhaité modifiant l'opinion publique. Ces travaux se sont surtout développés aux États-Unis à travers les publications de Lasswell ou en France à travers ceux de Le Bon. Cependant, ces idées ont rapidement été remises en question (Ibid:7) avec Cazeneuve en France ou Lazarsfeld aux États-Unis. D'une part, l'importance du contexte dans lequel le message était reçu était mise en relief, d'autre part, des études de Cazeneuve dans les années 1970 ont montré que le média pouvait surtout avoir un effet de renforcement des opinions déjà existantes et qu'une attaque frontale des croyances n'était pas efficace. Actuellement, les études tendent à comprendre les conséquences sur la socialisation et le contrôle social de l'action des médias. L'idée est que ceux-ci seraient un instrument de visibilité très efficace pour ceux qui détiennent le pouvoir (Ibid:12) et qu'ils auraient une influence sur la priorité accordée à certains débats menés dans la société (Ibid:14). Les médias semblent donc pouvoir influencer sur l'émergence d'une citoyenneté parmi l'opinion publique, au moins en relayant cette émergence qui serait déjà pré-observée au sein des masses.

Dans le cadre plus restreint de l'influence d'un discours médiatique sur la constitution d'une identité nationale ou d'une citoyenneté, Retzläff et Gänzle se sont intéressés dans *Constructing the European Union in Canada News* (2008) à la construction de ce qu'était l'identité européenne selon les journaux canadiens. Teubert s'est posé la question de savoir d'où pouvait venir le sentiment eurosceptique britannique dans *A Province of a Federal Superstate, Ruled by an unelected Bureaucracy, keywords of the Euro-Sceptic Discourse in Britain* (2001). Enfin, Higgins (2004) s'est interrogé sur le rôle des journaux écossais comme facilitateurs d'une émergence de l'identité écossaise dans *Putting the nation in the news : the rôle of location formulation in a selection of scottish newspapers*. Ces différentes études, parmi d'autres, font le postulat que la notion d'identité nationale ou supra nationale peut se construire dans les journaux écrits et télévisés. Cela peut s'expliquer par différentes raisons. D'une part, les nations et en conséquence l'identité nationale sont souvent considérées comme des constructions mentales, De Cillia cite notamment le concept de « imagined political communities » de Andersen (De Cillia 1999:153). La nation est donc un objet social, et en tant que tel, se construit nécessairement à travers le discours, ce que De Cillia précise en citant Stuart Hall : « A national culture is a discourse. » (De Cillia 1999:155).

La tradition en analyse du discours établit une relation réciproque, dialectale, entre

les personnes qui façonnent le discours, et la façon dont ce discours les façonne (Retzlaff et Gänzle 2008). S'ils sont façonnés, alors leur identité sera nécessairement affectée par le discours, mais en tant que producteurs d'un discours fréquemment consulté, les médias ont également une influence sur cette identité. Dans le cadre de l'analyse discursive, le discours est considéré comme un système complexe de textes écrits ou oraux dans lequel on reproduit une connaissance, un pouvoir ou un contrôle. Dans les sociétés modernes, l'analyse de discours montre que les relations de pouvoir sont souvent reproduites et légitimées à un niveau idéologique, sont répétées dans les discours et peuvent ainsi être « internalisées » par la masse des citoyens et lecteurs. Les médias sont un sujet typique de l'analyse discursive en ce que les représentations des journaux sont cruciales pour la représentation des cultures, des peuples, de la vie sociale et politique. Les journaux sont souvent la première source d'information de la masse des citoyens. Retzlaff et Gänzle citent ainsi Pietikäinen :

*« News representations contribute to ways in which people see themselves, their own identities and that of others and the relations between « us » and « them ». For any group [...] news coverage is a means of gaining wider attention for their agenda, of making their voices heard, and of possibly making a difference on issues important to them. »
(Pietikäinen in Retzlaff et Gänzle 2008:69)³*

D'autre part, il est intéressant de voir que les différents journaux d'une même nation peuvent avoir différentes représentations quant à l'identité de cette nation. Les études précédemment citées montrent ainsi que différentes stratégies ou réalisations discursives concernant l'identité nationale peuvent être mises en place. Selon De Cillia (1999), les journaux peuvent ainsi être tentés de suivre différentes stratégies dans l'élaboration d'une identité nationale. Il peut s'agir de mettre l'accent sur les réalisations communes d'une nation, sur son histoire commune, mais aussi sur ses différences avec d'autres pays, ou sur ses échecs. Retzlaff et Gänzle (2008) montrent que cela passe par différents types de réalisations linguistiques et donnent en exemple l'usage signifiant qui est fait dans les journaux de la métaphore, des pronoms, ou des parallélismes. Higgins (2004) considère pour sa part que les journaux sont des facilitateurs dans l'émergence d'une identité nationale et donne en exemple

3 « Les représentations des journaux contribuent à la façon dont les peuples se voient eux-mêmes, leurs propres identités et celles des autres, ainsi que les relations entre « nous » et « eux ». Pour n'importe quel groupe, la façon dont les journaux couvrent un événement est une façon de donner davantage d'attention à leur calendrier, de faire entendre leur voix et de potentiellement de faire une différence sur des questions importantes pour eux. » (ma traduction)

l'usage par les journaux écossais des indicateurs spatiaux.

Ces études ne sont pas des études en linguistique de corpus, il s'agit d'analyse de discours ne se basant pas sur les techniques habituelles de notre discipline. Néanmoins, leurs résultats semblent montrer une certaine influence des médias sur la constitution d'une identité dans l'opinion publique. La presse écrite et télévisée semble donc fournir un bon point de départ pour toute étude concernant l'identité nationale et son émergence chez la masse des citoyens. D'une part, l'identité est une construction mentale, donc un produit social et elle peut par conséquent être analysée de façon pertinente dans des réalisations discursives. D'autre part, les médias sont une source d'information primordiale et influente dans la conscience des citoyens. C'est pourquoi le projet IntUne a pris notamment comme source de matériau pour son corpus la presse écrite et orale des quatre pays où il devait être élaboré.

Le groupe média, un premier projet de création de corpus.

Cette thèse est un des travaux réalisés au sein du groupe média. Ce groupe se compose de linguistes de corpus intéressés par l'analyse du discours et travaillant selon des approches à la fois herméneutiques et empiriques. Il implique des chercheurs de quatre universités européennes : l'Université de Sienna dont la spécialité est l'analyse de discours sur corpus avec une équipe de chercheurs coordonnée par John Morley; l'Université de Cardiff spécialisée dans l'analyse de discours politique avec pour coordinateurs Joanna Thornborrow et Gordon Tucker; l'Université de Bretagne-Sud qui s'intéresse plus particulièrement à l'analyse des réseaux conceptuels prototypiques avec pour coordinateur Geoffrey Williams et l'Université de Lodz, spécialisée dans le traitement automatique de la langue avec pour coordinatrice Barbara Lewandowska-Tomaszcyk.

Le travail de collecte porte en conséquence sur des textes produits en Italie, au Royaume-Uni, en France et en Pologne. Ces pays font également l'objet de recherches de la part des groupes consacrés aux élites et aux masses, ces études prenant place en même temps que la collecte de données, afin de pouvoir par la suite procéder à un travail comparatif. Les périodes de collectes de données ont été en conséquence prévues du 5 février au 5 mai 2007 (correspondant en France à la période de campagne électorale présidentielle) pour le corpus presse, et du 12 février au 4 avril 2007 pour le corpus télévisé. Un deuxième corpus a été

réalisé du 2 février 2009 au 02 mai 2009 pour la presse et du 09 février au 30 mars 2009, cette deuxième période correspondant à la campagne pour les élections législatives européennes, mais surtout, dans les faits et sans que cela ait été initialement envisagé, à une crise économique internationale forte. Les études des autres groupes (masses et élites) ont été menées sur les mêmes périodes. L'objectif est en effet de voir si une interaction a lieu entre ces différents groupes d'opinions.

Dans l'annexe 1 du descriptif du projet IntUne, une des réalisations concrètes à fournir à l'Union Européenne était :

*« Data collection and coding TV and Newspapers media in the period of mass and elites surveys. »
(IntUne, work description, Annexe 1, 2005 :13)⁴*

Le corpus devait donc être préalablement élaboré à la même période que l'étude portant sur les élites et les masses (pour des raisons d'études comparatives déjà évoquées). Ce corpus devait également être collecté sous une forme électronique et rassembler des textes issus des médias mêlant à la fois journaux de presse écrite et journaux télévisés. Ensuite, il s'agissait d'indexer et de standardiser les corpus des quatre pays participant afin de pouvoir les utiliser et les analyser avec un logiciel approprié. Enfin, il convenait de donner des résultats quant aux analyses produites sur ce corpus. Nous allons rapidement préciser maintenant comment le corpus devait être construit, et quel type d'études devaient être menées selon la première réflexion du groupe média.

En règle générale, la construction du corpus selon le projet IntUne devait suivre trois principales étapes. La première concernait le téléchargement du matériau via Internet pour les journaux imprimés, et dans le cas de la télévision, la préparation de notes sur la situation du discours (locuteur, sexe, qualité professionnelle, etc.), puis la transcription du programme et sa transformation sous une forme électronique. La deuxième étape concernait la sélection et la compilation des quatre corpus de référence. Enfin, la troisième étape concernait la vérification, la standardisation et le formatage des corpus sous un format XML et TEI afin de produire des corpus indexés et unifiés de textes écrits et oraux pouvant être utilisés sur des logiciels (IntUne Work description 2005).

⁴ Le rassemblement et le codage de données issues des journaux télévisés et papiers à la même période que celle des enquêtes des groupes masses et élites. (ma traduction)

0- Introduction.

L'objectif fondamental du groupe média est d'utiliser les corpus électroniques afin d'identifier des tendances à grande échelle, des stratégies à long terme de la part des institutions et des développements dans les relations entre ces institutions et la masse des citoyens. Il s'agit aussi d'étudier comment la citoyenneté est construite dans des termes à la fois nationaux et européens, ce qui ne pourrait être atteint par une recherche à petite échelle, subjective, conventionnelle et non systématique. Il est précisé dans ce projet que les analyses des textes contenus dans le corpus doivent être menées en utilisant comme système de référence la linguistique systémique et fonctionnelle de Michael Halliday. Ce système voit le langage comme étant une pratique sociale et cherche à présenter un aperçu des systèmes, structures et fonctions du langage d'une façon qui tolère le côté inévitablement flou des catégories. C'est une théorie systémique, c'est à dire que cette grammaire est basée sur la notion de choix et que toute grammaire peut être pensée comme étant infinie et présentant un vaste réseau de choix. D'autre part, c'est une théorie fonctionnelle et ce dernier aspect est celui qui intéresse le plus le groupe média puisqu'il est engagé dans l'analyse de texte (Banks 2005). Les principales méthodes utilisées consistent en une combinaison des approches quantitatives et qualitatives : l'analyse classique de discours étant préconisée pour étudier des petits groupes de textes, et pouvant se combiner à une méthode d'analyse plus objective et statistique des textes écrits venant de la tradition théorique de la linguistique de corpus (IntUne Work description 2005).

Contenu de la thèse.

Pour mieux comprendre le projet et la ligne conductrice qui sous tendent notre réflexion, il paraît important d'expliquer le cheminement parcouru dans cette thèse, c'est l'objet de cette partie. Le premier chapitre, linguistique de corpus et linguistique générale, entame la réflexion sur la définition d'un corpus dans notre discipline. Il note d'une part les différents liens entre linguistique générale et corpus, montrant notamment que ces derniers ont longtemps été utilisés en linguistique, avant qu'une discipline n'en fasse l'objet. La relation entre linguistique et corpus est explicitée à travers une revue de la linguistique moderne, de Saussure à nos jours. Cette revue n'est pas une simple histoire de la linguistique, elle consiste en une mise en perspective qui permet de mieux comprendre pourquoi le corpus est davantage qu'un simple outil, et comment son utilisation sous forme informatique a pu conduire à l'émergence d'une nouvelle discipline. D'autre part, ce chapitre montre que

lorsqu'il s'agit du corpus, la réalité appréhendée peut grandement différer. Les différents types de corpus pouvant exister sont alors détaillés, afin de mieux situer notre propre corpus et les choix de recherche qui en découlent. On voit alors qu'il peut exister différents types de corpus, mais aussi différentes approches, et différentes applications dans cette discipline. Parmi ces différentes applications, nous nous attardons dans ce chapitre sur deux domaines principaux investis par les corpus avec succès, la traduction et l'enseignement de langue étrangère.

Le second chapitre, cadre théorique et méthodologie employée, s'attache à situer plus précisément l'environnement théorique des travaux menés dans cette thèse. La démarche de base du chercheur étant ici lexicographique, on voit qu'une vaste théorie existe déjà dans le domaine. Cette théorie est largement influencée par la pensée aristotélicienne qui a été appliquée avec plus ou moins de réussites dans la linguistique du vingtième siècle. Ce chapitre montre en quoi la théorie aristotélicienne présente de nombreux avantages, mais également les difficultés que celle-ci rencontre dans l'étude de la signification du langage. Différentes tentatives ont eu lieu pour résoudre ces difficultés. C'est le cas de la théorie du prototype ou de l'air de ressemblance qui est née d'une réflexion philosophique, et qui a été développée ultérieurement en psychologie puis en sémantique. Nous verrons là encore les atouts et les failles de ce modèle, et détaillerons une réponse qui a pu être apportée à ses lacunes à travers la théorie des prototypes sémantiques de Patrick Hanks. Ce lexicographe a donné une méthodologie pour l'étude de la signification se basant sur le modèle de l'air de ressemblance et des prototypes. C'est cette méthodologie qui est employée dans nos travaux. Enfin, le second chapitre donne un aperçu de l'application de la méthodologie utilisée dans la thèse à travers l'élaboration d'un premier prototype pour le mot « Europe ».

Le troisième chapitre revient sur les techniques employées lors de la constitution du corpus, ainsi que lors de son analyse. Il détaille d'une part, les différentes motivations du projet IntUne et les choix qui en ont découlé lors de la conception des corpus de presse et télévisés. D'autre part, il explique le fonctionnement et la structure de ce corpus à travers différentes illustrations. Enfin, ce chapitre donne des explications quant aux différentes manipulations et techniques qui ont été employées dans l'analyse de ce corpus.

Les chapitres quatre et cinq forment le cœur de la thèse, ils présentent les résultats des analyses qui ont été menées sur le corpus, ainsi que leur interprétation. Le quatrième

chapitre s'attache à l'analyse du concept « Europe » en développant un prototype basé sur une étude du réseau collocationnel de ce concept découvert à l'aide des collocations de ce mot ainsi que celles du lemme « européen ». Il détaille par la suite l'activation des différents sens qui forment ce prototype dans la presse française en 2007 et en 2009. L'analyse permet ainsi de voir comment l'Europe est représentée dans la presse française, et s'il y a une évolution sensible de cette représentation entre les deux années de collecte du corpus. Néanmoins, le second corpus n'ayant pu être constitué qu'à la moitié de l'année 2009, les analyses qui y ont été menées ne peuvent pas être aussi complètes que celles du corpus 2007, l'actualisation du prototype de « Europe » est en conséquence principalement vérifiée dans la presse de 2007.

Le chapitre cinq fonctionne sur le même principe, il s'attache à l'étude de deux autres pays dans la presse française : l'Italie et la Pologne. Un premier prototype est alors élaboré à l'aide du corpus pour les mots « nations » et « pays », sur le même mode de fonctionnement que celui qui est suivi dans le chapitre quatre, puis à l'aune de ces prototypes, l'activation de leurs sens pour les mots « Italie » et « Pologne » sont vérifiés dans la presse française en 2007 et en 2009. Cela permet alors de voir comment deux pays qui peuvent représenter l'Europe (à la fois de l'Est et de l'Ouest) peuvent être construits dans le discours des médias français, ainsi que les évolutions qui peuvent apparaître dans cette construction en deux années.

Enfin, le sixième chapitre s'intéresse à l'étude de deux reportages télévisés traitant du même sujet, concernant un pays européen, issus d'une chaîne publique et d'une chaîne privée françaises. La comparaison de ces deux reportages permet de voir les différents choix effectués dans la façon de représenter un pays d'Europe, ainsi que les différents traitements qui sont appliqués à un thème particulier. Cette étude enrichit en conséquence la vision que l'on a de la représentation que l'on a de l'Europe par la presse selon deux perspectives. La conclusion nous permet alors de voir ce qui a été appris dans cette thèse. D'une part, nous y reviendrons sur le nouvel angle de vision proposé par cette thèse dans le domaine de la linguistique générale, et plus particulièrement de la linguistique de corpus, l'étude ici menée permettant notamment de nouvelles perspectives en linguistique appliquée (enseignement de langues étrangères, traduction) . D'autre part, nous verrons quels apprentissages tirer de notre analyse en particulier dans la manière dont l'Europe est représentée à travers le discours qui

0- Introduction.

est tenu dans la presse française. Il sera alors possible d'envisager une nouvelle perspective concernant l'émergence d'une l'identité européenne.

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

I.0- Introduction.

La locution « linguistique de corpus » ou « corpus linguistics » pose d'emblée le corpus comme étant au centre de cette science.

Étant donné qu'il s'agit là d'une entité créée le plus souvent par le chercheur, il importe de se questionner quant aux modalités de production, aux choix, aux possibilités qu'impliquent cette création. Un cadre théorique sur le corpus existe déjà et peut donner des indications au chercheur, il va néanmoins s'agir avant toute étude sur corpus de savoir ce qu'est un corpus dès lors que l'on se situe dans le cadre de la linguistique actuelle, de connaître les différentes approches possibles dans la théorie, puis de s'approprier la conception de cet outil, en l'adaptant, dans la mesure du possible, aux besoins d'une étude menée.

Dans un premier temps, on cherchera donc à définir ce qu'est un corpus, sans pour autant entrer dans les détails des questions épistémologiques actuellement posées dans ce domaine (Rastier 2005, Jacques 2005, Alonso 2009, etc.), nous verrons ensuite quel usage il

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

en a été fait jusqu'à présent dans la linguistique, puis, on s'attardera davantage sur les différentes théories linguistiques qui ont pu se rapprocher de l'idée d'un travail en ayant ou non pour base un corpus, enfin, nous verrons en quoi la linguistique moderne a posé son lien avec le concept de corpus. Une deuxième grande partie de ce chapitre cherchera ensuite à détailler davantage les différents types de corpus qui peuvent être créés selon les buts de recherche que l'on se donne.

I.1/ Linguistique de corpus et linguistique générale.

I.1.1- Qu'est-ce qu'un corpus ?

Si nous nous interrogeons sur le sens que peut prendre ce mot, un dictionnaire comme le petit Robert nous en donne la définition suivante :

***CORPUS** n. m. (1863 ; "hostie", 1642 ; mot lat. "Corps"). 1. Dr. Recueil de pièces, de documents concernant une même discipline. Corpus d'inscriptions latines et grecques. 2. Ling. Ensemble limité des éléments (énoncés) sur lesquels se base l'étude d'un phénomène linguistique.*

Le Larousse en donne la définition suivante :

***CORPUS** n. m. (mot lat., corps). Didact. Ensemble de textes, de documents fournis par une tradition ou rassemblés pour une étude, en particulier pour une étude linguistique.*

Bien que restant très vague quant au mot lorsqu'il est utilisé en linguistique, la définition du petit Robert nous permet néanmoins de voir que c'est quelque chose qui recouvre l'idée d'un ensemble sur lequel va se baser une étude. La méthode de l'analyse ou ce qui compose cet ensemble ne sont pas abordés, ce qui pose un certain problème de clarté. Nous ne savons toujours pas ce qu'est concrètement un corpus. La deuxième définition, bien que ne comprenant que le sens linguistique du mot, n'est pas d'une plus grande utilité scientifique quant à ce qui compose le corpus, on y apprend simplement que les textes sont rassemblés dans le but de procéder à une étude. Il est utile de préciser ici que ces deux

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

définitions sont issues de dictionnaires basés sur une méthode introspective, n'ayant pas fait l'usage de corpus au sens que la linguistique moderne donne à cet outil. Il s'agit alors de « corpus » considéré dans le contexte de la linguistique générale française tandis que la linguistique de corpus est issue de la recherche britannique pour l'essentiel.

Selon McEnery et Wilson (1996:1), la linguistique de corpus est l'étude de la langue basée sur l'usage du langage de la vie réelle. Une dimension nouvelle apparaît alors, le corpus étant ici composé de langue « authentique ». Mais techniquement, cette définition ne donne pas de précisions sur ce qu'est ou non un corpus, sur la façon dont il est constitué concrètement.

Elena Tognini-Bonelli donne, quant à elle, du corpus la définition suivante :

« a collection of texts, assumed to be representative of a given language put together so that it can be used for linguistic analysis. (...) in general, there is a consensus that a corpus deals with natural, authentic language. » (Tognini-Bonelli, 2001:2)⁵

« a computerised collection of authentic texts, amenable to automatic or semi automatic processing or analysis » (Tognini-Bonelli, 2001: 55)⁶

L'idée d'une représentation authentique est donc bien essentielle au corpus, ainsi que celle d'un format informatique. Cette dernière notion est centrale dès lors que l'on parle de corpus en linguistique de corpus. C'est John Sinclair, théoricien majeur parmi les fondateurs de la linguistique de corpus qui a concrétisé cette idée d'une collection de textes informatisée, il en donne la définition suivante :

« Corpus has to have the material in electronic form » (Sinclair, 1991 : 14)

« ... a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample

5 Un ensemble de textes, acceptés comme étant représentatifs d'une langue donnée, et rassemblés afin de pouvoir être utilisés pour une analyse linguistique. En général, il y a un consensus sur le fait qu'un corpus est fait de langue naturelle, authentique. (ma traduction)

6 Un ensemble informatisé de textes authentiques, pouvant être analysés ou traités de façon automatique ou semi automatique. (ma traduction)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

of the language. (Sinclair 1996) »⁷

La nécessité pour un corpus d'être sous forme informatique lorsque l'on se situe dans le cadre de la linguistique de corpus moderne est alors clairement énoncée, d'autre part, il s'agit selon Sinclair d'un ensemble de textes qui sont organisés et sélectionnés, et ce afin de parvenir à une représentativité au moins relative d'un langage ou sous langage. Le corpus n'est pas un simple ensemble d'archives ni tout type de texte, de document que l'on pourrait trouver sous forme numérique, il doit obéir à des critères qui sont élaborés en vue d'une étude linguistique. La forme informatique du corpus est la réelle nouveauté en ce qui concerne la linguistique de corpus, puisque au cours des siècles précédents, de nombreux chercheurs avaient déjà travaillé sur des ensembles de textes dans un but linguistique. Ainsi, selon la tradition linguistique dans laquelle le chercheur se situe, le corpus pourra recouvrir de multiples réalités : l'objet corpus donnera un aperçu différent selon que l'on se situe dans telle ou telle école. La définition de travail que nous donnons du corpus s'inspire largement de celle de John Sinclair : le corpus est un ensemble de textes organisés et sélectionnés en fonction d'un but de recherche linguistique, regroupés de façon informatisée et permettant son analyse de façon automatique ou semi-automatique. Cette définition sera notamment affinée dans la seconde partie du chapitre.

Mais avant cela, il pourrait être utile de voir en quoi cet outil s'insère pleinement dans une réflexion qui s'est déroulée tout au long de l'histoire de la linguistique moderne. Dans la partie suivante, on rappellera dans un premier temps le fait que la pensée épistémologique de Ferdinand de Saussure s'est développée dès la fin du dix-neuvième siècle sur un support reconstruit et publié de façon posthume, support à travers lequel cette pensée a été véhiculée, pour l'essentiel, tout au long du vingtième siècle. Nous verrons également comment différentes interprétations de cette pensée fondatrice ont pu être menées à partir de la découverte d'écrits de la main de Saussure. Ensuite, nous verrons en quoi ces différentes interprétations ont pu mener à des questionnements et des pratiques différentes tout au long du vingtième siècle, d'abord par la pensée chomskyenne qui a beaucoup influencé les chercheurs dans le domaine de l'intelligence artificielle, puis, par le développement de l'école contextualiste britannique qui, développée par John Sinclair (1987, 1991) a donné naissance à la discipline de la linguistique de corpus. Enfin, nous montrerons comment les courants

⁷ Un ensemble de morceaux de langue qui sont sélectionnés et ordonnés selon des critères linguistiques explicites afin d'être utilisé comme un échantillon de la langue. (ma traduction)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

sausurriens et contextualistes peuvent être réunis à travers la synthèse proposée par Elena Tognini-Bonelli (2001)

I.1.2- La relation entre linguistique et corpus.

Ferdinand de Saussure est fréquemment considéré (d'autant plus en Europe) comme le père de la linguistique moderne. Son *Cours de linguistique générale*, paru aux éditions Payot de façon posthume en 1916 a connu une fortune rare attestée par de nombreuses traductions et interprétations. Selon Simon Bouquet (2005), cette dimension fondatrice de l'œuvre saussurienne prend sa source dans diverses causes. Ce n'est certes pas Ferdinand de Saussure qui emploie en premier le mot « linguistique » pour désigner une discipline qui naît au cours du dix-neuvième siècle, mais dès ses premiers pas dans ce domaine, il se taille une place de première importance grâce à son *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Il abandonne néanmoins cette réflexion comparatiste pour peu à peu se lancer dans une réflexion épistémologique nouvelle pour la linguistique moderne dont le *Cours de linguistique générale* deviendra la trace écrite vulgarisée. C'est à travers cet ouvrage que l'on considère généralement Saussure comme le père fondateur de la linguistique moderne. C'est aussi sur cet ouvrage que reposeront les différentes interprétations qui dirigeront la linguistique du vingtième siècle, que l'on pense au structuralisme, ou même au courant logico-grammatical (Rastier 2005).

Pourtant, ce *Cours de linguistique générale* n'a pas été écrit par Saussure lui-même. Il s'agit en fait d'une « reconstruction » par Bally et Séchehaye des notes de cours prises par les élèves du maître lors de ses leçons données en trois temps au Collège de Genève, avec le concours de l'un de ces élèves. Cette reconstruction peut alors prêter à polémique (Bouquet 2005) et en conséquence, les interprétations qui en ont été faites et qui ont influencé une partie de la linguistique au long du vingtième siècle peuvent tout autant être remises en question.

I.1.2.1- Saussure.

L'écrit qui a le plus contribué à vulgariser et répandre la pensée saussurienne est donc le *Cours de linguistique générale*, publié de façon posthume à partir de notes prises par les élèves du linguiste en 1916. D'autres écrits ont néanmoins par la suite été retrouvés,

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

notamment des sources de ce cours dans les années 1960, et des notes relativement complètes préparatoires à la publication d'un livre par Saussure regroupées sous le titre d' *Écrits de linguistique générale*, retrouvés en 1996 dans l'orangerie familiale de la famille de Saussure à Genève, et publiés en 2002 aux éditions Gallimard par l'institut Saussure.⁸ Selon que l'on base son interprétation de la pensée linguistique du maître sur le Cours publié en 1916 ou sur les Écrits publiés en 2002, c'est un Saussure sensiblement différent qui apparaît.

1.1.2.1.1- Le Cours de linguistique générale.

Saussure, dans le *Cours de Linguistique Générale* (Payot, 1916) notait que la linguistique devait avoir pour objet la langue, qu'il différenciait de la parole qui en était la réalisation concrète et partielle. La langue étant un fait social, une norme que chaque individu percevait à travers la masse, et la parole étant l'actualisation de la langue par chaque individu. Le linguiste avait dans cet ouvrage sa tâche clairement définie, la langue devait être étudiée séparément de toute autre considération, formant un système et ne pouvant être observée qu'à travers sa manifestation concrète mais imparfaite : la parole. Le langage se devaient de communiquer avec ces deux parties d'un même tout, selon un principe de vases communicants. La dichotomie chomskyenne de la compétence et de la performance que l'on a souvent associée à la distinction langue / parole est en fait une radicalisation de cette distinction : la compétence n'a quasiment plus rien à voir avec la performance, et l'étude de la performance pour atteindre la compétence est quelque chose de relativement complexe, qui demande des dons particuliers pour avoir une chance d'être réussie, ces deux entités sont plus ou moins coupées l'une de l'autre chez Chomsky.

Le *Cours de linguistique générale*, tel qui a été « reconstruit » par Bally et Séchehaye fait néanmoins en quelque sorte de Saussure un linguiste s'intéressant principalement à la langue, et de façon moindre à la parole, notamment à travers la dernière phrase du livre qui donne pour objet à la linguistique « la langue en elle-même et pour elle-même », bien que cette citation soit en fait parfaitement apocryphe et qu'on la doive à F. Bopp en 1816 (Rastier 2005).

La méthode d'approche du linguiste préconisée dans le Cours se voulait être la plus scientifique possible dans la mesure où il cherchait avant tout à circonscrire de la façon la

8 [Http://www.institut-saussure.org](http://www.institut-saussure.org)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

plus précise qui soit son objet d'étude. En rendant l'objet d'étude de la linguistique abstrait, il lui était également donné en quelque sorte une certaine caution scientifique, le rapprochant de sciences bien établies comme les mathématiques, qui fonctionnaient essentiellement dans l'abstraction. Cela faisait ainsi écho à un mouvement de la science à la fin du dix-neuvième siècle qui consistait à aller vers toujours plus d'objectivité et de scientificité pour atteindre le savoir, en associant ces idées de scientificité et d'objectivité à l'abstraction (Crystal 1985:144-156).

Comme le note Rastier, on a retenu à partir de ce Cours des dichotomies plus que des « dualités » (Rastier 2005) comme résumant la pensée de Saussure, ces dichotomies étant les oppositions classiques entre langue et parole, signifiant et signifié, synchronie et diachronie et axes paradigmatique et syntagmatique. Saussure y est présenté comme le constructeur d'une épistémologie de la linguistique qui devrait se préoccuper de la langue, en accordant davantage d'importance au signifié (l'esprit) qu'au signifiant (la matière) quand bien même on y reprend la métaphore saussurienne du signe linguistique comme un tout indivisible, tel le recto et le verso d'une feuille (Rastier 2005).

1.1.2.1.2- Les écrits de la main de Saussure.

Cependant, parmi les divers écrits retrouvés de Saussure (les notes préparatoires de ses cours au collège de Genève, ses notes préparatoires à la publication d'un livre de linguistique générale retrouvées en 1996, divers courriers) donnent à voir un aspect sensiblement différent de la pensée du maître.

Ainsi, si le cours peut laisser penser que Saussure était un linguiste de la langue, Rastier (2005) tempère ce point de vue :

« Or, pour Saussure, la parole est l'élément déterminant dans la dualité langue / parole. Dans son discours à l'occasion de la création de la chaire de pour Bally, il dit ainsi de la linguistique : « elle comporte deux parties : l'une qui est plus près de la langue, dépôt passif, l'autre qui est plus près de la parole, force active et véritable origine des phénomènes qui s'aperçoivent ensuite peu à peu dans l'autre moitié du langage. Ce n'est pas trop que les deux » (ELG, p. 273). » (Rastier 2005)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

La fameuse dernière phrase apocryphe du Cours est en conséquences pour le moins relative, et cette citation par Rastier des *Écrits de linguistique générale* montre à quel point Saussure était un linguiste de la parole. De nombreuses autres citations de Saussure ou de ses élèves, qu'il s'agisse de Regard ou de Riedlinger, le collaborateur de l'édition du Cours (Bouquet 2005), montrent ce même fait d'une pensée saussurienne qui aurait ainsi été faussée en faveur d'une linguistique plus abstraite, paraissant de cette façon plus scientifique. Cette vision postérieure d'un Saussure donnant une importance forte à la parole comme force active du langage et qui la réconcilie avec le concept de langue est reprise dans cette thèse : la langue ne doit pas être mise de côté, pas plus que la parole saussurienne, c'est bien à travers l'analyse de la seconde que l'on peut tirer certaines généralisations intéressantes la première.

Selon Rastier, plus que de dichotomies, c'est de « dualités » (Rastier 2005) dont il est question dans la pensée saussurienne : appréhender une partie de la dualité sans son pendant semble inefficace, parler de la langue sans longuement se pencher sur la parole le semble tout autant. Le fonctionnement de la langue telle qu'elle a été pensée par Saussure n'est pas selon Rastier à décrire dans les termes aristotéliens d'une grammaire séparée de son application, d'une langue séparée d'une parole ou d'une compétence séparée d'une performance, mais bien plutôt au niveau d'une parole alors séparée de la langue parce que Saussure, bien que très innovant dans sa pensée, n'avait pas encore eu les moyens suffisants pour étudier un espace des normes, faisant la jointure entre ces deux entités d'une même dualité.

Néanmoins, toute la linguistique du vingtième siècle ne va cesser de se réclamer de la parenté saussurienne, qu'il s'agisse des différents courants davantage axés sur une étude de la langue, du système, de l'abstrait, ou qu'il s'agisse de courants se basant sur les faits de langues authentiques observés dans leur multitude. Un rapide portrait va maintenant être brossé de ces différents courants pour voir comment ils ont pu influencer dans leur diversité les multiples pratiques et réflexions actuelles en linguistique de corpus.

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

I.1.2.2- Une linguistique de la langue.

I.1.2.2.1- Le structuralisme.

Ce que l'on a appelé le courant structuraliste au long du vingtième siècle s'est réclamé de la paternité de Saussure. La notion de structure trouve sa naissance dans la pensée saussurienne, selon Crystal (1985) :

« it is largely on account of Saussure that the idea of structuralism achieved the status which was to make it the major linguistic theme of the next thirty years. » (Crystal 1985:164)⁹

Les structuralistes présents à l'école de Prague ont cherché à décrire le système dont avait parlé Saussure, notamment à travers des études phonologiques du système en ce qui concernait Jakobson et Troubetzkoy (Crystal 1985:177). En fait, il s'agissait de considérer que la langue pouvait se décrire comme un système présentant des relations d'opposition et de rapprochement, Jakobson a notamment développé l'idée du phonème comme un concept abstrait reposant sur l'opposition (Jakobson 1973:123).

Aux États-Unis, toujours au début du vingtième siècle, un courant structuraliste s'est également développé, du fait de la découverte de pléthore de langues amérindiennes en voie d'extinction qu'il fallait pouvoir « enregistrer » de la façon la plus efficace et la plus objective possible. Bloomfield puis Harris ont contribué à cette linguistique. Bloomfield a élaboré un type d'analyse linguistique qui fonctionnait en constituants immédiats (ACI). Dans l'ACI, le langage n'est pas vu comme un ensemble fonctionnel, mais l'analyste cherche davantage à identifier et classer et segmenter les constituants d'une phrase, quelle que soit leur fonction, sans se préoccuper, finalement de l'ensemble de la phrase (Crystal 1985:204). Le mot formel est essentiel dans cette théorie, et il a influencé une grande partie des années post-bloomfieldiennes, or théoriquement, ce mot n'a aucun rapport avec l'idée de signification. Bloomfield n'avait aucune confiance en l'intuition du linguiste, il ne voulait pas que les constructions linguistiques aient un statut psychologique. Selon sa théorie, la signification ne pouvait s'approcher qu'au travers de configurations distinctives de stimuli et réponse béhavioriste (Ibid:234). En l'absence d'outils pour l'appréhender ou la classer

⁹ L'idée du structuralisme s'est essentiellement développée sur la base de Saussure et a accédé ainsi au statut qui en a fait un des thèmes linguistiques majeurs des trente années qui ont suivi. » (ma traduction).

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

précisément, Bloomfield a ainsi ôté la signification de ses recherches (Ibid:205). Il a en fait traité cette problématique en notant que la signification, plus qu'à la linguistique, était une tâche à réserver à d'autres disciplines ou d'autres chercheurs comme les zoologistes, par exemple, pour expliquer le sens de tel nom de plante (Bloomfield 1933:139). La question plus générale de la signification revenait de même davantage aux philosophes et logiciens qu'aux linguistes dans la pensée bloomfieldienne. Catford note que cela ne semble pas fonctionner, les définitions par des philosophes, logiciens ou autres scientifiques non spécialisés dans un domaine linguistique étant en effet le plus souvent peu pertinentes dans un usage quotidien de la langue (Catford 1983:15-6). En effet, lorsque l'on utilise un mot, cela n'a que rarement un rapport étroit avec la réalité scientifique de la chose nommée, le rapport s'établirait davantage avec la façon dont le locuteur conceptualise dans une situation précise la chose dont il parle (Ibid:14). En outre, séparer la question de la signification de la question de la description morphologique d'une langue, comme le préconisait Bloomfield, apparaît problématique à partir du moment où la signification est elle aussi systémique, c'est-à-dire que les significations de mots ou même de catégories grammaticales sont délimitées en fonction de leur rapport d'opposition à des significations d'autres mots ou d'autres catégories.

« language is systemic, and that the meaning of any particular word, and also of any grammatical category, are delimited by their opposition to other words or categories within the same system. This is particularly true of grammatical words and categories, which operate in strictly closed systems. But even in lexis, where words often belong to more loosely constituted open sets, there are clear cases of the delimitation of meaning by opposition to other members of the set. »¹⁰ (Catford 1983:23)

De même, au moment où Bloomfield arguait que la signification était un des points faibles de l'étude de la langue, Firth (1935) publiait un article majeur sur les techniques employées en sémantique et insistait sur le fait que la linguistique amputée de la signification

10 « La langue est systémique, et la signification de tout mot particulier ainsi que celle de toute catégorie grammaticale, est déterminée par leur opposition aux autres mots et catégories grammaticales au sein d'un même système. Cela est particulièrement vrai des mots grammaticaux et des catégories, qui opèrent dans des systèmes strictement étroits. Mais même au niveau de la lexie où les mots appartiennent à des ensembles ouverts moins fermement définis, on trouve des cas clairs de significations délimitées par opposition à d'autres membres de l'ensemble » (ma traduction).

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

devenait une discipline stérile (Stubbs 1996:35). Nous reviendrons plus tard sur ce courant linguistique qui a continué à travailler sur la signification, mais continuons maintenant avec la suite des linguistes structuralistes qui ont, à la suite de Bloomfield, exclu la signification de leur recherche.

Zellig Harris a poursuivi cette pensée bloomfieldienne en cherchant à voir comment ces constituants immédiats se répartissaient dans la phrase. Il recherchait en fait la distribution des éléments de la phrase. Pour ce faire, il se basait sur des observations de phrases qu'il analysait ensuite selon la fréquence avec laquelle tel ou tel constituant se distribuait auprès de tel autre constituant. Selon lui, la sémantique devait être considérée comme incluse dans la grammaire, cependant, ses études se sont essentiellement basées sur les classes de mots grammaticales qui étaient bien développées et relativement stables, contrairement à la dimension sémantique. Il cherchait en fait à décrire le sens de la phrase en notant une forme de représentation qui ne soit plus ambiguë. Pour réaliser cette représentation, il a travaillé sur les dépendances entre les différents mots d'un énoncé, les relatant sous la forme d'une transformation linéaire. La transformation est utilisée par Harris dans le sens d'une représentation qui préserve les combinaisons linéaires (Goldsmith 2004:721).

1.1.2.2- Le courant logico-grammatical.

En réaction au descriptivisme de la théorie distributionnelle, un nouveau courant est né, toujours aux États-Unis, notamment sous la direction de Noam Chomsky dans les années 1960. Chomsky a été l'élève de Zellig Harris. La représentation en analyse par constituants immédiats utilisant la transformation de Harris a fortement inspiré son élève dans les années 1950.

Chomsky, à partir de la fin des années 1950, a défini le paradigme compétence/performance qui différencie d'une part, la capacité de construire et de reconnaître l'ensemble des énoncés grammaticalement corrects d'une langue pour un locuteur (la compétence) et d'autre part, l'ensemble effectif des énoncés produits (la performance). Il souhaitait décrire la performance avant toute chose, s'inscrivant ce faisant dans une linguistique davantage axée sur la « langue » telle qu'on l'avait interprétée dans la pensée saussurienne. Il a ainsi radicalisé cette pensée. Nous avons vu précédemment que Saussure

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

avait distingué langue et parole, mais que les deux entités communiquaient ensemble, que l'une informait l'autre dans la recherche du linguiste. Ce n'est plus le cas chez Chomsky, l'étude du linguiste doit concerner uniquement la compétence au détriment de la performance :

« it is absurd to attempt to construct a grammar that describes observed linguistic behaviour directly » (Chomsky 1964a)

Pour ce faire, il a développé une grammaire générative transformationnelle, le mot « transformationnelle » semble ici hérité de Harris, néanmoins, le sens qu'il recouvre est bien différent de celui d'une transformation linéaire. Selon Chomsky, la performance se décrit essentiellement par une composante syntaxique : une base de connaissances partagées universellement par tout être humain est une première fois organisée en une structure profonde, des transformations se font sur celle-ci et on aboutit ainsi à une structure de surface qui est la chaîne phonétique que l'on entend lors d'un acte langagier. La composante sémantique devait être prise en compte au sein de la structure profonde, mais Chomsky ne s'est jamais réellement attaché à son étude, cette composante s'adaptant assez mal à son modèle génératif transformationnel (Crystal 1985:101-110).

Pour développer son modèle, Chomsky a postulé que l'introspection du linguiste était la première source fiable d'information. Il a en conséquence affirmé l'inutilité d'une étude se basant sur des faits de langue authentique (Hoey 1983:31) si ce n'était dans l'étude de langages inconnus (de type amérindiens, par exemple), et en a fait une stratégie au mieux sans intérêt, au pire dangereuse, préférant alors à l'approche empirique une approche selon lui plus rationnelle, et en réalité basée sur l'introspection avec des procédures de découverte déductives.

Chomsky souhaitait réaliser un modèle qui puisse rendre compte de tous les systèmes de langage humains. Il s'est ainsi opposé à l'analyse de corpus de langue particuliers aboutissant à des généralisations, un corpus ne pouvant jamais être représentatif d'un ensemble plus grand selon lui. Le langage étant infini, la production de langage croissant chaque jour, il est tout simplement impossible de créer un corpus qui puisse représenter la langue dans son ensemble, aucun corpus ne peut contenir toutes les occurrences possibles d'une langue. Il s'agit en conséquence d'un outil qui va dévier le

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

chercheur en lui donnant une vision fautive de ce qu'est la langue : certaines occurrences rares n'apparaîtront pas dans le corpus, et donc le linguiste en déduira que ces occurrences n'existent pas dans la langue.

*« There are, in other words, very few reliable experimental or data-processing procedures for obtaining significant information concerning the linguistic intuition of the native speaker. »
(Chomsky 1965 :18)¹¹*

En conséquence de cela, plutôt que d'utiliser des probabilités qui auraient été obtenues à l'aide d'une observation d'actes de langages authentiques, ou de la performance, Chomsky préférait se baser sur des certitudes syntaxiques obtenues selon une méthode hypothético-déductive. En effet, selon lui, la performance est par essence quelque chose d'imparfait, une sorte de représentation famélique et très appauvrie de la compétence, qu'il convient d'étudier. A partir de là, utiliser un corpus pour décrire la compétence risquait fortement de biaiser la conscience que le linguiste aurait eu du langage étudié :

*« It is almost impossible to generate a corpus without going beyond it. Any natural corpus will be skewed. Some sentences won't occur because they are obvious, others because they are false, still others because they are impolite. The corpus, if natural, will be so wildly skewed that the description would be no more than a mere list. (...) it would be a miracle if you could give any description shorter than the corpus itself. »
(Chomsky 1962 : 159)¹²*

De plus, les notions de fréquence de telles ou telles unités d'un langage, que l'on peut identifier précisément à l'aide des corpus, n'étaient selon lui pas utiles lorsque l'on souhaitait décrire la langue. En effet, la grammaticalité d'un énoncé est quelque chose d'essentiel dans la théorie chomskyenne, or il pose une indépendance totale entre cette grammaticalité et le fait qu'un énoncé apparaît fréquemment dans tel ou tel environnement :

11 « Il y a, en d'autres termes, très peu de procédés de traitement de données ou d'expériences qui soient fiables pour obtenir des informations significatives, en ce qui concerne l'intuition linguistique du locuteur natif. » (ma traduction)

12 « Il est presque impossible de créer un corpus sans en aller au-delà. N'importe quel corpus naturel sera biaisé. Certaines phrases n'apparaîtront pas car ce sont des évidences, d'autres, parce qu'elles sont fausses, et encore beaucoup d'autres, parce qu'elles sont impolies. Le corpus, s'il est naturel, sera biaisé à un tel point que la description ne pourra être rien plus qu'une simple liste. Ce serait un miracle si vous pouviez donner une description, quelle qu'elle soit, qui fut plus courte que le corpus lui-même. » (ma traduction)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

« There is no significant correlation between order of approximation and grammaticalness. If we order the strings of a given length in terms of order of approximation to English, we shall find both grammatical and ungrammatical strings scattered throughout the list . » (Chomsky, 1957 : 116)¹³

Ces arguments ont conduit certains universitaires et laboratoires à un basculement de la méthode empirique en linguistique générale vers un raisonnement plus rationnel (rationnel prend alors le sens de quelque chose se fondant avant tout sur le raisonnement) se basant sur l'introspection du chercheur. Si certains linguistes ont continué dans des domaines particuliers de la linguistique à utiliser des corpus (dans des domaines concernant l'acquisition de la langue par des enfants, par exemple, et dans la plupart des recherches en linguistique appliquée), de nombreuses recherches en linguistique générale ont alors adopté le modèle syntaxique de Chomsky, renonçant à des études qui se seraient basées essentiellement sur des faits de langue authentique. Hoey (1983:37) a notamment montré à travers une étude stylistique de certains des écrits de Chomsky comment la mise en place particulière de l'évaluation au sein de ses écrits scientifiques a pu conduire à cette dépréciation de toute étude qui se serait basée sur des faits de langue authentiques, sur la performance.

D'autre part, la théorie chomskyenne a continué à intéresser grandement les chercheurs qui souhaitaient modéliser la langue : cette théorie permettait en effet d'élaborer des modèles compréhensibles pour un ordinateur à partir du moment où une majeure partie des problèmes de variation ou d'instabilité de la langue orale n'était plus présente. Néanmoins, ce développement d'une linguistique descriptive se basant d'une façon exclusive sur l'étude de la « langue », interprétée comme étant la puissance ou la « performance » chomskyenne, n'a pas empêché la poursuite et le développement d'une autre tradition en parallèle, cette fois basée sur des actes de langages authentiques par des linguistes issus majoritairement des sciences appliquées, mais ayant également de nombreuses préoccupations descriptives.

13 « Il n'y a pas de corrélation significative entre la fréquence et la grammaticalité. Si nous classions des chaînes d'une longueur donnée en fonction de leur fréquence en anglais, nous trouverions à la fois des chaînes grammaticales et agrammaticales répandues à travers notre liste. » (ma traduction)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

1.1.2.3- Une linguistique de la parole.

Avant même le *Cours général de linguistique*, à la fin du dix-neuvième siècle, on peut trouver les prémices d'une linguistique qui se préoccupe d'étudier des faits de langues authentiques à travers l'usage de corpus. Cette tradition se situe le plus souvent dans une démarche de linguiste appliqué, mais peut également concerner des linguistes qui, tel Sweet, seront intéressés à la fois par la linguistique appliquée et sa dimension descriptive (Véronique 1992:180).

Cependant, une tradition descriptive existe depuis longtemps en Grande Bretagne, ainsi, dès le dix huitième siècle, en 1755, un dictionnaire est élaboré à partir de fiches de travail comportant des citations authentiques par Samuel Johnson (Williams 2006:152). Ce lettré anglais avait pour but d'écrire un dictionnaire prescriptif de l'anglais le plus complet qui soit et souhaitait rendre compte de la façon dont on le parlait. Rapidement, il s'aperceva du besoin de se baser sur des textes existants, pour illustrer de citations ses définitions. Cet ouvrage pourra ouvrir la voie à de nombreux autres dictionnaires se basant sur une étude de corpus (informatisés ou non). Cela montre aussi qu'au dix huitième siècle, quand bien même la plupart des dictionnaires se fondaient sur la perception qu'avait le lexicographe de la langue décrite, un besoin se faisait sentir de réaliser des descriptions plus en adéquation avec la réalité en s'aidant d'analyses et d'exemples extraits de faits de langues authentiques.

1.1.2.3.1- Les sources du contextualisme britannique.

Le domaine de l'enseignement des langues, et plus particulièrement des langues étrangères a été l'un des plus prolifiques en ce qui concerne la recherche linguistique se basant sur des faits de langue authentique. Les chercheurs spécialistes de la question de l'enseignement en langue sont souvent ceux qui, dans le même temps, ont pensé les termes et l'importance dans le domaine de la linguistique générale, d'une étude se basant sur des faits de langues authentiques. Nous allons maintenant voir que deux chercheurs impliqués dans l'enseignement des langues étrangères ont particulièrement contribué à donner naissance à une réflexion nouvelle qui permettra ensuite l'émergence du contextualisme britannique : Henri Sweet et Harold Palmer.

A la fin du dix neuvième siècle, et au début du vingtième, Sweet a apporté une grande contribution à la linguistique, notamment en phonétique. Après avoir réalisé de

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

nombreux travaux dans le domaine du vieil anglais dans la première moitié des années 1870, il s'est plus particulièrement intéressé à la phonétique et a notamment publié en 1877 un *Handbook of Phonetics* qui a eu une résonance au niveau international, et lui a valu une réputation de père fondateur de la phonétique dans toute l'Europe. Cette réflexion a eu lieu dans un contexte particulièrement favorable à l'émergence de nouvelles pratiques linguistiques. De nouvelles technologies sont en effet apparues, telles que l'avènement du magnétophone qui rendait possible l'enregistrement de discours oraux (D. Véronique 1992:174). Sweet a souhaité fonder une nouvelle science du langage reliant linguistique appliquée et descriptive. A travers son livre, *Practical Study of Languages* publié en 1899, il a énoncé différents principes se basant sur la nature du langage pour en permettre l'apprentissage. Il a notamment étudié le lexique et la phraséologie (Williams 2006:152), et proposait dans le domaine de l'apprentissage grammatical d'une langue de procéder de façon inductive, en confrontant tout d'abord les apprenants à des textes authentiques dont ils tireraient les principes grammaticaux de la langue (Véronique 1992:180). Le contexte est donc essentiel dans les principes de didactique du langage de Sweet.

Harold Palmer a également travaillé dans ce sens au début du vingtième siècle. C'est-à-dire que dans ses préoccupations de linguiste appliqué dans le domaine de l'apprentissage et de l'enseignement de l'anglais, il s'est beaucoup intéressé à la langue orale. Il s'est en effet illustré dans des travaux concernant la linguistique appliquée et plus particulièrement dans le domaine de l'apprentissage de la langue anglaise dans la première moitié du vingtième siècle. Il a effectué de nombreux travaux au Japon, où il a demeuré quatorze ans (à partir de 1922) comme « linguistic adviser » du ministère japonais de l'éducation et où il a réformé cette discipline en fondant, par exemple, l'institut pour la recherche en enseignement de l'anglais, concernant son l'apprentissage pour les étrangers. Ces ouvrages sont restés comme des référence dans la réflexion en didactique de l'anglais (Smith 1999:7-9). Lors de ses travaux au Japon, il s'est ainsi penché plus particulièrement sur la lexie et la mise en place d'un vocabulaire pour les apprenants. Une fois de retour au Royaume-Uni, (à partir de 1936), il a travaillé comme consultant pour les éditeurs Longman et Green. Il a aussi publié en 1937 un travail sur la lexie avec Hornby, créateur du premier *Advanced Learner's Dictionary*. Il a également produit un rapport traitant de la question des collocations, réemployant ainsi un terme déjà présent chez Sweet (Williams 2006:153). Il portait dans ses travaux l'idée que la langue est d'abord le véhicule des idées et condamnait

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

une « attention excessive de l'ancienne méthode à la construction grammaticale » (Véronique 1992:176).

Sweet et Palmer, de par leurs travaux en linguistique appliquée, ont soulevé des pistes d'étude intéressantes dans le domaine de la linguistique descriptive, notamment par l'intérêt porté au contexte des mots dans la création du sens. cependant, c'est avec Firth qu'une véritable école contextualiste sera développée en linguistique générale, notamment à travers l'étude du phénomène collocationnel (Williams 2006:153).

I.1.2.3.2- L'école contextualiste britannique.

Dans la suite de la première moitié du vingtième siècle, l'importance du contexte a été mis en avant par des linguistes britanniques ayant adapté une partie de la théorie anthropologique de Malinowski. John R. Firth est un linguiste de cette école, qui fut parmi les premiers à adapter les méthodologies fonctionnaliste et contextuelle à la linguistique générale, devenant ainsi le premier professeur de linguistique générale en 1944 à Londres (Honeybone 2005). Selon Malinowski, toute institution humaine, toute élaboration humaine (et le langage en fait explicitement partie selon lui) est lié à des besoins fonctionnels, et ne peut donc être comprise qu'au sein d'une culture, d'un contexte. De la même façon que l'on ne peut expliquer tel rite religieux ou de parenté qu'en fonction de tel ou tel contexte, de telle ou telle culture, on ne peut selon Malinowski expliquer tel ou tel énoncé que dans son contexte :

« La culture, où se tient le plus clair de la conduite humaine, est également importante pour tout le monde, pour le psychologue, pour le sociologue, pour l'historien, pour le linguiste. A mon avis, la linguistique de demain, et notamment la sémantique, sera l'étude de la langue dans le contexte d'une culture. » (Malinowski 1968:11)

Il ajoute plus loin :

« Nous avons posé en principe que toute théorie scientifique doit partir de l'observation et y conduire. Elle doit être inductive et se vérifier dans l'expérience » (ibid:60)

Firth s'est inspiré de ces travaux pour alimenter sa réflexion sémantique. Après

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

avoir appelé de ses vœux une théorie ethnographique de la langue, il va s'attacher à l'élaborer. Des années 1930 à 1950, il a ainsi postulé une théorie contextuelle du sens basée sur une procédure inductive, et en conséquent, sur une étude de phénomènes concrets, que l'on puisse appréhender. Il voyait le langage comme la répétition d'un procédé social, le travail du linguiste consistait ainsi à pouvoir dériver de l'impersonnel à partir du personnel qui était rendu typique. Ce passage du personnel à l'impersonnel était possible grâce à certaines structures, comme les collocations (occurrences répétées et signifiantes de deux mots ensembles) ou les colligations (la proximité répétée de même parties du discours).

« We must expect therefore that linguistic science will also find it necessary to postulate the maintenance of linguistic patterns and systems (including adaptation and change) within which there is order, structure, and function. Such systems are maintained by activity, and in activity they are to be studied. »¹⁴(Firth 1957:143)

Pour pouvoir décrire le langage, Firth se basait sur une observation de la langue, et non pas sur son introspection. D'autre part, il a donné toute son importance à l'étude du sens, comme étant accessible à travers des usages contextualisés, et en notant que les manifestations de la langue étaient observables plutôt que pré-supposables ou ontologiques. Le langage n'était pas dans sa théorie considéré comme un procédé mental ou une entité autonome, mais plutôt comme des événements exprimés par des locuteurs, comme un mode d'action.

« As we know so little about mind and as our study is essentially social, I shall cease to respect duality of mind and body, thought and word, and be satisfied with the whole man, thinking and acting as a whole, in association with his fellows. (...) Meaning, that is to say, is to be regarded as a complex of contextual relations, and phonetics, grammar, lexicography, and semantics, each handles its own components of the complex in its appropriate context. »¹⁵ (Ibid:19)

14 « Nous devons en conséquence attendre de la linguistique qu'elle trouve également nécessaire de postuler le maintien de structures et de systèmes linguistiques (qui incluent les adaptations et les changements) au sein desquels il y a un ordre, une structure, et une fonction. De tels systèmes sont maintenus grâce à l'activité, et c'est via cette activité qu'ils doivent être étudiés. » (ma traduction).

15 « Comme nous ne savons que très peu de choses sur ce qui se passe dans notre esprit, et étant donné que notre étude est essentiellement sociale, je ne reprendrai pas la dualité entre esprit et corps, pensée et mot, mais je me satisferai de l'homme dans son ensemble, qui pense et qui agit comme un tout, en association avec ses pairs. (...) Le sens, en conséquence, doit être considéré comme un ensemble complexe fait de relations contextuelles, de phonétique, de grammaire, de lexicographie et de sémantique ; chacune apportant

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

Selon lui, ces méthodes purement inductives devaient apporter beaucoup à l'étude de la langue, notamment par l'apport de nouvelles catégories permettant de décrire plus scientifiquement, plus précisément la langue :

« so that the situational and linguistic categories would not be unmanageable. Many new categories would arise from a systematic observation of the facts »¹⁶ (Ibid:28)

Il n'hésitait pas à dire que pour être scientifique, la linguistique se devait avant tout de fonctionner de façon inductive, empirique, sans quoi elle risquait de s'abîmer dans des hypothèses fallacieuses, ne rendant pas compte de la réalité du langage. L'influence de Sweet, qui lui même souhaitait que les apprenants procèdent par découverte en contexte de la grammaire après la lecture d'un texte, est présente lorsque Firth explique que la linguistique ne doit pas s'éloigner de la parole en la forçant à entrer dans des systèmes artificiels.

« Sciences should not impose systems on languages, it should look for systems in speech activity, and, having found them, state the facts in a suitable language. » (Ibid:144)¹⁷

Son travail a été largement reconnu en Grande Bretagne, que ce soit par le gouvernement (suite par exemple à l'élaboration de cours accélérés d'apprentissage du japonais au cours de la seconde guerre mondiale), ou par la communauté linguistique, il a ainsi été président de la Philological Society (Honeybone 2005:84). Il a influencé certains linguistes des années suivantes tels que Halliday qui écrira dans les années 1950 une thèse où il s'intéressait aux liens entre fréquence et sens dans une langue, ou bien Sinclair qui s'est toujours basé sur des observations de langage authentique en contexte dans ses travaux lexicographiques (Honeybone 2005:85, Williams 2006:153).

D'autres linguistes se sont également mis à travailler sur des ensemble de textes authentiques, observant de façon empirique les données accessibles du langage pour

sa propre composante au complexe selon son contexte approprié. » (ma traduction)

16 « Ainsi donc, les catégories linguistiques et situationnelles ne seront pas insurmontables. Beaucoup de nouvelles catégories émergeront d'une observation systématique des faits. » (ma traduction).

17 « Les sciences ne devraient pas imposer des systèmes sur les langues, mais plutôt rechercher des systèmes dans l'activité orale de langage, et, les ayant découverts, formuler les faits dans une langue appropriée. » (ma traduction)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

atteindre la langue. Dans un article, (Léon 2007) Jacqueline Léon donne un aperçu des différents courants théoriques des années 1950 et 1960. Des linguistes d'une filiation Bloomfieldienne, qui avaient pour but d'établir des taxonomies d'éléments linguistiques, ont au cours de ces années travaillé sur des corpus en utilisant des procédures de découverte de type inductif.

1.1.2.3.3- La naissance d'une linguistique de corpus moderne.

Dans la seconde moitié du vingtième siècle, des étudiants de Firth ont ainsi développé davantage le contextualisme : Halliday ou Sinclair n'ont eu de cesse de développer des techniques, des théories qui puissent permettre de rendre compte d'une langue de la façon la plus précise qui soit, sans en gommer les aspérités et en se basant avant tout sur une étude en contexte, qu'il s'agisse de grammaire (Halliday) ou bien de lexicographie (Sinclair).

Chomsky n'a pas formulé comme grief essentiel vis à vis du corpus le fait que ce soit une méthodologie qui dans les années 1960 pouvait prendre un temps considérable avant que les premiers résultats en soient disponibles, et que son traitement humain était générateur non seulement de coûts importants, mais également d'erreurs (Mc Enery et Wilson 1996:10). Firth lui même a mentionné les collocations, mais n'en a pas montré l'étendue dans le langage, peut être parce que n'ayant alors pas de technologies informatiques développées pour les étudier, il ne pouvait pas être conscient de cette étendue. L'avènement des nouvelles technologies, à travers l'utilisation d'ordinateurs capables de traiter de très grandes quantités de mots, a donc eu un impact positif sur l'utilisation des corpus (Williams 2006:155). Le développement des nouvelles technologies a permis de traiter des corpus d'une taille de plus en plus importante, dans des dimensions dont Chomsky ne pouvait probablement pas présager lorsqu'il adressait ses critiques à l'égard des études menées sur corpus.

Des linguistes comme Sinclair ou Halliday, ont montré que sans se départir de leur intuition, ils étaient néanmoins à même de recevoir les nombreuses preuves et informations authentiques accessibles à travers des corpus de langue authentique. Ils ont également veillé à rester ouverts aux évolutions technologiques qui ne peuvent qu'influencer toute attitude scientifique vis à vis des données observées. En gardant à l'esprit que les possibilités actuellement offertes par le traitement informatique de corpus de textes ne pouvaient

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

seulement être imaginées il y a de cela vingt ou trente ans, on conçoit mieux en quoi cette nouvelle approche a apporté une modification importante de la façon dont la linguistique était envisagée.

Halliday a notamment développé une grammaire systémique et fonctionnelle. Il en a emprunté les idées de base à son professeur, J. R. Firth et à l'école de Prague. Son approche cherche à découvrir comment fonctionne une langue, à la fois au niveau interne, c'est-à-dire la structuration de la langue en un système, par exemple, et au niveau externe : comment la langue fonctionne en tant que fait social, humain (Banks 2005:17). La langue est ainsi considérée en tant qu'un système dans lequel de multiples choix peuvent être effectués par le locuteur, le dernier type de ces choix étant d'ordre lexical (Ibid:19). C'est là que se trouve une des grandes nouveautés de la théorie de Halliday, dans laquelle la dimension sémantique est pleinement intégrée à la grammaire, on parle notamment d'une « lexicogrammaire » (Ibid :19). Les travaux de Sinclair ont concerné principalement le domaine de la lexicographie, nous allons y revenir dans la partie suivante.

Après avoir vu comment jusque dans les années 1970 une linguistique de la « langue » avait coexisté avec une linguistique de la « parole », nous allons maintenant étudier plus précisément comment cette « dichotomie » évoquée comme étant indissociable de la linguistique moderne s'est muée en une forme de dualité au sein même du paradigme de la linguistique de corpus (Rastier 2005).

I.1.3- La linguistique de corpus, une discipline aux multiples facettes.

Le paysage de la recherche britannique s'est intéressé aux études sur corpus (le corpus étant une collection de textes authentiques informatisée), et à ce qu'il était possible d'en retirer pour la linguistique non seulement appliquée, mais aussi générale de façon importante à partir des années 1970. D'autres pays ont également commencé à constituer des corpus informatisés (Williams 2006:151). Deux filiations semblent néanmoins apparaître dans le domaine de la linguistique de corpus moderne, une filiation firthienne suivant l'école contextualiste britannique dont les prémices ont été évoquées plus haut, et une filiation suivant davantage les travaux de Quirk et Leech travaillant aussi à partir de corpus informatisés, mais dans une approche moins contextualiste, plus proche du traitement automatique de la langue (Williams 2006:156). Nous allons étudier ici la façon dont ces deux

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

filiations ont développé le paradigme de linguistique de corpus, et dans laquelle cette thèse s'inscrit.

1.1.3.1- La filiation firthienne : John Sinclair d'Edimbourg à Birmingham.

1.1.3.1.1- Les projets OSTI et COBUILD, le corpus en lexicographie appliquée.

Dans les années 1960, John Sinclair, un ancien étudiant de Firth a développé un corpus dans le cadre d'un projet débuté en 1963 à l'université d'Edimbourg puis terminé à l'université de Birmingham, le corpus OSTI (Office for Scientific and Technical Information). Les résultats de l'analyse de ce premier corpus électronique ont été compilés dans un rapport OSTI dans les années 1970 qui a été peu diffusé jusqu'à sa re-publication en 2004. Le but était alors d'explorer la lexis au sein du paradigme contextualiste. Ce projet a notamment permis à Sinclair de mûrir l'idée des collocations, cette fois en les découvrant de façon automatisée et sans l'aide de l'introspection dans les données du corpus (Williams 2005:154)

Toujours à l'université de Birmingham, Sinclair a entrepris dans le début des années 1980 un nouveau projet d'une ampleur sans précédents, le projet COBUILD. L'histoire de ce projet est relatée par les principales personnes l'ayant supervisé dans le livre *Looking Up*. Le COBUILD résulte d'une communauté d'intérêts entre des partenaires à la fois industriels (la société d'édition Collins), universitaires (l'université de Birmingham) et gouvernementaux. Il s'est agi de créer un corpus informatisé afin de pouvoir réaliser le premier dictionnaire pour apprenant totalement fondé sur une analyse de corpus, l'idée étant de voir comment les mots étaient réellement utilisés (Sinclair 1987:vii).

De nombreux choix ont été réalisés dans la création de ce corpus. Par exemple, il a été décidé d'y entrer des textes authentiques et entiers, c'est à dire qu'ils n'étaient ni formés pour répondre à ce que l'on aurait souhaité trouver dans le corpus, et qu'ils n'étaient pas échantillonnés, mais que leur intégrité était préservée (Sinclair 1987:2). Au cours de ce projet, l'informatique avait une place de choix, de nombreux programmes ont été développés pour travailler sur le COBUILD, la dimension informatique du corpus était essentielle pour pouvoir mener le type de recherches empiriques qui étaient par la suite souhaitées. En effet, si l'idée d'utiliser l'informatique dans des recherches lexicographiques n'était pas neuve dans

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

le début des années 1980, celle de les utiliser pour avoir une vision large et compréhensive de l'anglais moderne l'était (Sinclair 1987:41). Jeremy Clear, le superviseur de la partie informatique du projet insiste notamment dans *Looking Up* sur le fait que des concordances obtenues en format KWIC (Key Word In Context) ont permis d'obtenir des données qui n'avaient jamais été accessibles auparavant, le paradigme contextualiste étant clairement présent dans l'utilisation de ce format. Ce projet a permis de constituer un premier corpus d'environ 7 millions de mots que de nombreux linguistes ont pu étudier à l'université de Birmingham.

1.1.3.1.2- Corpus et linguistique générale.

Ces projets (OSTI et COBUILD) ont amené John Sinclair à se questionner quant aux pratiques appliquées lors de la création d'un corpus à des fins linguistiques. La plupart de ces réflexions se trouvent dans *Corpus, Concordance, Collocation* (1991). Dans ce livre fondateur, il préconise tout d'abord d'utiliser des corpus qui soient lisibles et exploitables par des ordinateurs. Le corpus, dans le domaine de la linguistique moderne se doit donc avant toute chose d'être « machine-readable » comme nous l'avons déjà vu auparavant :

« Corpus has to have the material in electronic form » (Sinclair, 1991 : 14)¹⁸

Ce format électronique peut être atteint de différentes façons, il peut s'agir de texte déjà existant sous forme numérique que l'on adapte pour l'intégrer dans le corpus, il peut s'agir de texte imprimé que l'on scanne, ou tout simplement, de texte que l'on va saisir à l'aide d'un clavier pour lui donner une forme électronique. Sinclair préconise également l'utilité de fabriquer un corpus qui contienne à la fois du texte écrit et oral, bien que le texte oral puisse parfois être très proche de l'écrit (on peut parler de discours préparé pour les journaux télévisés, par exemple).

L'un des principaux conseils de John Sinclair concernant la conception des corpus porte sur leur taille, un corpus se doit d'être le plus grand possible, ce qui est évidemment relatif selon d'une part l'époque à laquelle on constitue un corpus, et d'autre part le type de langue que l'on transcrit : un corpus oral sera toujours moins grand qu'un corpus écrit car beaucoup plus long et coûteux à construire.

18 « Le corpus doit être un matériau présent sous une forme électronique. » (ma traduction)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

« *The only guidance I would give is that a corpus should be as large as possible, and should keep on growing* » (Sinclair, 1991 : 18)¹⁹

Sinclair explique ce conseil par un constat sur la répartition des mots au sein des textes fait par Zipf (1936)²⁰. Il note que la moitié des mots dans un texte ne sont présents qu'à une reprise (on parle alors d'hapax legomena), et que les mots les plus fréquents sont généralement des mots grammaticaux, sans valeur lexicale. D'autre part, lorsque l'on s'intéresse à un mot lexical et à ses usages, il y a le plus souvent un usage qui est majoritaire, et d'autres usages qui sont moins fréquents. Une recherche qui souhaiterait étudier de façon empirique les différents usages d'un mot ne peut donc faire l'économie d'un vaste corpus. C'est pourquoi, selon Sinclair, un corpus doit se composer, quelle que soit la façon dont il est structuré, de plusieurs millions de mots.

Dans *Corpus, Concordance, Collocation*, il note également que la plupart des corpus qui se forment alors ont tendance à utiliser un système d'échantillonnage afin d'être davantage représentatifs de la langue, c'est à dire que plutôt que de tenter de représenter une langue en prenant des textes entiers dans différents domaines, les corpus utilisant ce système ont recours à des échantillons de textes, ce qui leur permet, avec le même nombre de mot que s'ils avaient pris des textes entiers, de couvrir une plus grande variété de genres textuels. C'est l'approche du corpus Brown LOB (London-Oslo-Bergen Corpus), par exemple, ou plus récemment, du British National Corpus.

Cependant, Sinclair propose d'utiliser plutôt des documents entiers pour constituer un corpus. Le corpus est alors une entité qui n'a pas une taille finie, mais qui peut toujours continuer à croître. Cette méthode de constitution présente l'avantage non négligeable de pouvoir produire un outil qui pourra s'adapter aux logiciels plus efficaces à venir, ainsi qu'aux futures capacités des ordinateurs. Sinclair nomme ce type de corpus « monitor corpus », à la différence des « sample corpus » qui se basent sur des échantillons. Il préconise la construction de ce type de corpus pour toute langue ayant un statut international.

19 « Le seul conseil que je donnerai est qu'un corpus doit être aussi grand que possible, et doit pouvoir continuer à croître. » (ma traduction)

20 La loi de Zipf est une observation empirique de la fréquence des mots dans un texte. Elle a été élaborée par George Kingsley Zipf en 1936 à la suite de constats réalisés sur la répartition des mots au sein de l'*Ulysse* de Joyce. $f(n) = \alpha n^{-\beta}$ en exponentiel, où n est le rang d'un mot quand tous les mots sont ordonnés par ordre de fréquence décroissant, et $f(n)$ est la fréquence d'un mot de rang n . Alpha et beta sont constants. Il apparaît que cette loi est vraie qu'en tant que première approximation brute. (rapport Osti, Sinclair, 1970/2004:31-2)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

Le fait qu'il n'y ait pas de taille prédéfinie pour le corpus s'explique selon Sinclair par les développements de la technologie informatique : les textes sont de plus en plus présents sous forme numérique, il est donc de moins en moins compliqué d'obtenir du matériau pour constituer son corpus. En outre, les capacités de stockage et de traitement des ordinateurs suivent une croissance exponentielle qui assure le traitement de futurs corpus gigantesques. On constate enfin, que ce qui a été avancé en 1991 se vérifie, ainsi les possibilités offertes par les ordinateurs personnels croissent (Williams 2006:155), croissance qui se poursuivra sans doute dans les années à venir.

Selon Sinclair (1991), le corpus doit également rester vierge au maximum des annotations du chercheur, il doit être le plus standard dans la mesure du possible (chaque chercheur y imposant malgré tout nécessairement sa façon de voir, ne serait-ce que dans le choix de balises xml). A cela, trois raisons. Chaque recherche a une vision particulière du langage, donc chaque recherche va avoir tendance à annoter différemment chaque corpus, et de cette façon, le long travail de collecte du corpus ne pourrait bénéficier qu'à une seule recherche, ou à un seul type de recherches, toutes les autres ne trouvant pas leur intérêt dans un matériau annoté sans rapport avec les buts poursuivis. D'autre part, la linguistique utilise souvent des termes différents pour nommer les parties du discours, une annotation dont la visée est linguistique est donc risquée en ce qu'elle ne peut que difficilement être standardisée. Le corpus ne doit donc être annoté qu'au minimum, et si d'autres informations sont apportées au corpus en vue d'une étude précise, il convient de garder une forme première du matériau annoté au minimum. Enfin, les termes généralement utilisés pour annoter les corpus proviennent pour la plupart de la tradition linguistique, annoter le corpus selon une tradition vieille de plusieurs siècles comporte ainsi en substance le risque de ne pas pouvoir être conscient d'une réalité qui pourrait ne pas être dans une adéquation totale et complète avec cette tradition.

Les caractéristiques principales d'un corpus, telles que Sinclair les a appliquées lors de la création du corpus COBUILD qui a par la suite évolué en Bank of English²¹ sont donc une taille importante (le corpus comprend actuellement quelques 524 millions de mots), une capacité à toujours évoluer, une absence de taille finie, et un matériau le moins annoté possible.

21 <http://www.collins.co.uk/books.aspx?group=153> (dernière consultation le 23/09/2009)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

I.1.3.2- La filiation Quirk-Leech²².

Cette deuxième tradition compte des linguistes qui ont suivi le courant de pensée amorcé par les travaux de Quirk et sa grammaire en 1972 et ne se situent pas dans la tradition firthienne évoquée plus haut. Elle se caractérise notamment par la fabrication d'un autre corpus de grande dimension, le British National Corpus (BNC)²³. Dans cette partie, nous allons plus particulièrement nous intéresser à l'école de Lancaster où la tradition linguistique dont nous allons parler s'est développée, puis nous verrons un exemple de corpus suivant les préconisations de cette école à travers la description du British National Corpus.

I.1.3.2.1- L'école de Lancaster, l'apport de Geoffrey Leech.

L'université de Lancaster est un des pôles où la linguistique de corpus s'est développée au cours des dernières décennies. Geoffrey Leech qui y est professeur a notamment contribué à ce mouvement de recherche en participant à la création de corpus²⁴ tels que le Brown-LOB (Lancaster Oslo Bergen) corpus en 1978. A travers des études sur corpus, il a notamment élaboré avec Biber une grammaire de l'anglais qui prend essentiellement en compte des analyses de fréquences et des relevés statistiques de formes linguistiques. Cette grammaire a été publiée sous le titre de *Longman Grammar of Spoken and Written English*, by Biber, Johansson, Leech, Conrad and Finegan en 1999.²⁵ Au cours de ces participations à la conception de corpus, il s'est plus particulièrement intéressé à la problématique de l'annotation de corpus qu'il considère comme une valeur ajoutée pour ce corpus. En effet, l'annotation de corpus peut permettre lorsque l'on explore les données de trouver plus facilement le mot que l'on recherche exactement, selon que l'on s'intéresse à un verbe ou à un adjectif qui prennent la même forme graphique, l'annotation permettra d'obtenir les lignes de concordances pertinentes (Leech 2004). Il insiste dans ses recherches sur la nécessité de fournir une annotation de qualité pour un corpus, cette qualité pouvant se définir selon trois critères : la précision (le pourcentage de mots correctement annotés dans le corpus), la cohérence des étiquettes annotées entre elles, et la réalité des catégories attribuées dans la langue (Leech 2004). Le logiciel CLAWS mis au point dans son équipe est un

22 Stubbs (1993)

23 [Http://www.natcorp.ox.ac.uk](http://www.natcorp.ox.ac.uk) (dernière consultation le 23/09/2009)

24 <http://www.ling.lancs.ac.uk/profiles/296> (dernière consultation le 04/09/2009)

25 <http://www.ling.lancs.ac.uk/profiles/296> (dernière consultation le 04/09/2009)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

étiqueteur semi probabilistique qui attribue aux éléments lemmatisés du corpus une étiquette donnant la partie du discours à laquelle ils appartiennent.

L'annotation de corpus ne fait cependant pas le consensus au sein de la discipline linguistique de corpus, si certains chercheurs préfèrent annoter les corpus qu'ils utilisent afin d'en faciliter leur étude, d'autres, tels Sinclair, pointent les problèmes de ces annotations. En effet, chaque chercheur peut souhaiter annoter différemment le corpus sur lequel se fonde son étude en fonction de ses buts de recherche, une annotation complète et finie peut alors se révéler inapte à l'utilisation qui sera faite du corpus. Mais Sinclair note surtout que plus un corpus est annoté, et plus on risque de le biaiser avec sa propre vision de la langue : alors que l'on est conscient que la plupart des catégories bien établies au niveau grammatical, par exemple, depuis Aristote, ne conviennent pas à ce qui se passe réellement dans la langue en usage, les utiliser pour annoter le corpus risque de biaiser l'étude qui sera menée, et de fausser les analyses effectuées sur ce corpus (Sinclair 1991:21-2). Le BNC, situé dans la filiation Quirk Leech fait néanmoins le choix de l'annotation, tout en conservant une version non annotée pour l'utilisateur qui souhaiterait éviter les écueils évoqués par Sinclair.

1.1.3.2.2- Le British National Corpus.

Ce corpus a été fortement inspiré dans sa conception du COBUILD, il a notamment eu le même superviseur, Jeremy Clear dans le domaine informatique (Sinclair et al. 1987, Burnard 2007:3). Le projet, visant à élaborer le BNC est né d'un regroupement d'intérêts communs entre des éditeurs de dictionnaires (Longman et Chambers), des universitaires et chercheurs (la lexicographe Sue Atkins, l'Université d'Oxford, l'Université de Lancaster et la British Library) et le gouvernement britannique. Ces diverses entités éprouaient différents intérêts dans la mise au point d'un tel projet au cours des années 1990. Les éditeurs Longman et Chambers, encouragés par le succès de l'entreprise Collins COBUILD menée à Birmingham par l'équipe de John Sinclair, ont pu souhaiter financer l'expérience afin d'avoir au mieux des retombées en termes de nouveaux dictionnaires, au pire, en termes de meilleure image. Le gouvernement britannique pouvait quant à lui stimuler l'ingénierie en langue anglaise. Les chercheurs se trouvaient également intéressés par le projet du fait de la possibilité entrevue d'établir un nouveau modèle de développement de grands corpus, à la façon du corpus Bank of English de Birmingham (Burnard 2007).

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

Le BNC est donc un corpus d'anglais britannique moderne d'environ 100 millions de mots (sa dernière version, datant de 2007 comprend 98 363 784 éléments lemmatisés), et a été réalisé après The Bank of English, entre 1990 et 1994. Bien que moins important en taille que The Bank of English, son principal intérêt repose sur le fait qu'il peut être distribué de façon légale dans le monde entier, quand le corpus de Birmingham a un accès limité. C'est le plus grand corpus actuellement distribué de cette façon. Il comprend environ 10 % de textes oraux, cette faible proportion s'explique par le coût important de toute transcription de langue orale en comparaison avec la collecte de textes écrits. La collecte et la saisie des textes le composant ont été effectuées à l'université d'Oxford. C'est un corpus qui a une taille finie, il donne donc un aperçu de la langue anglaise dans les années 1990. Comme la plupart des corpus visant à représenter la langue dans sa variété et étant de taille finie, il se compose d'échantillons de divers textes dans divers domaines.

L'annotation du BNC s'est faite selon les standards de la TEI (Text Encoding Initiative)²⁶. La TEI est un ensemble de règles visant à standardiser les annotations généralement ajoutées aux textes formant un corpus pour décrire les informations relatives à ces textes. C'est également un ensemble de normes qui a été mis au point en parallèle du BNC. La TEI a été développée au même endroit, au même moment, et par les mêmes personnes que le BNC. Le BNC est donc parfaitement conforme à ces standards, il contient ainsi les informations relatives aux textes qui le composent en lui-même, et non pas dans une documentation extérieure au corpus, comme cela pouvait se faire dans les années 1990.

Il a également été étiqueté linguistiquement par le système CLAWS²⁷ (the Constituent Likelihood Automatic Word-Tagging System) au sein de l'université de Lancaster sous la coordination de Geoffrey Leech.

Actuellement, ce corpus est distribué avec un logiciel d'extraction, XAIRA (XML Aware Indexing and Retrieval Architecture)²⁸, à l'origine, le logiciel s'appelait SARA (Sgml Retrieval application) il est toujours disponible en open source sur le site <http://sara.natcorp.ox.ac.uk> (Burnard 2007:14). Il a été créé à l'aide de fonds du British National Corpus consortium et de la fondation Andrew Mellon. Ce programme a été développé par le Oxford University Computing Services (OUCS) précisément pour

26 [Http://www.tei-c.org/index.xml](http://www.tei-c.org/index.xml) (dernière consultation le 23/09/2009)

27 [Http://ucrel.lancs.ac.uk/claws](http://ucrel.lancs.ac.uk/claws) (dernière consultation le 23/09/2009)

28 [Http://www.xaira.org](http://www.xaira.org) (dernière consultation le 23/09/2009)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

fonctionner avec le BNC et a été distribué sous le nom de SARA à partir de 1994²⁹. Ce logiciel a été totalement remanié entre 2004 et 2005 pour prendre le nom de Xaira. Ce remaniement a notamment permis de le distribuer indépendamment du BNC, de le rendre accessible gratuitement pour l'usage de toute personne souhaitant interroger un corpus de grande taille. Ainsi, il peut actuellement fonctionner pour tout type de document dans un format .xml, néanmoins, il fonctionne mieux s'il s'agit d'un document conforme aux normes de la TEI. Ce logiciel permet entre autres possibilités d'extraire des concordanciers de corpus xml conformes aux normes de la TEI, de calculer des collocations, de créer des partitions, etc. dans des corpus de grande taille.

Le BNC n'est pas un monitor corpus, ou corpus suivi tel que celui élaboré par John Sinclair. Il a une taille finie, et par exemple, si l'on recherche l'expression « World Wide Web » dans le corpus, elle n'apparaît qu'à deux reprises et devrait donc être considérée comme quelque chose de marginal au sein de l'anglais britannique moderne. Néanmoins, deux autres versions ont été éditées depuis la première version de 1994. En 1999, The British National World Edition a été diffusé. Les principales modifications ont concerné pour cette version une amputation d'une cinquantaine de textes afin de pouvoir le diffuser légalement dans le monde entier, une amélioration a également été apportée à l'étiquetage par CLAWS du corpus. La troisième version est le British National Corpus XML édition de 2007. Outre l'adaptation du corpus au format XML alors largement répandu, cette édition a également modifié l'annotation afin de la rendre plus conforme aux normes de la TEI, le corpus a été amélioré au niveau de son utilisation avec d'autres outils XML qui se sont développés, et son balisage a été simplifié.

On constate ainsi à travers l'exemple de ces deux corpus de référence que si une grande taille semble toujours être de mise lorsque l'on parle de corpus en linguistique moderne, ainsi qu'une grande variété de textes, des divergences peuvent apparaître. La tradition firthienne opte pour un corpus suivi quand une tradition plus orthodoxe a préféré un corpus fini, la première a préféré un corpus annoté au minimum quand la seconde a étiqueté le BNC.

Le paradigme de la linguistique de corpus, actuellement en plein essor (Williams 2006) réunit en quelques sortes deux linguistiques dont le développement a pris place tout au

29 <http://drh2004.ncl.ac.uk/abstract.php?abstract=218> (dernière consultation le 23/09/09)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

long du vingtième siècle. Selon la « dualité » saussurienne (Rastier 2005) de la langue et de la parole, on pourrait dire qu'une linguistique de la langue s'est développée parallèlement à une linguistique de la parole.

La linguistique générale s'est beaucoup préoccupée de la langue, interprétée comme un modèle devant être abstrait, notamment à travers le courant structural, tandis qu'une linguistique de la parole n'a cessé de se développer, tout d'abord au sein de la linguistique appliquée britannique, puis de façon beaucoup plus généralisée avec l'utilisation d'outils informatiques à partir des années 1970.

La théorie actuelle de la linguistique de corpus prend sa source dans différents projets de linguistique appliquée, qu'il s'agisse de lexicographie (projet COBUILD, BNC) ou d'enseignement de langue vivante (Williams 2006). Elle fonctionne au sein du paradigme contextualiste, si l'université de Lancaster, dans une filiation Quirk-Leech semble davantage s'intéresser au Traitement Automatique de la Langue, et sur la création d'outils informatisés notamment, la plupart des groupes de chercheurs ou de laboratoires en linguistique de corpus quel que soit le lieu où ils prennent place se situent dans une démarche contextualiste. (Williams 2006). La tour d'ivoire grammaticale dans laquelle on a parfois reproché au linguiste de s'enfermer n'est plus, et la dimension sémantique du langage est pleinement étudiée (Rastier 2005).

Un des aphorismes les plus connus de Saussure pose que c'est le point de vue qui crée l'objet. Sinclair (1991:100) note que la façon de voir le langage peut changer considérablement lorsque l'on peut en voir une grande quantité en une seule fois, Tognini-Bonelli poursuit cette idée en disant que parfois, c'est le point de vue qui crée la discipline. Dans le domaine de la linguistique de corpus, cela signifie que l'impact des nouvelles technologies a totalement modifié le point de vue que l'on avait de ce sujet qu'est le langage. Grâce aux nouvelles technologies, aux corpus informatisés, aux logiciels d'extractions et aux concordanciers, il est devenu possible d'observer la langue à travers la « masse parlante » saussurienne, ce n'est plus un seul fait de parole qui est donné à voir, mais une véritable masse de parole à travers laquelle se réalise la langue.

« langue has become now observable for us on the vertical axis of the concordance. It is the awareness of the social habit underlying the

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

presence of repeated patterning. The concordance can thus reunite for us the instance of parole on the horizontal axis and the awareness of langue of the vertical one » (Tognini-Bonelli 2001:169)³⁰

Langue et parole, ainsi que les différents points de vue ou différentes linguistiques que ces deux pendants d'une même « dualité » (Rastier 2005) se rejoignent ainsi au sein de la linguistique de corpus.

C'est dans ce paradigme que s'inscrit cette thèse, travaillant sur un corpus qui a été formé notamment selon les principes généraux élaborés par Sinclair à travers les nombreux corpus qu'il a construit, et adoptant vis à vis des données du corpus une approche largement lexicographique. La seconde partie de ce chapitre va maintenant développer davantage les différents types de corpus qui peuvent être créés selon les buts de recherche que l'on se donne, et la façon dont le corpus IntUne, sur lequel les analyses de cette thèse sont faites, a été élaboré.

I.2- Le corpus, une entité à géométrie variable.

Cette thèse se base sur un ensemble d'études et d'analyses qui ont été menées sur corpus. Il nous semble dès lors essentiel de définir plus précisément comment cette entité est le plus souvent utilisée en linguistique de corpus, ainsi que ses applications dans des domaines aussi nombreux que variés (la lexicographie, la traduction automatique, l'apprentissage des langues, etc.). Une explication des choix qui ont été faits plus précisément concernant la constitution du corpus dans le projet IntUne sur lequel se base notre étude doit également être posée avant de procéder à toute analyse, tant il est essentiel pour tout analyste d'être conscient des atouts, mais aussi des limites de ce sur quoi se fonde son analyse.

Nous verrons donc dans un premier temps que les corpus électroniques, loin de présenter une réalité homogène peuvent prendre de multiples formes, que ces corpus peuvent

30 « La langue est devenue maintenant observable pour les chercheurs sur l'axe vertical des concordances. C'est la conscience de l'usage social qui sous-tend la présence de structures répétitives. Les concordances peuvent en conséquence réunir pour nous des instances de la parole sur l'axe horizontal, et la conscience de la langue sur l'axe vertical » (ma traduction).

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

être étudiés selon différentes approches. Nous nous poserons dans un second temps la question des limites que peuvent présenter les corpus. Enfin, nous verrons quelles applications les corpus peuvent avoir.

I.2.1- Les corpus électroniques : multiplicité de formes et d'approches.

Les corpus électroniques, ou machine-readable (Sinclair 1991) connaissent un développement certain actuellement (Kennedy 1998), ce développement a notamment contribué à la conception de corpus de plus en plus nombreux, et prenant des formes variées. Nous allons donc distinguer quelles sont les principales formes de corpus que la tradition en linguistique de corpus utilise, afin de situer plus précisément le corpus IntUne parmi ces différents choix. Il sera ainsi possible de mieux comprendre quelles possibilités ce corpus offre.

I.2.1.1- Différents types de corpus.

Selon le type de recherche que l'analyste mène, le domaine dans lequel il se situe, différents corpus peuvent être conçus : un chercheur en lexicographie s'intéressant à l'élaboration de dictionnaire spécialisé en biologie dans une langue précise (Williams 1998) ne concevra pas ou du moins, n'utilisera exactement le même type de corpus qu'un traducteur professionnel, un enseignant en langue étrangère, ou un linguiste travaillant à une meilleure description d'une langue en général. Il existe ainsi différents types de corpus. Sans entrer dans une réflexion épistémologique détaillée à ce sujet (Rastier 2005, Jacques 2005, Alonso 2009), nous allons voir quels sont les principaux types de corpus qui existent. Le plus souvent, on distingue entre corpus monolingue et corpus multilingues (Tognini Bonelli 2001, Déjean et Gaussier 2002, Pearson 2003). Étant donné que le corpus IntUne est multilingue (anglais, français, italien et polonais), nous allons décrire plus particulièrement les deux types de corpus généralement reconnus dans cette catégorie : les corpus parallèles et comparables.

I.2.1.1.1- Les corpus parallèles.

Les corpus parallèles sont des ensembles de textes traduits dans plusieurs langues

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

que l'on aligne à un niveau syntagmatique. Il s'agit de corpus qui ont essentiellement une utilité au niveau de la traduction en linguistique de corpus. Les premiers ensembles de ce genre étaient réellement parallèles, Tognini-Bonelli donne l'exemple d'un corpus d'anglais et de français parlementaire enregistré de façon très rigoureuse au Canada dans lequel les deux corpus étaient alignés phrase à phrase (2001:6). Tognini-Bonelli précise ensuite (2001:7) qu'avant de mettre au point ce type de corpus, on a créé des corpus qui étaient de traduction libre. Tout comme pour le corpus parallèle, on traduisait un même texte dans plusieurs langues, mais cette fois sans qu'un souci particulier ait été porté sur sa correspondance rigoureuse d'une langue à l'autre, notamment du point de vue syntagmatique. Ces ensembles étaient en conséquence beaucoup moins faciles à traiter et à étudier de façon informatisée, puisqu'ils n'avaient pas une rigueur suffisante dans leur conception.

Pearson (2003:167) cite aussi les corpus parallèles comme étant très utilisés dans le domaine de la traduction, mais elle affine ce point de vue en notant que s'ils sont très présents au niveau de la recherche fondamentale en traduction, on ne les utilise que relativement peu dans le domaine de l'apprentissage ou de l'entraînement à la traduction.

Néanmoins, malgré leur utilité, ces corpus semblent poser quelques difficultés. Déjean et Gaussier montrent ainsi que si ces corpus peuvent présenter une utilité manifeste dans différents domaines (apprentissage, traduction), ils posent un réel problème d'accès.

« Il est plus facile d'accéder à un corpus comparable dans un domaine donné qu'à un corpus parallèle de bonne qualité. » (Déjean et Gaussier 2002:315)

Un autre type de corpus apparaît dans cette citation, le corpus comparable, auquel il semble plus aisé d'accéder, nous allons maintenant voir ce que sont les corpus comparables.

1.2.1.1.2- Les corpus comparables.

Les corpus comparables ont été imaginés, selon Tognini-Bonelli (2001:7), notamment du fait de l'absence de représentativité de corpus strictement parallèles et du peu de ressources traduites avec suffisamment de rigueur qui étaient disponibles, rejoignant alors l'argument de Déjean et Gaussier (2002). Outre leur plus grande accessibilité, les corpus comparables regroupent par ailleurs des textes dans leur langue originale, et présentent donc des faits de langue authentiques, non traduits. Les corpus comparables sont constitués de

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

différents éléments qui sont choisis pour représenter des catégories semblables dans les différentes langues qui constituent le corpus Tognini-Bonelli (2001:7) donne comme exemple de catégories communes le fait de choisir du langage oral ou écrit dans telles proportions pour chaque langue, celui de choisir tel registre. Déjean et Gaussier (2002:314) confirment ce point de vue en expliquant que le corpus comparable peut se baser sur ce type de critères qualitatifs (genre, auteur, période, média) qui seraient les mêmes dans les corpus des différentes langues collectées, ils ajoutent que les corpus comparables peuvent aussi utiliser des mesures quantitatives pour en évaluer la proximité. Pearson définit brièvement le corpus comparable comme suit :

« Comparable corpora, i.e. collections of texts in two or more languages containing texts of the same type and dealing with the same domain, » (Pearson 2003:168)³¹

Un exemple de corpus comparable qui est établi au niveau international est l'ICE (International Corpus of English), sa conception a débuté en 1990, notamment à l'initiative de Sidney Greenbaum, et il est actuellement coordonné par Gerald Nelson. Ce corpus a pour but de mener des études comparatives sur les différentes façon dont on peut parler anglais dans les pays anglophones à travers le monde. Afin de permettre une étude comparative, les différents corpus élaborés dans les pays étudiés doivent adopter des critères communs de conception, ainsi, tous se composent d'environ un million de mots d'anglais écrit et oral (avec une majorité de langue orale), selon des proportions et des registres ou types de textes clairement énoncés³². Ce type de corpus fonctionne donc sur des échantillons d'une taille exactement identique d'une langue à une autre, cela n'est pas le cas du corpus IntUne sur lequel cette thèse se base, en effet, d'un pays à un autre, les journaux utilisés pouvaient être plus ou moins prolixes, de plus, ce sont des textes entiers qui figurent dans ce corpus, néanmoins, nous postulons qu'à partir du moment où des critères rigoureux ont été mis en place pour construire le corpus IntUne et où ces critères ont été suivis dans les quatre pays du groupe média, le corpus IntUne est un corpus qui permet de mener des études comparatives.

Le corpus IntUne a été conçu pour pouvoir être étudié de façon comparative à la

31 « Les corpus comparables, c'est à-dire des ensembles de texte dans deux langues ou plus contenant des textes du même type et dont le sujet concerne le même domaine » (ma traduction).

32 <http://ice-corpora.net/ice/design.htm> (consulté le 10/09/2009)

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

fois au niveau diachronique et au niveau national. Il comprend des textes qui proviennent des médias écrits et télévisés, collectés à des périodes et sur des types de journaux semblables pour les quatre pays du groupe média, il est par conséquent un corpus comparable. Mais nous reviendrons plus tard dans le détail de sa conception ainsi que les difficultés que celle-ci a pu poser. Avant, il semble important de définir les différentes approches qu'un chercheur peut avoir vis à vis du corpus qu'il étudie.

1.2.1.2 – Différents types d'approches.

Le corpus est une entité qui permet au chercheur d'avoir une multitude d'approches dans son étude. A l'approche qualitative, peut s'opposer ou se combiner une approche quantitative, à une démarche déductive ou corpus-based peut s'opposer une démarche inductive ou exploratoire, corpus-driven. Il convient alors d'explicitier ces différentes possibilités.

1.2.1.2.1- L'approche qualitative.

L'approche qualitative concerne des quantités limitées de données. Bien que se situant au sein d'un corpus, ces données peuvent être regroupées en des sous corpus en fonction du pôle d'intérêt de l'étude menée. De nombreux paramètres conservés dans l'annotation du corpus peuvent alors être pris en compte, qu'il s'agisse du locuteur, de son âge, de son sexe, de la nature du discours (oral ou écrit), de la situation géographique du locuteur, etc. Cette approche se situe essentiellement dans le cadre de l'analyse critique de discours ainsi que des Corpus Assisted Discourse Analysis (Partington 2006). Elle permet notamment d'étudier les corpus oraux qui sont généralement de taille insuffisante pour mener une approche quantitative de façon significative, et ce, du fait du coût impliqué par la réalisation de tels ensembles (Sinclair 1991:16).

L'analyse critique du discours tend de plus en plus à s'associer à la linguistique de corpus (Mauntner 2009:32). Ces deux disciplines semblent pouvoir collaborer de façon fructueuse. Stubbs a ainsi fourni une analyse se basant tout d'abord sur des textes courts de quelques centaines de mots (Stubbs 1996:81), puis sur des textes plus longs, (allant d'un livre entier à des dizaines de milliers de mots), allant de corpus courts à des corpus plus importants. Au cours de cette étude, il a montré que l'analyse critique du monde social peut

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

s'enrichir d'analyses lexicales ou grammaticales menées sur un vaste corpus, et assistées par un ordinateur. Il prend ainsi l'exemple de textes de Baden-Powell pour montrer les différentes stratégies qui peuvent être mises en place dans un discours véhiculant une idéologie sexiste. Il note que ces stratégies peuvent être mieux analysées par une confrontation à des données issues de vaste corpus avec l'exemple des mots « carrière », « travail » et « foyer ». Il étudie dans un large corpus les connotations dont ceux-ci sont généralement affectés, et en déduit plus facilement dans les discours analysés si ces mots sont porteurs ou non de stratégies spécifiques (Stubbs 1996:84). Ainsi, il procède à une analyse de discours en s'appuyant non seulement sur certains textes sélectionnés pour son étude, mais également sur de larges corpus dont le linguiste peut rechercher les mots clefs principaux à travers notamment les listes de fréquences qui peuvent être fournies par différents concordanciers.

1.2.1.2.2 – L'approche quantitative.

L'approche quantitative pose comme premier postulat l'existence d'un corpus de taille importante (plusieurs millions de mots selon Sinclair en 1991). En effet, pour avoir une représentation importante d'un mot lexical, il faut obligatoirement avoir un corpus d'une grande taille. Cela tient essentiellement au fait que les mots se distribuent régulièrement dans les textes, et 50% des occurrences d'un texte ne sont présentes qu'à une reprise. La méthode scientifique élaborée par John Sinclair en lexicographie suit ce type d'approche, c'est en comparant les fréquences des mots, de leurs usages, que l'on peut élaborer des dictionnaires représentant avec davantage de précision la réalité de la langue. Il peut être plus utile pour un apprenant de connaître le sens le plus généralement employé d'un mot ou d'une expression que d'en avoir des sens, présentés dans un dictionnaire sans avoir d'information sur leur usage. L'approche quantitative permet d'avoir des informations précises sur ce qui est typique ou ce qui ne l'est pas dans la langue (Leech 2002:6).

Cette approche permet également d'apporter de nombreuses réponses aux questions soulevées en ingénierie de la langue, les étiqueteurs (dont CLAWS, dont il a été question précédemment) et une multitude d'autres outils sont des applications d'études statistiques sur la langue. Leech (2002:6-7) insiste sur le fait que le traitement automatique de la langue ne saurait se passer de ces méthodes statistiques.

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

I.2.1.2.3- La démarche « corpus-based »³³

L'approche de type « corpus-based », ou déductive, considère le corpus comme un inventaire de données langagières. Il s'agit donc pour le linguiste d'en extraire le matériau approprié en fonction de sa recherche, ce qui va pouvoir venir conforter une intuition, vérifier certaines attentes, quantifier un phénomène linguistique et surtout, lui donner une assise supplémentaire en l'illustrant par des exemples authentiques issus d'un corpus, et donc en le prouvant. En fait, le corpus y est utilisé pour confirmer des ensembles de savoirs linguistiques déjà prêts. Comme le note Elena Tognini-Bonelli dans *Corpus Linguistics At Work* (2001), il s'agit d'une approche qui peut affiner la connaissance que l'on a déjà de la langue, mais qui ne peut pas révolutionner cette connaissance :

« In this case, however, corpus evidence is brought in as an extra bonus rather than as a determining factor with respect to the analysis, which is still carried out according to preexisting categories; although it is used to refine such categories, it is never really in a position to challenge them as there is no claim made that they arise directly from the data. » (Tognini-Bonelli 2001 : 66)³⁴

Cette approche peut néanmoins être utile pour étudier des phénomènes linguistiques atypiques qui ont peu de chances d'apparaître dans les structures les plus présentes en terme de fréquence dans une langue. Cependant, elle pose le problème de risquer de passer à côté de certaines choses, en restant cloisonnée dans l'étude à la première hypothèse que l'on soumet au corpus :

« There might be a large number of potentially meaningful patterns that escape the attention of the traditional linguist; these will not be recorded in traditional reference works and may not even be recognised until they are forced upon the corpus analyst by sheer visual presence of the emerging patterns in a concordance page. » (Tognini-Bonelli 2001 : 86)³⁵

33 La terminologie « corpus-based » « corpus-driven » vient du livre de Tognini Bonelli qui a mis en évidence ces deux types de démarches différentes.

34 « Dans ce cas, cependant, les preuves trouvées dans le corpus sont considérés comme des bonus plutôt que comme des facteurs déterminants de l'analyse, qui elle, est toujours menée selon des catégories pré-existantes ; bien que le corpus soit utilisé pour affiner de telles catégories, il n'est jamais en position de les défier, puisque qu'à aucun moment, on ne postule que ces catégories pourraient émerger directement des données. » (ma traduction)

35 « Il peut y avoir un grand nombre de structures potentiellement signifiantes qui échappent à l'attention du

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

I.2.1.2.4- L'approche « corpus-driven ».

Une autre approche existe dans le domaine de la linguistique de corpus, il s'agit de l'approche « corpus driven » ou inductive, exploratoire. Dans cette méthode, le corpus est considéré comme une base empirique à partir de laquelle les linguistes peuvent extraire des données et des phénomènes sans avoir une quelconque attente ou hypothèse préalable à leur recherche. Leur conclusion ne peut venir que de l'observation des données. Ce type de recherche se base dans un premier temps sur le corpus seul et non sur une introspection menée par le linguiste. L'étude porte donc sur un mot dont on en étudie les collocats et les concordances, afin de voir au sein de quel contexte, de quelles structures et avec quel usage ce mot apparaît généralement dans la langue étudiée. Cette méthode permet donc de mieux appréhender ce qui est véritablement typique (ce qui est répété dans des termes firthiens) dans une langue en ce qu'elle se base d'abord sur les résultats statistiques fournis par l'ordinateur et le corpus. Souvent, dans ce genre d'approche, la nature de ce qui est trouvé peut être sensiblement différente de ce à quoi s'attendait le linguiste (Sinclair 1991). Son intuition entre tout de même en compte puisque c'est à lui que revient la tâche d'interpréter les résultats statistiques.

« Some very common usages are often not featured in descriptions because they are so humdrum and routine; this method brings them to the fore. Especially in lexicography, there is a marked contrast between the data collected by computer and that collected by human readers exercising their judgement on what should or should not be selected for inclusion in a dictionary. » (Sinclair 1991:4)³⁶

Cette interprétation peut mener à deux types de résultats : les résultats directs, qui sont donnés par l'analyse automatique des collocations d'un mot et qui montrent certains collocats comme étant statistiquement signifiants, et les résultats indirects, qui peuvent être obtenus en menant une analyse plus approfondie des collocations et des concordances, afin de faire ressortir des phénomènes moins saillants statistiquement, mais néanmoins révélés

linguiste traditionnel. Elles ne seront pas relevées dans les travaux habituels de référence et elles ne pourront seulement pas être reconnues tant qu'elles ne forceront pas la main à l'analyste de corpus par leur présence brute et visible de structure émergente au sein d'une page de concordances. » (ma traduction)

36 « Souvent, certains usages très communs ne sont pas marqués dans les descriptions car ils sont routiniers et banals, cette méthode les met au premier plan. Plus particulièrement dans le domaine de la lexicographie, on note un fort contraste entre des données collectées par un ordinateur et celles collectées par un lecteur humain qui exerce son jugement sur ce qui devrait être sélectionné ou non pour être inclus dans un dictionnaire. » (ma traduction).

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

par le corpus. Cette approche semble alors plus fiable scientifiquement qu'une méthode où l'intuition serait à la base de la démarche de recherche, quand bien même l'intuition demeure cruciale, notamment lors de la phase d'interprétation des résultats (Sinclair 1991:4).

C'est en conséquence cette approche inductive et exploratoire qui a été adoptée lors des études menées au cours de cette thèse, influençant certaines des caractéristiques de la structuration du corpus, notamment. Le corpus sur lequel elle se base a été créé dans le cadre d'un projet de l'Union Européenne, IntUne, *Integrated and United in an ever closer Europe*. Ce projet a également influencé le mode de création, la composition, la structure et le type d'études menées sur le corpus, mais avant de détailler ces points, nous allons montrer que si les corpus peuvent prendre différentes formes et s'ils peuvent être étudiés selon différentes approches, ils peuvent également connaître quelques limites.

I.2.2-Les limites des corpus.

I.2.2.1- Un problème de définition.

La popularité croissante des corpus (Sinclair 1991:1) ne doit pas faire oublier les problèmes auxquels ils sont toujours confrontés. D'une part, l'engouement pour leur usage peut risquer d'en diluer le concept, et il est alors essentiel de veiller à ce que sa définition soit précise et rigoureuse. D'autre part, savoir si une collection de textes peut être considérée ou non comme un corpus demeure un problème épineux. Tognini-Bonelli (2001:53) en donne pour exemple une suite de définitions fournies par différents linguistes de corpus dans les années 1980 et 1990. Tour à tour, le corpus sera considéré comme se voulant représentatif de la langue (Francis, 1982), comme étant naturel (Sinclair 1991:6), structuré de façon électronique et selon des critères explicites pour un but spécifique (Atkins, Clear et Osler 1992). Cela peut également être un ensemble d'échantillons de toute forme de langage (Aarts 1991) ou encore un ensemble de « morceaux » de langue qui sont sélectionnés et structurés pour pouvoir être utilisés comme des échantillons de la langue (EAGLES 1993³⁷).

On voit alors que la définition de cet outil demeure souvent vague et diffère parfois

37 EAGLES ou Expert Advisory Group on Language Engineering Standards est projet financé à l'initiative de la Commission Européenne visant à établir des normes pour la transcription électronique de ressources langagières à grande échelle (telles que les corpus) et qui a élaboré un ensemble de normes, le CES, Corpus Encoding Standard. cf. <http://www.ilc.cnr.it/EAGLES/home.html>

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

grandement d'un chercheur à l'autre. Le corpus ne saurait être un simple grand « sac de mots » (Rastier 2005a) compilé sous forme informatique, et les questions soulevées par sa définition posent de nombreux problèmes : doit-il s'agir d'échantillons ou de textes apparaissant en entier ? Doit-il s'agir de texte de la langue courante, ou les proverbes peuvent-ils en faire partie ? Quelles sont les caractéristiques réelles d'un échantillon et d'un « morceau » ?

Il est essentiel d'avoir ces questions à l'esprit dès lors que l'on se lance dans la création ou l'utilisation d'un corpus, quel qu'il soit.

1.2.2- Le problème de la représentativité.

Biber pose que le travail sur corpus en linguistique peut être considéré comme fiable à partir du moment où on travaille sur un corpus « représentatif », c'est-à-dire qui permette de tirer des observations que l'on pourrait généraliser à un comportement langagier plus important que celui précisément sélectionné dans le corpus. Il existe en fait un consensus dans le monde de la recherche linguistique pour dire qu'un corpus est représentatif d'une population bien précise et qu'en conséquence, les observations dérivant de son analyse seront applicables à un ensemble plus vaste. Par exemple, si l'on souhaite analyser la façon dont le langage est utilisé en général, le corpus devra enregistrer l'usage de locuteurs ordinaires, si l'on souhaite étudier tel type de média, on choisira les textes du corpus en fonction (Sinclair 1991:17). Cependant, comme le précise Sinclair (1991:9), on ne peut jamais savoir si on a été suffisamment sélectif ou au contraire si on l'a trop été dans sa constitution de façon définitive en regard du but que l'on poursuit, on se doit donc de préciser au maximum les choix qui ont été faits lors de cette création, ainsi que leurs motivations.

La question de la fréquence est également très importante dès lors que l'on parle de représentativité. En effet, pour des linguistes se situant dans la filiation firthienne, on peut considérer que chaque acte de communication montre le système linguistique en opération, et la notion firthienne d'événement répété (la fréquence) est alors cruciale par rapport à la notion de généralisation. Il est en conséquence important pour le chercheur d'établir un lien important entre les critères de représentativité établis pour le corpus et le retour qu'il a dans son travail sur celui-ci, parfois, des résultats peuvent suggérer de modifier ces critères, quel

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

que soit le cadre théorique remis en question dans cet acte. Il peut alors être intéressant d'établir un premier ensemble de critères qui soient obtenus à partir d'une recherche théorique, puis ensuite, de revenir sur ces critères et de les modifier à partir d'une recherche empirique. Le travail sur corpus se fait de cette façon sur un mode cyclique, un premier corpus, appelé corpus pilote est assemblé, et après avoir effectué des recherches sur celui-ci, on peut affiner les critères d'assemblage pour créer un nouveau corpus plus en adéquation avec la réalité (Tognini-Bonelli 2001:61).

Comme nous venons de le voir, il y a une grande variété d'approches et d'opinions concernant cette question, l'essentiel pour le chercheur travaillant sur corpus semble alors d'être le plus précis possible quant aux raisons qui ont précédées le choix d'un corpus ou sa constitution :

« the important point is that workers in the field should be as explicit as possible, particularly about what they have done in assembling their corpus, and if possible some discussion about why they have taken these decisions » (Tognini-Bonelli 2001:63)³⁸

C'est ce que nous nous efforcerons de faire dans le troisième chapitre de cette thèse, en décrivant précisément le mode opératoire qui a été appliqué dans la création du corpus IntUne. Mais avant cela, nous allons voir que si les corpus peuvent prendre différentes formes et être adaptés à différentes approches, ils peuvent également donner lieu à plusieurs applications fructueuses dans des domaines aussi divers que la traduction ou l'enseignement des langues.

I.2.3- Différentes applications.

I.2.3.1- Différentes applications.

Les corpus multilingues décrits dans le paragraphe précédent peuvent avoir de nombreux atouts que ce soit au niveau de la linguistique générale ou appliquée. Des disciplines telles que la traduction ou la didactique des langues en linguistique appliquée sont notamment propices à l'utilisation efficace des corpus.

38 « Le point important est que toute personne travaillant dans ce domaine se devrait d'être aussi explicite que possible, particulièrement par rapport à ce qui a été fait lors de l'assemblage du corpus, et si possible, à certaines discussions qui ont eu lieu à propos des choix de tel ou tel décisions » (ma traduction).

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

1.2.3.1.1- Dans le domaine de la traduction, et de son enseignement.

Dans le domaine de la traduction, le corpus se révèle ainsi être un outil d'une grande valeur, permettant un nouveau mode de fonctionnement pour le traducteur, et une démarche de traduction plus efficace. L'acte de traduction suppose de trouver entre deux langues des unités équivalentes de sens, or, ces unités ne se correspondent pas forcément d'une langue à l'autre sous la forme de mots, comme on peut le constater à la lecture de tout dictionnaire bilingue. Tognini Bonelli (2001:134) note ainsi que ce n'est qu'après avoir établi une unité de signification complète et non ambiguë dans une langue 1 que l'on pourra tenter de la traduire dans une langue 2. Le corpus est un outil précieux dans la délimitation de ces unités de significations qui peuvent couvrir un ou plusieurs mots, en effet, il permet d'appréhender tout mot en contexte, c'est à dire à un niveau syntagmatique. En voyant quelles associations fréquentes entre un mot clef et son contexte peuvent correspondre à des sens précis, on peut circonscrire des unités de significations y correspondant, ce sens étant par la suite découvert au moyen d'une étude des différentes lignes de concordances les unes par rapport aux autres, à un niveau cette fois plus paradigmatique, permis par les concordanciers. Le corpus permet ainsi dans le domaine de la traduction de mieux circonscrire les unités de significations auxquelles le traducteur devra par la suite rechercher des équivalences en maintenant sa recherche sur des données de langage authentique.

Toujours dans ce domaine de la traduction, Kenny (2000:271) montre que les corpus sont des ressources de grande valeur. En effet, ils permettent de distinguer des tendances concernant les comportements et les stratégies mises en place par des traducteurs professionnels. Selon elle, les recherches orientées dans le domaine de la traduction et de la créativité lexicale peuvent grandement bénéficier de l'usage de preuves issues de corpus et de l'intégration de certaines notions issues de la linguistique de corpus comme les collocations, la préférence sémantique ou la prosodie sémantique (Ibid:281).

Pearson (2003) montre en outre que les corpus, et notamment les corpus parallèles peuvent être d'excellents auxiliaires dans le domaine de l'enseignement de la traduction. En effet, ces corpus parallèles, étant alignés de façon rigoureuse sur un axe syntagmatique, permettent aux étudiants d'analyser directement les stratégies mises en place par tel ou tel traducteur professionnel pour passer d'une langue à une autre (Pearson 2003:169). Les corpus comparables peuvent également apporter des informations à ces mêmes étudiants en

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

traduction, notamment pour avoir davantage d'informations sur la phraséologie naturelle de tel ou tel mot dans la langue traitée (Ibid:168).

1.2.3.1.2- Dans le domaine de l'apprentissage des langues étrangères.

Le corpus est également un outil qui peut revêtir une grande utilité dans le domaine de l'apprentissage des langues. D'une part, dans la mesure où des innovations technologiques ont eu entre autres conséquences à la fin du dix-neuvième siècle et au début du vingtième siècle une vague de réformes dans la didactique des langues (Véronique 1992:172), les corpus peuvent également influencer grandement les pratiques dans ce domaine en ce début de vingt et unième siècle. Henri Sweet insistait notamment lors du mouvement de réforme à la fin du dix-neuvième siècle sur le fait que toute étude, tout apprentissage d'une langue nouvelle devait se fonder sur des textes authentiques (Véronique 1992:179-180). Harold Palmer préconisait également une utilisation essentielle de textes authentiques dans une situation d'apprentissage d'une langue nouvelle (Crystal 1985:142). Ce besoin se fait toujours sentir à notre époque et l'accès à ces textes a été grandement facilité par la création de plus en plus importante de corpus dans de multiples langues (Kennedy 1998). L'intérêt d'utiliser des textes authentiques réside également dans le fait que c'est à cette langue que l'apprenant va se confronter, et qu'elle diffère bien souvent de la langue enseignée ou présente dans les manuels d'apprentissage, ces différences s'expliquant notamment par des procédures de découverte du fonctionnement de la langue largement introspective et ne se basant que trop peu sur des faits de langue réels (Tognini-Bonelli 2001:15).

L'utilisation de corpus dans une salle de cours est par ailleurs plutôt agréable pour les apprenants, elle permet la mise en place d'un apprentissage actif de l'étudiant, de type bottom-up (Ibid:14), l'étudiant n'a pas besoin d'une grande expertise dans la langue étudiée avant de pouvoir se confronter à des faits de langues authentiques. Cette plus grande participation de l'étudiant à son apprentissage contribue ainsi à faciliter l'acquisition de la langue étudiée, et contribue à lui donner une plus grande autonomie pour d'autres apprentissages à venir (Holec 1990).

L'analyse de corpus peut aussi permettre la création d'un matériel pédagogique plus adapté. Par exemple, dans le cadre de l'acquisition du vocabulaire d'une langue, Tognini-Bonelli (2001:40) note que l'utilisation du corpus, en ce qu'elle s'intéresse aux fréquences des

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

mots utilisés de façon naturelle, peut aider le linguiste à élaborer des syllabus beaucoup plus efficaces pour les apprenants. Sinclair et Renouf (1988) notent ainsi que ces nouveaux syllabus doivent effectivement prendre en compte les mots avec un fort usage, et les façons les plus communes de les utiliser.

En outre, l'usage du corpus permet de sensibiliser l'apprenant à la variation de la langue (Tognini-Bonelli 2001), il existe, comme nous l'avons précisé dans la partie précédente différents types de corpus, on peut noter parmi ceux-ci la présence de corpus de langue de spécialité qui permettent aux étudiants souhaitant apprendre un sous langage particulier de le faire avec une plus grande acuité :

« In the field of LSP, many corpora of restricted varieties are being assembled for teaching purposes very much in line with a shift towards vocational language training where students of economics, for example, are exposed preferentially to the type of language they will be using, i.e. The language of economic texts, journals and/or spoken negotiations. » (Tognini-Bonelli 2001:8)³⁹

L'utilisation du corpus compte ainsi de nombreux avantages dans divers domaines de la linguistique appliquée, qu'il s'agisse de traduction ou de lexicographie. Son emploi dans le domaine de la didactique des langues semble poser de nombreux atouts. Néanmoins, cet outil n'a rien d'un remède miracle à administrer en toute occasion sans sens critique, il connaît certaines limites que nous allons maintenant aborder de façon plus détaillée.

I.3- Conclusion.

La linguistique de corpus, bien qu'étant une discipline récente dispose d'une littérature abondante et diverse. Nous avons précisé dans ce premier chapitre ce que nous entendions par cette locution, en retraçant le cadre historique de la linguistique générale, et en remettant en perspective la tradition linguistique depuis Saussure, nous avons vu que la linguistique de corpus réunissait deux linguistiques qui avaient coexisté l'une à côté de l'autre

39 « Dans le domaine des langues de spécialité, beaucoup de corpus de variétés restreintes de langue sont rassemblés dans un but pédagogique plus en phase avec les besoins d'une formation professionnelle où par exemple, les étudiants en sciences économiques sont confrontés davantage aux types de langages qu'ils devront utiliser, c'est à dire le langage des textes économiques, des journaux, et/ou de négociations orales » (ma traduction).

1- Linguistique de corpus et linguistique générale, qu'est-ce qu'un corpus ?

tout au long du vingtième siècle : langue et parole se rejoignent sans s'exclure au sein de ce paradigme (Tognini-Bonelli 2001).

Savoir ce que l'on entend par corpus, ce qu'est cette entité dès lors que l'on se situe dans la discipline de linguistique de corpus est important pour pouvoir concevoir plus clairement ce sur quoi la recherche menée se fonde. Nous avons ainsi également éclairci ce point en notant que le corpus dans cette discipline est avant tout électronique, et que si différentes formes en existent, le corpus IntUne peut être considéré comme un corpus comparable. Nous reviendrons dans le troisième chapitre sur les détails de la conception de ce corpus.

Comme nous l'avons vu, diverses méthodologies sont envisageables lorsque l'on se situe dans le paradigme de la linguistique de corpus, si l'usage d'une approche corpus-driven a déjà été explicitée dans le cadre de notre étude (cf. partie 2.1.2.4), le chapitre suivant va aborder de façon plus détaillée et précise le cadre théorique et méthodologique dans lequel les recherches de cette thèse ont été menées, ainsi que les procédures d'analyse et d'étude qui y ont été mises au point.

2- Cadre théorique, méthodologie employée.

II.0- Introduction.

Le projet du groupe média IntUne est de voir comment l'Europe est représentée linguistiquement dans la presse de quatre pays : le Royaume-Uni, la Pologne, la France et l'Italie. Le choix de la presse, comme cela a déjà été expliqué dans l'introduction se justifie par l'influence que ce type de discours peut avoir sur la construction d'une identité pour des citoyens. Lors des chapitres précédents, nous avons vu quels étaient les rapports entre linguistique générale et linguistique de corpus, et ce que l'on entendait dans cette thèse par linguistique de corpus. Nous avons également précisé ce que l'on entendait par corpus. Mais à partir du moment où l'on dispose d'un corpus, il s'agit de s'interroger sur la méthodologie à suivre et les processus à mettre en place pour l'analyser.

Dans le cadre de cette étude, la méthode choisie pour analyser la représentation lexicographique et sociolinguistique de l'Europe dans la presse est d'étudier le ou les sens dont différents mots associés de façon thématique au concept d'Europe pourraient être affectés. En effet, Wierzbicka note que l'étude des mots permet d'appréhender les valeurs données par une culture à tels et tels concepts (tels l'Europe) :

*« Words are society's most basic cultural artefacts,
and -properly understood- they provide the best key*

2- Cadre théorique, méthodologie employée.

to a culture's values and assumptions »⁴⁰ (Wierzbicka 1996:237)

Notre approche, se basant sur le sens d'un mot, son association à un ou plusieurs concepts, est largement lexicographique. En cela, il est nécessaire d'explicitier une tradition longue de plusieurs siècles dans ce domaine qui pourrait influencer les méthodes, techniques, et concepts-outils utilisés lors de l'analyse.

Dans ce chapitre, nous aborderons donc le cadre théorique sur lequel la méthodologie employée dans cette étude se fonde. Tout d'abord, nous partirons de la théorie aristotélicienne des conditions nécessaires et suffisantes qui demeure une influence majeure dans le domaine de la lexicographie, mais qui pose également de nombreux problèmes. Nous verrons ensuite les réponses, à travers l'élaboration d'une théorie du prototype, qui ont pu être proposées à ces problèmes par la philosophie (via les travaux de Wittgenstein), la psychologie (via les travaux de Rosch), et la synthèse qui a été proposée pour cette théorie. Puis, nous aborderons dans une troisième partie une adaptation des prototypes par Patrick Hanks qui a proposé de les appliquer dans le domaine de la lexicographie. Enfin, nous construirons un premier prototype à l'aide d'une matière lexicographique pour le mot clef que nous étudions dans cette thèse « Europe » (Williams 2009) qui pourra être complété ou modifié à la lumière d'une analyse de corpus dans les chapitres suivants.

Tout d'abord, nous allons voir comment, depuis des années, des concepts ont été associés à une unité sémantique largement utilisée, le mot. Cette unité, bien que très répandue dans les approches lexicographiques (les mots sont les unités d'entrée de la plupart des dictionnaires), présente de nombreuses failles que nous allons explorer.

II.1- Le mot, une unité problématique.

Le mot est une des unités sur lesquelles se fonde notre acquisition du langage et des concepts qu'il permet d'appréhender⁴¹. Si quelqu'un cherche à comprendre ce que c'est que

40 « Les mots sont les artefacts culturels les plus basiques d'une société, et, correctement compris, ils fournissent la meilleure clef qu'il soit pour comprendre les valeurs et les positions d'une culture » (ma traduction).

41 Cf. les programmes de l'école maternelle, p. 165, l'enfant doit être capable de « identifier immédiatement les mots usuels des activités de la classe et du vocabulaire acquis dans le cadre d'autres enseignements (date, nom d'exercices ou d'activités (...) vocabulaire mémorisé dans les autres domaines » dans le cadre de son acquisition du langage. Les programmes de l'école primaire peuvent se consulter sur le site suivant :

2- Cadre théorique, méthodologie employée.

L'Europe, un des premiers réflexes sera probablement de rechercher une définition pour ce mot. Le langage qui permet d'élaborer des concepts, d'en rendre compte, de leur conférer telle ou telle signification se constitue de mots. Mais cette unité, si elle est très répandue et populaire, pose un réel problème de définition. En effet, comment circonscrire aussi bien au niveau graphique, phonétique que sémantique un objet dont la réalité ne peut s'approcher de façon empirique et systématique ?

Si la tradition populaire définit le mot comme étant *grosso modo* une unité de sens que l'on distinguerait par une séparation dans la chaîne langagière (une pause à l'oral, ou un blanc au niveau typographique), la réalité de ce concept apparaît, dès lors qu'on l'interroge plus avant, beaucoup plus complexe à circonscrire (Crystal 1985:185-9). En effet, de nombreuses langues n'utilisent aucun signe typographique pour séparer les mots les uns des autres, quand bien même ce blanc typographique existe, des situations mixtes peuvent surgir telles que l'emploi de l'apostrophe dans les langues latines. Au niveau phonétique, il est quasiment impossible de distinguer les différents mots d'un énoncé que l'on entendrait dans une langue inconnue, alors que d'autres unités, comme le phonème, semblent pouvoir être analysées beaucoup plus facilement. Au niveau sémantique, enfin, il semble complexe d'attribuer un voire plusieurs sens bien circonscrits à un mot en particulier, que ce soit parce que cette tâche demande un long travail de recherche (Wierzbicka 1996:256) ou parce que la nature de la signification rend tout simplement cette recherche impossible (Hanks 2000:205). Cette unité pose donc de nombreux problèmes d'identification et de définition. Une longue tradition linguistique s'est formée pour tenter de répondre à ces difficultés. Il est intéressant de passer en revue certaines des solutions qui ont pu être proposées pour expliquer les mots dès lors que l'on souhaite étudier la représentation de l'Europe dans la presse française. Dans la mesure où l'étude porte sur un concept représenté dans la langue par un mot, il est essentiel d'avoir présent à l'esprit les diverses méthodes d'analyse de ces unités, les mots, qui ont été élaborées. Si la réflexion sur le mot est probablement aussi ancienne que celle sur le langage, nous partirons dans cette partie d'une théorie encore largement répandue aujourd'hui, le modèle aristotélicien des conditions nécessaires et suffisantes, nous verrons quels sont ses avantages pour rendre compte de la façon dont un concept est présenté sous une forme linguistique, mais également les limites dont cette théorie peut souffrir.

II.1.1- Aristote, la théorie des conditions nécessaires et suffisantes.

Tenter d'appréhender le sens d'un concept, la représentation qui peut être donnée à ce concept sous forme de mot est quelque chose de complexe : selon qu'on se situe au niveau psychologique, on s'attachera à une valeur interne du concept. Si l'on veut une définition scientifique et opérationnelle, le modèle d'une formalisation intensionnelle semblera plus approprié : une description formelle par partie sera en effet plus efficace qu'une description extensionnelle dans laquelle des sens seraient ajoutés à un mot. Le sens en philosophie, enfin, s'approche davantage de la recherche d'une vérité (Wierzbicka 1996:258), recherche d'une vérité que le modèle des conditions nécessaires et suffisantes d'Aristote s'est efforcé de mettre en application. La théorie classique aristotélicienne est celle qui a dominé la psychologie, la philosophie et même la linguistique (que l'on soit dans un paradigme structuraliste ou générativiste) jusqu'à nos jours (Hanks 1994:90).

Aristote a élaboré une théorie de la classification pour répondre au problème de la définition du concept. Il a ainsi distingué pour un concept ce qui était du domaine de l'essence, et ce qui était du domaine de l'accident. L'essence correspond à ce qui ne peut manquer à un concept sans que celui-ci n'existe plus, et l'accident à ce qui peut être dit de façon vraie du concept, mais qui peut manquer à sa définition sans en affecter le sens. Dans ce cadre, définir un concept revient à trouver les conditions qui lui sont essentielles. Pour dire qu'un objet A du réel est le concept B, il faut s'assurer que A vérifie toutes les conditions essentielles de B.

Aristote pose dans cette théorie la chose, qui est le *definiendum* et la définition, le *definiens*. Le *definiens* se compose d'un *genus proximum* et d'une *differentia specifica*, c'est-à-dire que la chose X est en fait équivalente à Y (le *genus proximum*, qui est la chose la plus proche à un niveau hiérarchique supérieur de ce que l'on veut définir) qui se différenciera par une caractéristique spécifique. Pour reprendre le célèbre exemple de Platon, la chose « homme » pourrait avoir pour *definiens* « bipède sans plume » où « bipède » serait une forme de *genus proximum*, et « sans plumes » la *differentia specifica*⁴². Si un être veut pouvoir nommer un concept, il doit avant tout en connaître la définition. Ainsi, si un enfant par exemple veut pouvoir nommer un chat, il doit en connaître une définition vérifiant les conditions nécessaires et suffisantes pour être un chat, c'est-à-dire qu'il doit savoir que pour

42 Cours donné par Williams dans le cadre d'un Master II à l'Université de Bretagne Sud en 2008.

2- Cadre théorique, méthodologie employée.

qu'un animal soit chat, il doit par exemple s'agir d'un mammifère de la famille des félidés, muni de griffes rétractiles, nyctalope, qui peut être domestique ou sauvage. S'il ne connaît pas l'ensemble de ces conditions nécessaires et suffisantes, il risque fort, selon cette théorie, de mal nommer ou concevoir ce qu'il aperçoit.

Le *genus proximum* répond à une conception du sens intensionnelle, c'est-à-dire qu'il différencie les choses se situant au sein d'une même classe. Cette idée même d'un sens intensionnel suppose en conséquence l'existence d'un ordre naturel dans lequel des catégories existent. Pour définir un concept, il suffit alors de relever tous ses traits caractéristiques, toutes les conditions qu'il remplit, puis ensuite de vérifier à quelle catégorie il appartient du fait de ces conditions. L'ensemble des conditions nécessaires et suffisantes du concept devient le « sens » du mot. Kleiber (1990) note ainsi que selon ce modèle :

« (i) Les concepts ou catégories sont des entités aux frontières clairement délimitées »

(ii) L'appartenance d'une entité particulière à une catégorie répond au système du vrai ou faux : un x quelconque est un chien ou n'est pas un chien, selon qu'il satisfait ou non aux conditions critères de la catégorie chien.

(iii) Les membres d'une même catégorie ont un statut catégoriel égal, puisque chaque membre possède les propriétés requises par la définition de la catégorie. Eu égard à la catégorie à laquelle ils appartiennent, chaque membre est aussi « bon » membre que les autres. » (Georges Kleiber 1990:22)

Cette théorie est donc plutôt confortable, elle propose un modèle strict et précis du fonctionnement de la définition. En outre, elle règle le problème de la catégorisation des concepts, comme l'ajoute Kleiber plus loin :

« Le problème de l'appartenance à une catégorie se trouve réglé : pour décider de l'appartenance d'un objet à une catégorie, il suffit de vérifier si cet objet possède les CNS de cette catégorie ou non.

Le sens ou intension (c'est-à-dire le faisceau des CNS) détermine l'extension ou la référence. Le modèle des CNS est donc conforme à la règle fréguenne de la détermination de la référence par le sens. Le corollaire en est que, pour pouvoir utiliser un mot, il faut en connaître le sens (c'est-à-dire les CNS). » (Ibid 1990:25)

2- Cadre théorique, méthodologie employée.

Ce modèle répond en fait de façon précise et explicite à la double intuition ou croyance qu'il existe un sens exact pour chaque mot, et que les catégories sont des entités discrètes que l'on peut aisément identifier.

En effet, le modèle aristotélien, en posant les conditions nécessaires et suffisantes comme définissant le concept, pose aussi le fait qu'une chose ne peut pas être et ne pas être, qu'elle ne peut pas appartenir à une catégorie et ne pas y appartenir : une chose possède ou non une caractéristique, une condition, en fonction de quoi elle appartient ou non à une catégorie (Taylor 1995:23). Ces conditions, ces traits caractéristiques sont donc binaires elles sont présentes ou non dans la définition du concept en vertu de quoi, les frontières des catégories sont nettes et précises.

« A category, once established, divides the universe into two sets of entities – those that are members of the category, and those that are not. There are no ambiguous case, no entities which « in a way » or « to some extent » belong to the category, but which in another way do not. » (Taylor, 1995:23)⁴³

Enfin, cette théorie aristotélienne pose le fait que tous les membres de la catégorie ont un statut égal : il n'y a pas un membre plus proéminent au sein de la catégorie qu'un autre, pas de degrés d'appartenance à cette catégorie, un chat est un mammifère, au même titre qu'une baleine ou qu'une chauve-souris. Nous allons maintenant voir comment cette théorie a influencé la pratique linguistique notamment au long du vingtième siècle.

II.1.2- L'application de ce modèle dans la linguistique de la seconde moitié du vingtième siècle.

Lors de la deuxième moitié du vingtième siècle, ce modèle a continué à influencer la linguistique générale, notamment du fait de sa fortune dans le domaine de la phonologie. Le but était alors de montrer que le système phonologique de toute langue pouvait se décrire et s'expliquer selon cette idée de conditions nécessaires et suffisantes qui, une fois structurées en un ensemble, pouvaient former une catégorie se comportant au sein d'un système.

43 « une catégorie, une fois qu'elle est établie, divise l'univers en deux ensembles d'entités – celles qui sont des membres de la catégorie, et celles qui n'en sont pas. Il n'y a pas de cas ambigus, pas de catégories qui « d'une certaine façon » ou « dans une certaine mesure » appartiennent à la catégorie mais n'y appartenant pas d'une autre façon ». (ma traduction).

2- Cadre théorique, méthodologie employée.

Ainsi, toute langue, au niveau phonologique peut s'appréhender, qu'on la connaisse ou non en des unités discrètes, les phonèmes, selon la théorie formaliste. Chaque être humain peut séparer une chaîne de langage en ces unités discrètes (les phonèmes) qu'il sera ensuite possible de classer dans différentes catégories selon les conditions nécessaires et suffisantes qu'elles remplissent ou non (Jakobson 1973:123). Les conditions nécessaires et suffisantes prennent ici la forme de traits distinctifs. Ces traits distinctifs ont quatre principales caractéristiques, ils sont primitifs, c'est-à-dire qu'il est impossible de descendre à un niveau inférieur, on peut les comparer en cela au concept de l'atome. Deuxièmement, ils sont universels, chaque langue humaine peut se définir selon ces traits distinctifs au niveau de la phonologie. Troisièmement, ils sont abstraits (Jakobson 1962:294).

Cette troisième assomption peut poser problème, on peut se demander pourquoi on ne se base pas sur l'observable pour catégoriser ces unités, Taylor (1995:27) résume ce point en expliquant que se baser sur des critères de l'observable, plutôt que sur des critères élaborés de façon abstraite est beaucoup moins économique pour rendre compte de la façon dont ces unités fonctionnent dans le système. Quatrièmement, enfin, ces traits distinctifs, selon le paradigme générativiste, sont innés. Cette dernière caractéristique découle logiquement des autres, puisque ces traits sont abstraits, et qu'ils sont partagés de façon universelle, il semble naturel de leur donner une sorte de réalité psychologique qui préexiste à tout apprentissage.

Ce modèle a donc été appliqué et élaboré avec une certaine fortune à la phonétique. La sémantique se l'est également approprié à travers la constitution de catégories sémantiques dans lesquelles on a pu distinguer des traits distinctifs ou sèmes. Pour prendre un exemple, le mot anglais « bachelor » devrait vérifier les sèmes [+animé], [+mâle], [+humain], [+adulte], [-marié], etc. Le modèle aristotélicien ainsi développé apporte de cette façon une clarté binaire rassurante et de nombreuses informations sémantiques sur les concepts ou les énoncés. Cependant, nous allons voir que ce modèle a posé de nombreux problèmes qui ont été mis en exergue et auxquels des solutions ont été recherchées dans la seconde moitié du vingtième siècle, notamment.

II.1.3- Les limites du modèle des CNS dans le domaine de la sémantique et de la lexicographie.

Malgré le confort apporté par cette théorie, des critiques peuvent lui être adressées. D'une part, l'acte de classification peut certes être scientifique, mais il peut aussi être plus subjectif, influencé par le cadre culturel. Par exemple, pourquoi choisir de mettre les choses « chat » ou « chien » dans une catégorie « animaux domestiques », plutôt que dans une catégorie « outil de travail » ou « bétail » ? La définition d'un siège avec la théorie aristotélicienne pourrait être « meuble pour s'asseoir », mais s'agirait-il nécessairement d'un « fauteuil » dès lors que l'on y ajouterait des accoudoirs⁴⁴ ? Cette théorie n'est pas le résultat d'une étude empirique, mais elle est plutôt une position philosophique à laquelle on est parvenu à la suite d'une spéculation *a priori* (Lakoff 1987:6). De ce fait, elle pose certaines difficultés que nous allons maintenant évoquer.

Dans un premier temps, outre que la condition nécessaire et suffisante demande en premier lieu à être vérifiée, c'est un modèle qui possède un pouvoir explicatif fort, mais qui a davantage de limites au niveau descriptif : savoir qu'un homme est un bipède sans plumes ne nous renseigne que peu sur ce qu'est ce concept, savoir qu'un chat est un « mammifère carnivore au museau court et arrondi, aux griffes rétractiles, dont il existe des espèces domestiques et des espèces sauvages de la « famille des félidés » »⁴⁵ ne sera que de peu d'intérêt pour expliquer à un jeune enfant ce qu'est cet objet s'il n'en a jamais vu. Un des problèmes de ce modèle est donc son faible pouvoir descriptif (Kleiber 1990).

La principale limite de ce modèle ne semble cependant pas résider dans sa difficulté à décrire le concept dont il est question ni dans le besoin où il est de vérifier la condition nécessaire et suffisante, mais dans sa rigidité. Ainsi, les tenants du modèle aristotélicien prennent parfois l'exemple anglais « bachelor » pour expliquer que ce mot fonctionne très bien avec la théorie des conditions nécessaires et suffisantes (CNS), puisque pour entrer dans ce concept, il suffirait d'être un homme adulte non marié. Mais quand bien même cela pourrait fonctionner pour un mot, cela n'est pas le cas pour la majorité des mots d'une langue, Aitchison note ainsi :

« There are a small number of words such as

44 Cours donné par Williams dans le cadre d'un Master II à l'Université de Bretagne Sud en 2008.

45 Définition donnée par le Petit Larousse 1992 pour le mot chat en première position.

2- Cadre théorique, méthodologie employée.

bachelor which appear to have a fixed meaning, that is, they are words for which we can specify a set of necessary and sufficient conditions. The majority of words, however, do not behave in this way. They suffer from one or more of the following problems: first, it may be difficult to specify a hard core of meaning at all. Second, it may be impossible to tell where « true meaning » ends and encyclopaedic knowledge begins. Third, the words may have « fuzzy boundaries » in that there might be no clear point at which the meaning of one word ends and another begins. Fourth, a single word may apply to a « family » of items which all overlap in meaning but which do not share any one common characteristic. » (Aitchison, 1987:49)⁴⁶

Et selon Hanks, c'est encore accorder trop d'efficacité au modèle des CNS, puisque « bachelor » lui même pose de nombreux problèmes dès lors qu'on définit ce concept d'un simple « homme adulte non marié ». Que dire alors d'un « vieux garçon » ou d'un homme qui est en couple homosexuel ? Si le modèle des CNS peut parfois apporter beaucoup de clarté à la conception des choses, il apparaît comme non satisfaisant quand il est appliqué tel quel pour définir les mots courants d'une langue vivante, dont les usages réels sont en mouvement constant et dont les frontières sémantiques risquent fort d'être floues, par conséquent. Hanks note qu'au niveau de la lexicographie :

« Any attempt to write a completely analytical definition for any common word in a natural language is absurd. Experience is far too diverse for that. What a good dictionary offers instead is a typification : the dictionary definition summarizes what the lexicographers finds to be the most typical common features in his [or her] experience, of the use, context, and collocations of the word. » (Hanks, 1994:91)⁴⁷

46 Aitchison J., (1987), *Words in the Mind*, Blackwell. « Il y a peu de mots tels que « bachelor » qui semble avoir un sens fixe, c'est-à-dire qu'il y a peu de mots pour lesquels on puisse spécifier en ensemble de conditions nécessaires et suffisantes. La majorité des mots, cependant, ne se comporte pas ainsi. Il souffrent d'un ou plusieurs des problèmes suivants : premièrement, il peut être difficile de spécifier un noyau dur de sens. Deuxièmement, il peut être impossible de dire à partir de quand le « vrai sens » s'arrête et où la connaissance encyclopédique commence. Troisièmement, les mots peuvent avoir des frontières floues en ce qu'il peut ne pas y avoir un point clairement établi où le sens d'un mot commence et celui d'un autre termine. Quatrièmement, un seul mot peut s'appliquer à une famille d'objets qui se chevauchent au niveau des sens, mais qui ne partagent aucun caractéristique commune » (ma traduction)

47 « Toute tentative d'écrire une définition complètement analytique pour n'importe quel mot d'une langue naturelle est absurde. L'expérience est par trop diverse pour cela. Ce qu'un bon dictionnaire fournit à la place est une typification : la définition du dictionnaire résume ce que le (ou la) lexicographe trouve comme

2- Cadre théorique, méthodologie employée.

C'est pourquoi de nombreux chercheurs se sont attachés à trouver une théorie permettant davantage de flexibilité, ainsi qu'une plus grande adéquation pour décrire les concepts tels qu'ils sont utilisés et compris par la majorité des personnes, c'est ce que nous allons voir dans cette seconde partie avec le modèle des prototypes, tel qu'il a été élaboré.

Dans cette thèse, c'est également à travers cette volonté de rendre compte de la flexibilité et de la multiplicité des approches d'un concept, l'Europe, que nous avons fait usage d'une méthode s'inspirant d'un modèle plus flexible : nous allons maintenant tenter de la définir. Dans cette seconde partie, nous allons rapidement évoquer dans un premier temps la réponse philosophique qui a été proposée par Wittgenstein à la théorie aristotélicienne des CNS via le concept de « ressemblance de familles ». Puis, nous verrons comment cette solution philosophique a influencé la psychologie à travers les travaux de Rosch qui a proposé une théorie du prototype. Enfin, nous étudierons la synthèse qui en a été proposée par Taylor et Kleiber.

II.2- La théorie des prototypes.

Nous venons de voir que le modèle des conditions nécessaires et suffisantes, tel qu'il a été conçu puis élaboré, ne suffit pas pour rendre compte de l'acte humain de catégorisation, de compréhension du réel. Certains concepts semblent en effet poser des problèmes pour être intégrés dans ces catégories, et pour remplir de façon claire et précise les conditions nécessaires et suffisantes. Une réponse philosophique a été apportée à ces difficultés dans les travaux de Wittgenstein qui a mis au point le concept de « l'air de famille » après avoir constaté que les explications, selon le modèle des CNS ou en fonction de sèmes, ne fonctionnaient pas dans la pratique si elles étaient confrontées aux mots.

Nous allons dans un premier temps expliciter cette théorie, puis nous verrons comment elle a conduit à son développement dans le domaine de la psychologie avec Rosch, ainsi que la synthèse qu'il en a été proposé. Enfin, nous verrons quelles failles cette théorie peut rencontrer.

étant les traits les plus typiques dans son expérience sur l'usage, le contexte, et les collocations du mot. »
(ma traduction).

II.2.1- Wittgenstein et l'air de famille.

Ludwig Wittgenstein a notamment porté ses investigations dans le domaine du langage dans les années 1950. S'étant interrogé sur les différentes possibilités que l'être humain a pour définir un mot ou un concept à un de ses congénères, il rejette la définition ostensible, qui consiste à montrer un objet. Par exemple, si on montre deux noix à quelqu'un en lui expliquant que « ceci est deux », il pose la question de savoir comment être sûr que la personne concernée va bien comprendre la réalité du concept « deux » et ne pas interpréter « deux » comme une couleur, ou un fruit sec ? On peut objecter que si la « catégorie » de ce qui est nommé est précisée, la personne risque de moins se tromper, mais comment définir cette catégorie ? La plupart du temps, on la définit au moyen d'autres mots, la difficulté est que l'on tombe alors dans une circularité de la définition qui rend problématique l'identification précise des concepts désignés.

Wittgenstein propose alors comme solution le fait d'identifier la signification du mot à son usage : le sens du mot se trouvera dans son usage en contexte, et de fait sera relié aux actes de la vie humaine. Sa définition n'a pas nécessairement à être précise, mais elle peut néanmoins exister, et a une utilité (Wittgenstein 2004).

C'est dans ce sens que l'idée des ressemblances de familles prend toute sa dimension. Cette idée repose sur un exemple, celui du mot « game » :

«Consider for example the proceedings that we call « games ». I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all ? - Don't say: « There must be something common, or they would not be called « games » - but look and see whether there is anything common to all. - For if you look at them you will not see something that is common to all, but similarities, relationships, and a whole series of them at that. (...) And the result of this examination is: we see a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail. » (Wittgenstein 1953:31) ⁴⁸

48 « Considère, par exemple, les processus que nous nommons « jeux ». Je veux dire les jeux de pions, les jeux de cartes, les jeux de balle, les jeux de combat, etc. Qu'ont-ils tous de commun ? - Ne dis pas : « Il doit y avoir quelque chose de commun à tous, sans quoi ils ne s'appelleraient pas des « jeux » »- mais regarde s'il

2- Cadre théorique, méthodologie employée.

En prenant cet exemple, Wittgenstein montre qu'on ne peut trouver une caractéristique qui soit commune à tous les membres de la catégorie « game ». On peut penser au fait de gagner ou de perdre, mais ce trait tombe lorsqu'on parle d'un enfant qui joue seul à lancer sa balle ou d'un groupe d'enfants récitant une comptine. On peut encore penser au fait d'exercer une certaine habileté, mais il est notable que l'habileté nécessitée pour jouer au tennis diffère sensiblement de celle requise dans une partie de jeux d'échecs. A partir de ce constat, se pose alors la question de savoir pourquoi nous appelons malgré tout ces concepts « game » et comment nous reconnaissons que telle chose est un jeu. Selon Wittgenstein, c'est parce que cette « catégorie » prend la forme d'un continuum où un premier jeu, par exemple les jeux de pions, peut posséder des caractéristiques AB, un deuxième jeu (par exemple, les jeux de carte) posséder cette fois les caractéristiques BC et un troisième jeu (par exemple, le tennis) posséder les caractéristiques CD. Le premier jeu n'a plus aucune caractéristique de commune avec le troisième, cependant, on note une proximité entre ces deux choses, ce que Wittgenstein appelle un « air de famille ».

Cette expression signifie alors que parmi les différents membres du concept « game », il y a des ressemblances existantes qui se chevauchent et s'entrecroisent dans un réseau complexe. Pour renforcer cette idée d'un continuum, il utilise une métaphore où la catégorie est vue comme la technique de filage :

« Et nous étendons notre concept de nombre de la même façon que nous enroulons, dans le filage, une fibre sur une autre. Or la solidité du fil ne tient pas à ce qu'une certaine fibre court sur toute sa longueur, mais à ce que de nombreuses fibres se chevauchent. » (Wittgenstein 1953/2004:65)

De cette façon, le concept est une entité relativement floue, au contour imprécis. Pour autant, cela ne pose pas de problèmes, du moins pas plus que d'y donner de façon abstraite une frontière précise qui serait en désaccord avec la réalité. Donner des limites au concept « game », par exemple, peut être juste dans un certain contexte, mais se révéler totalement impropre dans un autre, en règle générale, ce concept a des frontières floues. Mais cela n'est en rien incompatible avec le fait d'utiliser le concept de façon efficace : la

y a quelque chose de commun à tous.- Car si tu le fais, tu ne verras rien de commun à tous, mais tu verras des ressemblances, des parentés, et tu en verras toute une série. (...) Et le résultat de cet examen est que nous voyons un réseau complexe de ressemblances qui se chevauchent et s'entrecroisent. Des ressemblances à grande et à petite échelle. » (traduction française de Wittgenstein (2004), Recherches Philosophiques, Gallimard, p. 64)

2- Cadre théorique, méthodologie employée.

plupart du temps, le concept « game » est employé sans que l'on en délimite précisément les contours, et cela n'entrave pas la compréhension que l'on peut avoir de ce concept :

« Qu'est-ce qui est encore un jeu, et qu'est-ce qui n'en est plus un ? Peux-tu indiquer des limites ? Non. Tu peux tracer certaines limites, car jusqu'ici aucune n'a été tracée. (Tu n'as pourtant jamais ressenti aucune gêne lorsque tu as appliqué le mot « jeu ».) » (Wittgenstein 1953/2004:65)

Le mot ou concept « game » n'est en effet pas délimité sous tous ses rapports par des règles ou des critères, mais on peut l'utiliser de façon efficace, tout comme le jeu de tennis ne stipule pas si la balle doit être propulsée à une vitesse minimale ou maximale sans que la pratique en soit rendue impossible. Ce n'est pas parce qu'on ne peut pas définir avec « exactitude »⁴⁹ un concept que l'on ne peut pas l'utiliser ; le concept flou n'est pas forcément inutile.

En fait, on peut très bien définir un concept de façon « exacte » ou « inexacte » à la fois quand la réalité que l'on cherche à appréhender prend la forme d'un continuum parfait où les limites semblent impossible à trouver, mais il est alors essentiel selon Wittgenstein de se rappeler de quelle façon ce concept nous a été enseigné lorsqu'on cherche à le délimiter, cet enseignement ayant nécessairement une forte influence sur la perception que l'on en a :

« Quand tu butes sur cette difficulté, demande-toi toujours : Comment avons-nous appris la signification de ce mot (« bon », par exemple) ? Sur quel type d'exemples ? Dans quels jeux de langage ? (Tu verras alors plus facilement que le mot doit avoir toute une famille de significations.) » (Wittgenstein 1953/2004:70)

Ainsi, les nombreux exemples de « jeu » que l'on connaît nous permettent d'appréhender la réalité de ce concept, et il en va de même pour tous les concepts que l'on a rencontrés, c'est à travers des exemples que nous les percevons. Ces différents exemples, ces différents usages des concepts interviennent alors directement dans notre compréhension du réel, et forment l'air de famille dont il est question.

49 Sur le propos d'exactitude, Wittgenstein note encore que « Est-ce seulement ainsi que nous rendons un concept utilisable ? Certainement pas ! Si ce n'est dans un but particulier. Pas plus que la mesure de longueur « 1 pas » n'a été rendue utilisable par la définition : 1 pas) 75 cm. Et si tu disais : « Mais avant cette définition, ce n'était cependant pas une mesure exacte », je te répliquerais « En effet. - Encore que tu me sois redevable d'une définition de l'exactitude. » (2004:66)

2- Cadre théorique, méthodologie employée.

Cette pensée a inspiré la théorie des prototypes de la psychologue Eleanor Rosch, des linguistes comme Lakoff s'y sont également intéressés, l'idée de contour imprécis de la catégorie a notamment été reprise. Cependant, une des différences importantes entre ces deux modèles va résider dans le fait que le prototype suppose une périphérie et un cœur, contrairement au continuum de l'air de famille où pas un « jeu » n'est plus « jeu » qu'un autre, où il n'y a pas de représentant prototypique.

II.2.2- Le prototype en psychologie, l'apport d'Eleanor Rosch.

Le modèle des prototypes est une théorie qui a été élaborée par une psychologue, Eleanor Rosch, après les travaux sur l'air de famille de Wittgenstein. Taylor et Kleiber reprendront les différents travaux qui ont été faits sur ce modèle dans deux ouvrages de synthèse sur la question. Nous allons maintenant voir comment cette théorie a été pensée et élaborée dans le domaine de la psychologie .

Eleanor Rosch a pensé et mis en œuvre un protocole expérimental qui visait à mieux comprendre ce que les personnes ont à l'esprit quand elles utilisent des mots qui se réfèrent aux catégories. Ainsi, en pratiquant ces expériences sur environ 200 étudiants de son université, elle a cherché à savoir si des oiseaux pouvaient être davantage « oiseau » que d'autres, si des chiens étaient davantage « chien » que d'autres, etc. S'adressant à son public d'étudiants, elle propose la question suivante :

« Let's take the word red as an example. Close your eyes and imagine a true red. Now imagine an orangish red... imagine a purple red. Although you might still name the orange red or the purple red with the term red, they are not as good examples of red... as the clear « true » red. In short, some reds are redder than others. The same is true for other kinds of categories. Think of dogs. You all have some notion of what a « real dog », a « doggy dog » is. To me, a retriever or a German shepherd is a very doggy dog while a Pekinese is a less doggy dog. Notice that this kind of judgement has nothing to do with how well you like the thing; you can like a purple red better than a true red, but still recognize that the colour you like is not a true red. You may prefer to own a Pekinese without thinking that it is the breed that best represents what people mean by

2- Cadre théorique, méthodologie employée.

dogginess. » (Eleanor Rosch 1975:198)⁵⁰

A partir de ces questions, un protocole expérimental est construit. La première étape consiste en un questionnaire d'une dizaine de pages que les étudiants remplissent. Une page représentant une catégorie, parmi les dix catégories expérimentées, se trouvent notamment celle des oiseaux, des légumes, des fruits, les vêtements, des armes, des meubles, etc. Chaque page se compose du nom de la catégorie, puis, en dessous, d'une cinquantaine d'exemples de choses pouvant être ou non des membres de la catégorie. Ainsi, pour la catégorie « fruit », on peut trouver les membres orange, citron, pomme, pêche, melon, etc. Chaque liste présente ces exemples dans un ordre différent afin que l'ordre de présentation des membres n'influence pas les résultats fournis par les différents étudiants.

Les étudiants participant à l'expérience donnent alors une « note » à chacun des cinquante membres proposés, selon qu'ils sont de bons représentants de la catégorie ou non. La note allant de un pour un excellent représentant de la catégorie à sept pour un représentant beaucoup moins bon, dont l'appartenance même à cette catégorie pourrait se discuter.

Les résultats ainsi rassemblés ont montré une grande cohérence. Le consensus était d'autant plus élevé sur les membres qui obtenaient les meilleures notes en tant que bon exemple de la catégorie. Par exemple, l'accord était général sur le fait qu'un rouge-gorge était un membre excellent pour la catégorie « oiseau », qu'un « petit pois » était un bon exemple de « légume » ou qu'une « chaise » était un bon exemple de « meuble ». De même, considérant la catégorie « oiseau », il apparaissait avec un fort consensus que des membres comme « autruche », « pingouin » ou « chauve-souris » avaient un degré d'appartenance à cette catégorie beaucoup plus discutable. Des résultats semblables ont été notés dans les différentes catégories, qu'il s'agisse des armes (où les pistolets sont davantage des armes que les arcs) ou des vêtements (où les T-shirts sont davantage des vêtements que les chaussures).

50 « Prenons comme exemple le mot « rouge ». Fermez vos yeux, et imaginez un vrai rouge. Maintenant, imaginez un rouge orangé... Imaginez un rouge violacé. Bien que vous appelez toujours le rouge orangé ou violacé du terme rouge, ce ne sont pas d'aussi bons exemples du terme « rouge » que le clairement « vrai » rouge. En bref, certains rouge sont plus rouge que d'autres. La même chose est vraie pour d'autres sortes de catégories. Pensez aux chiens. Vous avez tous la notion de ce qu'est un « vrai chien », un « chien très chien ». Pour moi, un retriever ou un berger allemand sont des « chiens très chien », tandis qu'un pékinois n'est un peu moins « très chien ». Notez que ce type de jugement n'a rien à voir avec le fait d'aimer ou non la chose ; vous pouvez aimer un rouge violacé davantage qu'un vrai rouge, mais vous reconnaîtrez tout de même que la couleur que vous aimez n'est pas un vrai rouge. Vous pouvez préférer posséder un pékinois sans penser que c'est l'espèce qui représente le mieux ce que les gens entendent par « chien très chien ». (ma traduction)

2- Cadre théorique, méthodologie employée.

La même expérience a été menée sur la côte Est des États-Unis, afin de vérifier si les résultats ne pouvaient pas être influencés par l'environnement dans lequel les étudiants évoluaient, mais comme les conclusions furent sensiblement les mêmes, Rosch a pu exclure le fait que ces résultats auraient été une simple réaction particulière des étudiants californiens.

D'autres expériences ont ensuite été menées, toujours dans le but de confronter les résultats aux premiers obtenus. Une autre expérience était notamment de faire dire aux étudiant si tel objet faisait ou non partie de la catégorie, et de voir, en fonction du membre potentiel, quel était le temps de réponse de l'étudiant. Par exemple, un étudiant devait dire si un merle était ou non un oiseau, même chose pour une autruche, et en fonction de son temps de réponse on pouvait déduire de la plus ou moins grande évidence de l'appartenance de ce membre à la catégorie. Là encore, les résultats étaient cohérents.

Rosch (1978) a émis une réserve sur le fait que les étudiants auraient pu répondre plus rapidement simplement lorsqu'il s'agissait de mots plus communs, on rencontre plus fréquemment des merles que des autruches. Néanmoins, dans le cas de la catégorie « meuble », des membres tels que « ottomane » arrivaient avant d'autres membres comme « réfrigérateur ». Dans ce cas, il est manifeste que la fréquence des mots étudiés n'impactait pas les résultats. En règle générale, donc, s'il peut être gardé à l'esprit que la fréquence d'un membre peut influencer son résultat, cela est loin d'être systématique. C'est une réelle intuition plutôt cohérente chez les différents étudiants ayant participé à l'expérience de ce qui est membre d'une catégorie ou ne l'est pas qui a ainsi pu être mise en évidence. D'autre part, on aurait pu penser que les membres d'une catégorie avaient une proximité visuelle, par exemple, si le petit pois était un des meilleurs représentants de la catégorie légumes, alors, on aurait pu en déduire que les étudiants, influencés par cet élément, auraient catégorisé d'autres légumes ressemblant visuellement au petit pois comme étant aussi de bons représentants de la catégorie, hors, cela n'a pas été observé. Les carottes bien que différentes en tout point visuellement des petits pois étaient un autre des meilleurs représentants de la catégorie légumes. La proximité visuelle n'a donc pas plus influencé de façon systématique cette classification que la forte fréquence d'un item.

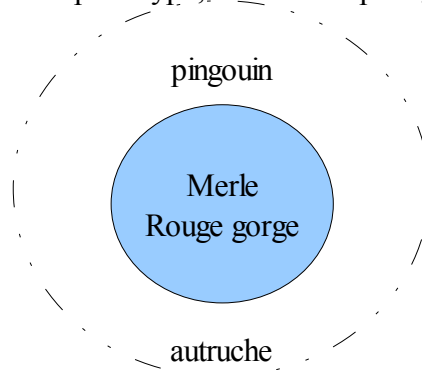
Au niveau fonctionnel, la proximité entre différents items n'a pas davantage influencé leur classement au sein de la catégorie par les étudiants. Ainsi, on aurait pu croire

2- Cadre théorique, méthodologie employée.

que si les chaises sont de bons membres pour la catégorie « meuble », alors les tabourets ou les bancs auraient pu avoir un comportement semblable, mais au niveau des résultats de l'expérimentation, il est apparu que cela n'était pas le cas, des objets comme les étagères à livres apparaissant bien plus proches de la chaise.

Cet ensemble d'expérience a donc amené Eleanor Rosch à formuler différents postulats. D'une part, cette classification était avant tout effectuée grâce à une intuition des étudiants, et non pas à divers autres facteurs extérieurs qu'ils avaient vis à vis de la catégorisation du réel. Cette catégorisation étant relativement homogène pour cette population. D'autre part, selon ces résultats, Eleanor Rosch a déduit l'existence de prototypes. C'est-à-dire qu'il existe différentes catégories qui ont des frontières floues et perméables, et au centre de chacune de ces catégories, se trouve un prototype qui représente très bien la catégorie à laquelle il appartient. De cette façon, lorsqu'un enfant souhaite définir une catégorie, il pourra utiliser par métonymie son membre prototype et ainsi expliquer qu'un légume, c'est comme un petit pois ou une carotte. Le prototype est vu ici comme l'exemplaire en quelque sorte idéal de sa catégorie, et on mesure à l'aune de ce prototype si d'autres objets figurent en plus ou moins bonne place au sein de la catégorie.

En termes de conditions nécessaires et suffisantes, cela reviendrait à dire que pour être un bon prototype, il faut réunir toutes ou du moins un maximum des conditions nécessaires et suffisantes de la catégorie, et que certains exemplaires de la catégorie réunissant un nombre moins important de ces conditions pourront tout de même appartenir à la catégorie, mais à un niveau périphérique. Enfin, le prototype peut se manifester pour toutes les catégories, qu'il s'agisse de catégories d'objets du réel (comme oiseau, légumes) ou de catégories d'actions plus abstraites (tuer, etc.). La catégorie ne repose donc pas sur une réalité quelconque, sur une inscription dans le monde naturel, mais sur une perception de l'esprit humain. Le prototype peut se présenter de façon schématique comme suit, le cercle bleu représentant le cœur du prototype, et le cercle plus grand, la périphérie de la catégorie :



2- Cadre théorique, méthodologie employée.

George P. Lakoff s'est également intéressé au problème de la catégorisation et à l'aide que pouvaient y fournir les prototypes. En effet, en tant que cognitiviste, la théorie des prototypes lui apporte une perspective nouvelle sur la façon dont la raison humaine fonctionne. Il a notamment relié la catégorisation en prototypes au mécanisme de la métaphore. Ainsi, selon lui, les métaphores que nous utilisons de façon récurrente structurent réellement notre façon de penser ainsi que nos actes, et, associées aux prototypes, permettent de rendre compte de la façon dont on organise le monde :

« Prototype theory, as it is evolving, is changing our idea of the most fundamental of human capacities, the capacity to categorize, and with it, our idea of what the human mind and human reason are like. Reason, in the West, has long been assumed to be disembodied and abstract, distinct on the one hand from perception and the body and culture, and on the other hand from the mechanisms of imagination, for example, metaphor and mental imagery. » (Lakoff 1987:7)⁵¹

Il remet alors en cause cette séparation de l'imaginaire et de la raison, et utilise la théorie des prototypes dans sa démarche de linguiste cognitif. Selon lui, la théorie classique considère que les êtres humains catégorisent le réel en fonction de propriétés intrinsèques des membres de ces catégories. Il réfute ce point de vue, arguant que c'est plutôt en fonction de la façon dont les êtres humains vont comprendre les éléments du réel qu'ils effectueront leur catégorisation. La catégorisation se fonde donc sur la perception humaine, que cette perception soit physique (visuelle, motrice) ou mentale: mécanismes intellectuels, dont la métaphore, par exemple, (Lakoff 1987:8). En conséquence, l'esprit et le corps n'ont plus à être séparés. Reconsidérer l'acte de catégorisation (particulièrement à l'aide des prototypes) implique également qu'il n'y a plus une bonne façon (la vraie façon) de voir les choses ou de les comprendre, ce que Lakoff appelle « God's eye view » (Lakoff 1987:9), mais au contraire les systèmes conceptuels peuvent grandement évoluer d'une culture à une autre. Ainsi, la signification n'est plus basée sur une vérité, sur des relations entre des symboles et des référents, mais sur la perception qu'en ont les humains.

51 « La théorie des prototypes, telle qu'elle évolue, modifie notre perception d'une des plus fondamentales des capacités humaine, la capacité à catégoriser, et avec elle, notre perception de ce à quoi ressemblent l'esprit et la raison humaine. La raison, dans le monde occidental, a longtemps été considérée comme étant désincarnée et abstraite, distincte d'une part de la perception, du corps et de la culture, et d'autre part, distincte des mécanismes de l'imagination, par exemple, la métaphore et l'imagerie mentale. » (ma traduction).

2- Cadre théorique, méthodologie employée.

Cette théorie a ensuite été synthétisée, notamment pour le domaine de la sémantique par les ouvrages de Taylor, *Linguistic Categorization* (1995) et de Kleiber, *La sémantique du prototype* (1990). Nous allons maintenant voir quelle synthèse ces deux ouvrages proposent, puis nous exposerons les failles de cette théorie du prototype.

II.2.3- La proposition d'une synthèse de la théorie du prototype.

La théorie du prototype affecte directement l'ancien modèle aristotélicien sur lequel s'était basée la sémantique jusqu'à la deuxième moitié du vingtième siècle, comme l'a noté Lakoff (1987). Cette remise en cause a donc naturellement été suivie d'une adaptation, une synthèse de cette théorie pour divers domaines académiques dont notamment la linguistique, et plus particulièrement la sémantique en a été réalisée.

Kleiber, dans son ouvrage *la sémantique du prototype* (1990), note que ce modèle a pris une importance croissante dans la linguistique depuis les années quatre vingt, et il souhaite en conséquence dans son travail clarifier les diverses acceptions qui peuvent être entendues sous cette dénomination unique, ainsi que les diverses améliorations ou applications que ce modèle peut avoir en sémantique. Il liste entre autres avantages de ce modèle celui d'avoir un vaste domaine d'application. En effet, cela peut s'appliquer non seulement à la dénomination des catégories, mais aussi à la grammaire, à la phonologie, et la sémantique. En grammaire, par exemple, la règle ne doit pas être strictement suivie ou non, des graduations peuvent apparaître dans son application. La théorie du prototype a également l'avantage de proposer des catégories moins rigides :

« Elle se révèle, en second lieu, suffisamment souple pour s'adapter aux conditions changeantes de la réalité, pour incorporer de nouvelles données dans les catégories existantes. » (Kleiber 1990:105)

La fixité centrale du prototype assure une stabilité structurale pour permettre la catégorisation. Et c'est alors selon Kleiber cette conjonction entre flexibilité d'adaptation et stabilité structurale qui fait toute l'efficacité du prototype.

Un autre des avantages de cette théorie est qu'elle rend compte du caractère non homogène des catégories. Tous les membres d'une catégorie ne sont pas équivalents, et s'il est aisé d'aller des membres prototypiques vers les membres de la périphérie, l'inverse est

2- Cadre théorique, méthodologie employée.

plus complexe. Cette théorie permet aussi de prendre en compte pour un concept des « propriétés sémantiquement pertinentes mais non nécessaires » (Ibid:109), libérant ainsi les lexicographes du carcan des définitions analytiques. Kleiber note qu'il y aurait selon Wierzbicka deux types de propriétés, certaines essentielles et d'autres prototypiques. Les essentielles ne sont dans ce cas pas nécessaires, pas obligatoires (par exemple, voler pour un oiseau) mais elles ont une pertinence qu'on peut appeler « linguistique ».

Cette théorie permet également une forme d'organisation hiérarchique du lexique. Enfin, Kleiber montre qu'elle permet une réelle avancée en sémantique lexicale à travers diverses découvertes telles que la mise en relief d'une structure intercatégorielle organisée autour d'instances prototypiques, la détermination d'un niveau catégoriel privilégié (le niveau basique), ou encore la vision positive du sens lexical et non contrastive, le sens lexical intervenant davantage notamment dans la construction du sens d'un texte.

Taylor a également travaillé dans *Linguistic categorisation* (1995) sur les apports que cette théorie pouvait amener dans le domaine de la linguistique. Selon lui, cette théorie influence notamment la linguistique au niveau de la sémantique et des équivalences au sein d'une langue ou entre deux langues différentes. Il prend l'exemple des synonymes, on peut parfois croire qu'au sein d'une langue, deux mots soient des synonymes parfaits, mais en fait, après différentes recherches, des prototypes bien distincts peuvent être mis à jour. C'est le cas des mots « vernielen » et « vernietigen » du danois au dix neuvième siècle, Gerraerts a montré que bien qu'étant fréquemment considérés comme des synonymes parfaits, ces deux mots avaient en fait deux prototypes différents (Taylor, 1995:55,56). La même remarque peut se faire concernant le domaine de la traduction, ainsi, on pourrait être tenté de traduire l'anglais « furniture » par l'allemand « möbel », mais après une expérimentation reprenant une partie du protocole d'Eleanor Rosch, il est apparu que les étudiants, selon qu'ils étaient américains ou allemands n'avaient pas fait intervenir les même types de meubles. On peut donc supposer que la traduction est ici inexacte.

Selon Taylor, le prototype peut recouvrir deux réalités, d'une part, il peut s'agir de(s) (l') objet(s) meilleur(s) représentant(s) de la catégorie, et d'autre part, de l'abstraction que forme le cœur de cette catégorie. Il préfère cette deuxième acception, se basant notamment sur des travaux où il traitait des catégories « tallness » et « cowardice » et arguant du fait que trouver un concept prototypique pour ces catégories est pour le moins ardu. Il note encore

2- Cadre théorique, méthodologie employée.

que l'appartenance d'une entité à une catégorie se fait en vertu des similitudes que l'on observe entre cette entité et le prototype. Cela pose des difficultés puisque juger de similitudes entre deux concepts est avant tout une opération subjective à partir du moment où l'on a rejeté les conditions nécessaires et suffisantes en tant que traits distinctifs binaires. Il est donc important de garder à l'esprit que s'il est possible de décomposer un concept ou un mot en attributs plus basiques pour le comparer, il semble impossible d'atteindre des traits primitifs et binaires. Les mots, les concepts sont des objets dynamiques, en perpétuel mouvement, influencés par le comportement de chaque homme de la société, c'est pourquoi leur décomposition en traits fixes, insécables et binaires pose problème.

Taylor note que dans le modèle de Rosch, pour faire partie de la catégorie, un membre doit remplir des attributs qui sont ceux du prototype. Il modère ce point de vue en arguant que le fait de n'avoir pas tous ces attributs n'est pas incompatible avec l'approche prototypique pour un mot ou concept :

« But is not the presence of essential attributes -attributes which are necessarily shared by all members of a category- inconsistent with the prototype approach? The answer, I think, must be no. Attributes are differentially weighted; some might be essential, other can be overridden with varying degrees of facility. Yet the existence of an essential attribute, perhaps even a set of essential attributes, does not in itself lead to all-or-nothing membership in a category. » (Taylor 1995:63)⁵²

Il relève également le fait que coexistent des catégories d'experts et des catégories populaires. Les premières rencontrent assez bien le modèle aristotélicien, tandis que les secondes s'accommodent davantage de la théorie des prototypes. Mais ces différentes façons de catégoriser peuvent coexister sans peine au sein d'une même catégorie. Taylor prend ainsi l'exemple des nombres pairs et impairs, dont la définition technique est simple : un nombre pair est divisible par deux, tandis qu'un nombre impair fournira toujours un reste de un. Cependant, une expérience menée par Armstrong (Armstrong et al. 1983) selon le protocole établi par Eleanor Rosch a montré que les personnes interrogées trouvaient que certains

52 « Mais est-ce que l'absence d'attributs essentiels – attributs qui sont nécessairement partagés par tous les membres de la catégorie -est incohérente avec l'approche prototypique ? Je pense que la réponse doit être non. Les attributs sont plus ou moins importants, certains peuvent être essentiels, d'autres peuvent être mis en position dominante avec différents degrés de facilité. Et même, l'existence d'un attribut essentiel, voire même d'un ensemble d'attributs essentiels, ne mène pas en soi à l'appartenance ou non de façon absolue à une catégorie. » (ma traduction).

2- Cadre théorique, méthodologie employée.

nombres étaient plus pairs que d'autres, même chose pour les nombres impairs. Cela peut s'expliquer par le mécanisme de reconnaissance des nombres pairs et impairs tel qu'il est appris par les masses. Ainsi, le sens, quelle que soit la catégorie qui est concernée, est avant tout une affaire de croyance et de connaissance, et il est possible de donner aux mots des définitions qui seront au choix plutôt populaires ou plutôt pour experts.

Malgré tous ces avantages, la théorie classique aristotélicienne est celle qui a largement prédominé dans l'histoire de la sémantique. Lakoff (1987:5) explique en effet que l'on continue en règle générale à penser que les catégories se définissent en fonction de ce que leurs membres ont en commun, ce qui peut être considéré comme une théorie populaire de la catégorisation. Cette persistance peut s'expliquer par le poids de ce qui a été appris (et qui est directement issu de la tradition aristotélicienne):

« Over the centuries it simply became part of the background assumptions taken for granted in most scholarly disciplines. In fact, until very recently, the classical theory of categories was not even thought of as a theory. It was taught in most disciplines not as an empirical hypothesis but as an unquestionable, definitional truth. »⁵³ (Lakoff 1987:6)

Cette persistance peut être expliquée encore par la prégnance du livre de la Genèse dans les sociétés occidentales, puisque Dieu crée le monde en séparant les entités puis en créant diverses catégories. Le langage, également, selon Wittgenstein (1953) entre en jeu dans ce phénomène de persistance : puisqu'il n'existe qu'un mot pour « game », c'est que les concepts ainsi dénommés ont probablement au moins un point commun. Cependant, Taylor relativise ce point de vue, selon lui, le langage est justement riche de ce qu'il appelle des « hedges » sorte de modalisateurs permettant de donner différents degrés d'appartenance à une catégorie aux concepts dont on parle tels que « à strictement parler », « par excellence », « *grosso modo* » : dire qu'un merle est l'oiseau par excellence semble ainsi beaucoup moins étrange que de dire qu'une dinde est l'oiseau par excellence. Par contre, le langage permet tout à fait de dire que la dinde est, *grosso modo*, un oiseau.

Cependant, si la théorie des CNS demeure aussi utilisée et aussi stable, c'est

53 « Au cours des siècles, (cette théorie) est simplement devenue une affirmation considérée comme garantie dans la plupart des disciplines académiques. En fait, jusqu'à très récemment, la catégorie classique de la catégorisation n'était tout simplement pas envisagée comme une théorie. Elle était enseignée dans la plupart des disciplines non comme une hypothèse empirique, mais comme quelque chose de non discutable, une vérité définie. » (ma traduction).

2- Cadre théorique, méthodologie employée.

également en partie parce que celle des prototypes a pu rencontrer certaines difficultés que nous allons maintenant évoquer plus précisément.

II.2.4- Les failles de la théorie du prototype en sémantique.

Wierzbicka (1996) a nuancé l'idée d'une approche du prototype en sémantique qui permettrait d'oublier le modèle des conditions nécessaires et suffisantes. Selon elle, les deux théories ne s'excluent pas l'une l'autre et si les prototypes peuvent présenter des avantages, la clarté offerte par la pensée aristotélicienne (qui dans ses termes consiste en la recherche d'invariants sémantiques, c'est-à-dire, *grosso modo*, de conditions nécessaires et suffisantes) demeure utile et même vitale dans le domaine de la sémantique (1996:148).

La théorie du prototype fait souvent appel à des exemples tels que le mot anglais « bachelor » pour montrer que le modèle des CNS ne suffit pas pour expliquer ce mot. Wierzbicka montre que sans avoir besoin d'élaborer un prototype, il suffit dans ce cas de compléter, de reformuler la définition qui aurait été faite pour ce mot afin qu'elle soit complète. Selon Wierzbicka, de nombreux abus sont effectués dans l'utilisation de la théorie des prototypes qui n'est pas toujours justifiée (Ibid:160). De la même façon, elle réfute le postulat suite aux travaux de Rosch qui ferait d'une chauve-souris un membre non prototypique de la catégorie oiseau, oiseau est une catégorie taxonomique, en cela des frontières très précises peuvent en être dressées (Ibid:155). Au contraire, la catégorie « meuble » est bien un concept vague, puisqu'il est fondé sur la conscience collective que l'on en a et non sur une existence naturelle, cependant, on ne peut se servir de ce type de concepts flous pour rejeter en bloc le modèle des CNS selon Wierzbicka. Elle réfute particulièrement l'idée d'un degré d'appartenance à une catégorie, peu importe qu'un merle soit plus fréquemment ou plus spontanément associé à la catégorie oiseau qu'une autruche, l'important est que ces deux membres soient des oiseaux.

D'autre part, elle note que selon cette théorie, la signification des mots est considérée comme étant un ensemble flou de composantes sémantiques. Elle refuse ce point de vue en prenant l'exemple du mot « arbre » qui peut se décrire de façon très précise, avec un sens aux frontières bien délimitées (Ibid :168). De plus, si des définitions précises ne peuvent plus être données des mots et concepts, si on ne peut plus tracer une frontière nette entre le sens de tels et tels mots, alors il devient impossible de comparer les significations et

2- Cadre théorique, méthodologie employée.

les valeurs dans différentes cultures (Ibid : 245). C'est pourquoi elle propose de rechercher davantage que de vagues composantes qui constitueraient le sens d'un mot, des invariants sémantiques universels, sortes de concepts basiques se trouvant dans absolument toute culture humaine de façon innée qui puissent rendre compte du sens des mots dans toutes les langues (1996:238). Néanmoins, bien que très intéressante et étayée de nombreuses études empiriques, un des problèmes de la pensée de Wierzbicka demeure dans la complexité de ses définitions, et le fait qu'elles soient relativement obscures d'un premier abord. Elle définit par exemple le mot « games » comme suit :

*« (a) many kinds of things that people do
(b) for some times
(c) for pleasure (i.e. because they want to feel something good)
(d) when people do these things, one can say these things about these people :
(e) they want some things to happen.
(f) if they were not doing these things, they wouldn't want these things to happen
(g) they don't know what will happen
(h) they know what they can do
(i) they know what they cannot do. » (Wierzbicka 1996:159⁵⁴).*

Lakoff a également trouvé une limite principale au modèle des prototypes. Comme nous l'avons déjà vu, il parle de « God's eye view » pour expliquer que selon la culture dans laquelle on se situe, on voit les choses différemment, en fonction de quoi il apparaît problématique de donner des membres de catégorie qui soient plus prototypiques que d'autres dans l'absolu. Ce membre prototypique est purement culturel (Lakoff 1987:446), et il est en conséquence essentiel de situer un prototype dans la culture où il est construit.

Si certaines des limites de la théorie des prototypes telles que le problème de la typicité ou gradation de l'appartenance à une catégorie apparaissent effectivement, il demeure néanmoins possible d'envisager une solution alternative à ce problème, autre que la recherche des primitifs sémantiques de Wierzbicka. Hanks a ainsi proposé une nouvelle

54 « (a) des types de choses que les gens font (b) pendant un certain temps (c) pour le plaisir (c'est-à-dire parce qu'ils veulent se sentir bien) (d) quand les gens font ces choses, on peut dire ces choses de ces gens : (e) ils veulent que des choses arrivent. (f) S'ils ne faisaient pas ces choses, ils ne voudraient pas que ces choses arrivent. (g) ils ne savent pas ce qui va arriver (h) ils savent ce qu'ils peuvent faire (i) ils savent ce qu'ils ne peuvent pas faire. » (ma traduction).

2- Cadre théorique, méthodologie employée.

forme de prototypes, notamment en reprenant ces notions à la lumière de la pensée fondatrice de Wittgenstein, et en y intégrant le concept de coopération conversationnelle gricéenne ainsi que les apports des corpus électroniques modernes.

II.3- Les prototypes selon Patrick Hanks.

Patrick Hanks est un lexicographe qui a dirigé de nombreuses éditions de dictionnaires en anglais, notamment le COBUILD ou des projets au sein de l'Université d'Oxford. Son domaine de recherche l'a amené à s'intéresser plus particulièrement aux analyses lexicales et à la façon d'étudier les sens en usage, en travaillant sur corpus.⁵⁵ Il a à cette fin développé une théorie se basant sur le principe des prototypes en remontant à une autre pensée en étant assez proche, celle de l'air de famille de Wittgenstein. Nous allons voir dans cette partie comment Hanks a, à travers de nombreux articles, élaboré des prototypes sémantiques permettant d'appréhender avec davantage de précision les mots et concepts dotés de multiples acceptions dans l'usage courant, tels le mot « Europe » que nous pourrions ensuite étudier à l'aide des prototypes ainsi définis.

Au fil de divers articles, Patrick Hanks a mené une réflexion au niveau lexicographique sur les moyens de parvenir à une meilleure description du fonctionnement de la langue, et notamment de la signification, du sens dans les dictionnaires. Il a ainsi noté que concernant la langue dans son état naturel, l'imprécision, la redondance et son dynamisme sont trois propriétés principales, et qu'elles devraient davantage être prises en compte dans l'élaboration de dictionnaires ou de théories de la sémantique. La dynamique du langage fait ici appel à la notion gricéenne d'interaction conversationnelle. Concernant le propos de cette thèse, cette imprécision touche aussi de fait l'idée que l'on se fait d'un concept tel que l'Europe dès que l'on en parle puisque pour en parler, on utilise ces entités floues que sont les mots. Une méthode plus en adéquation avec la réalité pour représenter les sens que prennent ces mots permettrait donc de mieux voir ce qu'est le concept « Europe », cette réalité pouvant être approchée notamment via l'usage de grands corpus informatisés. C'est cette méthode, proposée par Hanks que nous allons maintenant détailler en voyant tout d'abord les lacunes qui ont conduit à son élaboration, puis la façon dont un prototype peut se

55 cf. <http://www.patrickhanks.com>

construire, et en en donnant un exemple avec le mot « bank ».

II.3.1- Les problèmes rencontrés en lexicographie.

Les lexicographes s'attachent souvent au problème de l'ambiguïté des sens d'un mot selon Hanks (1994). Il note qu'en fait, le principal problème ne réside pas dans la recherche d'une désambiguïsation parmi différents sens possibles donnés par un dictionnaire, mais plutôt dans une façon plus réaliste de rendre compte du côté flou de la langue. En effet, un mot peut très bien mettre en œuvre plusieurs des sens qui sont listés dans un dictionnaire, mais de façon plus ou moins vague, dans ce cas, il est inutile de chercher à savoir quel sens est appliqué :

« For the vast majority of words in use, including the two words spelled bank, one meaning shades into another, and indeed, the word may be used in a perfectly natural but vague or even contradictory way. In any random corpus based selection of citations, a number of delicate questions will arise that are difficult to resolve or indeed are unresolvable. » (Hanks 2000:206)⁵⁶

De plus, le fait de vouloir présenter un dictionnaire sous la forme d'une liste de significations du mot parmi laquelle il suffirait de choisir pour avoir une interprétation *salva veritate* pose un problème par rapport à la nature dynamique de la langue (Hanks 2000). Cette idée suppose en effet que la langue puisse être décrite sous la forme d'un menu de choix fini, ce qui ne correspond pas à une réalité dynamique de cette même langue. Hanks se réfère à Grice (Hanks 1994:92) pour montrer que la dynamique de la langue a déjà été mise en exergue.

Grice a en fait pensé une théorie de la communication selon laquelle toute conversation ou toute communication se base sur un principe de coopération (Levinson 1983:101). Cette coopération veut que tout locuteur dans une situation de communication tend à faire en sorte que le message soit le plus efficace possible, en fonction de cette volonté, quatre maximes doivent être respectées pour obtenir un message le plus coopératif possible. Il s'agit de la maxime de qualité qui suppose que le locuteur produise une

56 « Pour la grande majorité des mots en usage, incluant les deux mots orthographiés « bank », un sens peut se mélanger à un autre, et en fait, le mot peut être utilisé d'une façon tout à fait naturelle mais aussi vague et contradictoire. Dans toute sélection aléatoire basée sur corpus de citations, nombre de questions délicates ou même impossible à résoudre surgiront. » (ma traduction).

2- Cadre théorique, méthodologie employée.

contribution qui soit une vérité, la maxime de quantité, qui suppose que le locuteur apporte toutes les informations nécessaires et pas davantage, la maxime de pertinence, et la maxime de manière qui requiert d'éviter les propos obscurs ou l'ambiguïté (Ibid :102). Il est évident que ces principes ne sont pas appliqués à la lettre dans toutes les situations de communication. Parfois, la communication peut en effet se faire de façon moins limpide, mais l'idée essentielle de ce principe gricéen de la coopération conversationnelle tient au fait que le récepteur d'un message tend à interpréter le message reçu en fonction de ce principe de coopération, même si le message est éloigné de celui qu'il aurait pu attendre. Dans ce cas, le récepteur produit une inférence que Grice appelle une implicature conversationnelle, c'est-à-dire qu'il interprète le message en fonction de ce qui est conforme au principe de coopération.

Lorsque les maximes ne sont pas observées à la lettre, il y a une forme de floutage, une exploitation -dans des termes gricéens- de la maxime, c'est le cas de l'ironie. Le locuteur dit quelque chose qui rompt la progression de la conversation en défendant superficiellement une thèse qui n'est pas la sienne, le récepteur, appliquant le principe de coopération, et partant de la progression logique de la situation de communication comprend que la thèse ainsi défendue superficiellement est en réalité condamnée, et l'ironie fonctionne (Ibid :109). Grice distingue également les implicatures conventionnelles, qui sont des inférences du récepteur qui ne sont pas conditionnées à une vérité, il prend l'exemple des pronoms « tu » et « vous » en français, les deux peuvent présenter une même personne du monde réel, une même vérité, mais ils expriment néanmoins une différence au niveau de la relation sociale que l'on a vis à vis de la personne représentée par le pronom. De la même façon, les items lexicaux peuvent fréquemment avoir des caractéristiques non conditionnées à une vérité, mais néanmoins conventionnelles dans leur signification, et les règles syntaxiques seront notamment sensibles à ce type d'éléments conventionnels de la signification (Ibid :130).

Plus que le problème de l'ambiguïté entre différents sens à choisir, plus que le fait de déterminer quelle signification parmi la liste fournie est la bonne, un dictionnaire doit viser à décrire au mieux le côté imprécis et dynamique de la langue. Hanks se réfère en cela à Anna Wierzbicka qu'il cite :

« An adequate definition of a vague concept must aim not at precision but at vagueness: it must aim at precisely that level of vagueness which characterizes

2- Cadre théorique, méthodologie employée.

the concept itself » (Hanks 1994:96)⁵⁷

Il ajoute plus loin avec ses propres mots :

“A good prototype will not be so vague as to be meaningless, but will show quite precisely the combinations of conventional syntactic and semantic features that go to make up the conventional usage and meaning potential of a word” (Ibid: 96)⁵⁸

Il peut sembler inconfortable de quitter la forme de certitude que proposent la plupart des dictionnaires (choisissez parmi ce nombre de significations que l'on vous propose pour un mot, et si vous faites le bon choix, vous obtiendrez une interprétation correcte, *salva veritate* dans des termes leibniziens (Hanks 1994:90)). Mais c'est néanmoins ce qu'il convient de faire si on souhaite décrire au mieux une réalité mouvante et imprécise, dynamique, et néanmoins dotée de conventions, telle que la décrit Patrick Hanks en ce qui concerne la langue.

Le dictionnaire a donc pour but de nous renseigner sur la réalité de la langue dont les principales caractéristiques sont selon Hanks (2000:205) l'imprécision, le dynamisme et la redondance. Puisque cette langue est un phénomène qui se prête à la répétition, il est également très important pour le lexicographe de rendre compte de la typicité d'un usage en contexte. Ainsi, le linguiste travaillant sur corpus se doit avant toute chose de poser la question de savoir s'il est normal de prononcer tel énoncé dans telle langue. Pour trouver ce qui est habituel, pour décrire la convention présente dans l'usage des items lexicaux, il apparaît utile de trouver les sens potentiels d'un item lexical, ce qui peut exister selon les conventions généralement recensées, lorsque l'on emploie tel ou tel mot. Nous allons maintenant voir ce que sont ces sens potentiels et comment ils s'insèrent dans la construction de prototypes sémantiques.

II.3.2- Signification et sens potentiel.

Hanks pense ainsi que plus que des significations, ce sont des sens potentiels qui sont mis à disposition dans un dictionnaire ordinaire. Dans des termes gricéens, la

57 « Toute définition adéquate d'un concept vague doit tendre non pas à la précision, mais à l'imprécision. Cela doit tendre précisément au même niveau d'imprécision que celui qui caractérise le concept en lui-même. » (ma traduction).

58 « Un bon prototype ne sera pas imprécis au point d'y devenir obscur, mais il montrera précisément les combinaisons des traits sémantiques et syntaxiques conventionnelles qui construisent l'usage conventionnel et les sens potentiels d'un mot. » (Ma traduction).

2- Cadre théorique, méthodologie employée.

signification est en fait un événement du réel, à partir duquel chaque locuteur va tirer ses propres croyances quant au sens du concept impliqué dans l'acte de communication. Il y a nécessairement parmi cet univers de croyance disparate une part commune, sans quoi la communication serait impossible entre les différents locuteurs, et c'est cette part commune qui forme l'objet des recherches d'un lexicographe (Hanks 1994:92). Cette partie commune peut alors se présenter sous la forme de sens potentiels pour un mot. Il peut y avoir différentes façon d'atteindre ces sens potentiels, le lexicographe peut par exemple se fier à son instinct de locuteur natif d'une langue: puisque je comprends telle chose en général quand j'entends tel mot, je peux en déduire que cela représente un sens potentiel. Mais Hanks souligne que cette méthode pose problème, quand on entend un mot, la façon dont on réagit effectivement, et celle dont on croit que l'on a réagi ne sont pas équivalentes, l'introspection est donc à éviter dans cette recherche. Un des arguments pour cette méfiance est le fait que l'on tend habituellement à se souvenir plus facilement du non familier plutôt que du familier :

« Psychologically, human beings tend to register the unfamiliar rather than the familiar, the unusual rather than the usual. Thus, the files of modern dictionary publishers are full of citations for tachograph and ayatollah – words which have come into prominence within the past decade. [This was written in 1985]. But nobody notices that take has a common but previously unrecorded sense. » (Hanks 1994:92)⁵⁹

Le linguiste, pour rechercher ces sens potentiels a alors la possibilité de se baser sur un corpus, en plus d'indications non biaisées par sa propre expérience linguistique, il disposera ainsi de nombreuses informations sur ce qui est typique ou non dans tel usage, et pourra déterminer pour chaque mot ou concept une liste de différents sens potentiels.

Ces sens potentiels ne sont pas la même chose que les différentes significations habituellement listées dans un dictionnaire, elles ne sont pas considérées comme incompatibles, notamment. Hanks (2000) prend l'exemple de « bank », ce mot est habituellement considéré comme ayant deux sens en anglais : celui d'une bande de terre, et

59 « D'un point de vue psychologique, l'être humain tend à enregistrer ce qui est non familier plutôt que ce qui est familier, l'inhabituel plutôt que l'habituel. Ainsi, les fichiers des éditeurs de dictionnaires récents sont emplis de citations pour ayatollah ou tachygraphe -mots qui ont émergé au cours de la dernière décade. [Ceci a été écrit en 1985]. Mais personne n'a noté que « take » a un sens commun qui n'a jusqu'alors pas été relevé. » (ma traduction)

2- Cadre théorique, méthodologie employée.

celui d'une institution financière qui garde de l'argent en sécurité. Il note qu'en fait, il n'y a d'une part pas un mot avec deux sens principaux, mais bien deux mots qui n'ont en commun que leur graphie. D'autre part, parmi les sens repérés pour l'institution financière, différentes possibilités existent : il peut s'agir d'une banque monétaire, mais aussi d'une banque de données, ou encore d'une banque de sperme. Il peut être question du bâtiment dans lequel la banque se situe : « Cette Banque de France est magnifique ! », et surtout, parfois, plusieurs de ces sens potentiels peuvent être activés en même temps « A quelle heure la banque ferme-t-elle ses portes ? ». Les sens potentiels qui doivent être relevés par le lexicographe ont en conséquence la caractéristique de n'être pas incompatibles.

Ainsi, quand on veut définir le mot « bank » (=entité qui permet de garder quelque chose en sécurité), on liste les différents sens potentiels sous formes de composantes sémantiques de ce mot. Dans la réalité, ces composantes peuvent être activées de façon combinée ou non. Cette approche componentielle prend donc en compte l'imprécision inhérente à un système langagier naturel, et grâce à l'analyse sur corpus, elle peut avoir une force prédictive. De plus, ce type de démarche définitionnelle se rapproche fortement du modèle des prototypes, il y a divers sens potentiels qui sont autant de membres de la catégorie que serait l'objet à définir.

II.3.3- Un exemple de prototype, le cas de « bank ».

Hanks a notamment proposé des études de cas pratiques sur différents mots fréquemment étudiés en lexicographie, tels que le verbe « climb » ou le mot « bank ». Nous allons nous attarder plus particulièrement sur ce deuxième exemple, puisqu'il est fait sur un substantif, ce qui est plus proche de l'objet d'étude de cette thèse qu'un verbe. Les composantes relevées pour créer le prototype ne sont pas plus nécessaires qu'une condition commune ne le serait pour définir un jeu dans la pensée de Wittgenstein. Le prototype présente alors pour le mot « bank » (institution financière) les entités qui peuvent être séparées ou combinées de la façon suivante :

*« BANK
Is an institution
Is a large Building
For storage
For safekeeping
Of finance / money*

2- Cadre théorique, méthodologie employée.

*Carries out transactions
Consist of a staff of people »(Hanks 2000:209)*

Il fournit le même type de prototype pour le deuxième mot s'écrivant « bank ». C'est sur ce modèle, élaboré dans un premier temps à l'aide du matériel lexicographique déjà existant, puis complété ou modifié à l'aide de corpus que nous baserons nos réalisations de prototypes et que nous verrons comment les différents sens potentiels y figurant se trouvent activés dans la presse française. Dans la partie suivante, nous allons en conséquence commencer par proposer un prototype sémantique issu de la matière lexicographique pour le mot « Europe », afin de constituer une base à laquelle les analyses du corpus IntUne viendront se confronter.

II.4- Application du cadre théorique dans la thèse, création d'un prototype lexicographique.

Dans cette thèse, le concept étudié est la représentation de l'Europe à travers une étude de la presse écrite et télévisée française. Il va donc s'agir de voir quel prototype on pourrait construire pour le concept ou mot « Europe » ou « pays » voire pour différents pays de cette entité. Ce type de prototype a déjà été construit à plusieurs reprises (Williams 2008, 2009), en s'appuyant sur les modèles précédemment évoqués, tout en les adaptant au contexte européen. Nous allons ici nous poser la question de savoir comment ces prototypes se construisent.

II.4.1- Le prototype pour mieux appréhender la représentation de l'Europe.

Si comme on l'a vu plus haut, il est compliqué dans le domaine de la lexicographie de définir de façon précise et adéquate la plupart des mots communs, il est d'autres types de mots qui posent davantage de problèmes encore : les noms propres. Ainsi, Williams (2008) note que les dictionnaires qui existent dans le domaine de la toponymie ne donnent le plus souvent que des informations de peu d'utilité dès lors que l'on souhaite avoir des éléments historiques ou culturels sur le mot recherché. En effet, ces dictionnaires se basent davantage sur l'étymologie et l'origine des noms en question. C'est pourquoi l'approche prototypique, telle qu'elle a été élaborée par Hanks semble revêtir un intérêt certain : outre sa capacité à

2- Cadre théorique, méthodologie employée.

rendre compte de la variation au niveau de la signification, elle permet une évolution de la description du mot en fonction de nouvelles recherches, ou de nouvelles preuves rencontrées :

« In looking at place names, it is proposed that the best approach is to adopt the prototype methodology developed by Hanks (2000) to cover the various related features that may be attached to a place name. In this way, the word can be seen in geographical, but also cultural and historical terms. In addition, the name of a country may represent the people of a nation, or an idealisation of that nation » (Williams 2008:399)⁶⁰

Dans sa conception, ce prototype doit alors obéir à différentes règles. Premièrement, il doit se composer d'énoncés simples. Dans le cadre de la création d'ontologies multilingues, Williams suggère en outre que ces énoncés doivent être traduisibles. Ces énoncés reprennent la forme telle que montrée plus haut avec le prototype de « bank » fourni par Hanks.

Ce prototype sémantique se construit par le biais d'une analyse de l'item lexical concerné et s'écrit ainsi en termes de caractéristiques sémantiques basiques. Chacune de ces caractéristiques basiques appartient ensuite, par un jeu de relations thématiques entre différents mots clefs, à un prototype plus large qui permet de mieux rendre compte de la représentation qui est donnée dans la presse du concept complexe étudié. Lorsque l'on parle de caractéristiques sémantiques basiques, il n'est pas question ici des caractéristiques primitives de Wierzbicka (1996) où le but est de réduire toutes les données sémantiques à un ensemble d'universaux primitifs, mais, comme Williams l'explique (2009) leur simplicité se situe davantage à un niveau définitionnel puisqu'elles doivent facilement être compréhensibles par un être humain, ainsi que traduisibles sans ambiguïté.

II.4.2- Un prototype de base.

II.4.2.1- Méthode d'élaboration.

La méthodologie appliquée dans cette thèse suit différentes étapes dans l'élaboration

60 « Concernant les noms de lieux, il est proposé que la meilleure approche à adopter est la méthodologie du prototype développée par Hanks (2000) pour recouvrir les différentes caractéristiques reliées qui peuvent être attachées à un nom de lieu. De cette façon, le mot peut être considéré à un niveau géographique, mais aussi culturel ou historique. En outre, le nom d'un pays peut représenter le peuple d'une nation, ou une idéalisation de cette nation. » (ma traduction.)

2- Cadre théorique, méthodologie employée.

d'un prototype sémantique pour rendre compte des différents concepts étudiés. Dans un premier temps, il s'agit d'élaborer un premier prototype, sur lequel l'étude se basera, et auquel seront confrontées les diverses preuves découvertes à travers l'analyse du corpus. Ce prototype peut notamment s'élaborer avec le concours des différentes ressources lexicographiques selon les caractéristiques évoquées précédemment (suite d'énoncés simples). Ainsi, Williams (2009, Poitiers) utilise dans cette première étape différents dictionnaires qu'ils soient français (Le Larousse, Le Petit Robert) ou anglais (l'Oxford Dictionary of English). Le recours à des ontologies comme WordNet peut aussi être significatif dans la construction de ce premier prototype que l'on va par la suite confronter à la fenêtre sur la réalité captée dans le corpus.

Au sein de ce premier prototype, on peut déjà distinguer un sens, un énoncé plus prédominant que les autres, il sera par exemple situé en premier dans l'entrée du dictionnaire concernant le mot étudié. On peut alors considérer, tout comme Eleanor Rosch que ce centre est en quelque sorte prototypique, en ce qu'il est au cœur du concept, de la catégorie étudiée. Cela peut aussi se retrouver au sein du corpus à travers tel sens plus dominant que les autres en termes de fréquence, notamment. Néanmoins, il est essentiel d'avoir à l'esprit qu'une différence de taille subsiste quant à l'approche roschienne. En effet, si on peut accepter une signification cœur au sein d'un corpus, d'un sous corpus ou d'un dictionnaire donné, aucun sens seul ne peut être considéré comme prototypique puisque comme pour le « jeu » de Wittgenstein ou le « bank » de Hanks, dès que l'on parle d'une entité comme « Europe », une multitude de sens interviennent de façon simultanée.

II.4.2.2- Un prototype lexicographique pour le mot « Europe ».

Dans le cadre d'une publication du groupe IntUne, « Nation and Supernation : a tale of three Europes » (Williams et al. 2009), un prototype a déjà été élaboré pour ce mot, se basant sur de multiples ressources lexicographiques aussi diverses que le Larousse 2005, l'Oxford Dictionary of English 2005, ou le Zingarelli de la même année. Ce prototype se base sur la méthodologie énoncée par Hanks décrite dans le chapitre 3 de cette thèse. L'utilisation de dictionnaires dans différentes langues et faisant appel à des techniques de construction différentes permet de créer un prototype qui se constitue de différentes composantes, de différents sens potentiels tels que décrits par Hanks (2000). En effet, le fait

2- Cadre théorique, méthodologie employée.

d'avoir des ressources dans des langues différentes permet de créer des énoncés simples qui soient aisément traduisibles. D'autre part, le fait d'avoir des dictionnaires qui aient été construits à l'aide de corpus ou au seul moyen de l'introspection permet de couvrir un vaste spectre de significations possibles, qui seront une base solide à l'étude sur corpus.

Ce prototype fait apparaître les composantes suivantes pour le mot « Europe » :

- *« Une masse terrestre.*
- *Situé dans l'hémisphère Nord.*
- *Entouré d'îles plus ou moins grandes.*
- *Une péninsule d'Eurafrasie.*
- *Une péninsule d'Eurasie.*
- *Un continent.*
- *Situé entre l'Asie et l'océan Atlantique.*
- *Délimité à l'ouest par l'océan Atlantique.*
- *Délimité au Sud par la méditerranée.*
- *Délimité au Sud par la chaîne du Caucase.*
- *Délimité à l'Est par la mer Caspienne et l'Oural.*
- *Composé de plusieurs pays.*
- *Dont les pays sont regroupés dans des blocs géographiques.*
- *Dont les pays sont regroupés dans des blocs politiques.*
- *Partiellement composé d'une unité politique : l'Union Européenne.*
- *L'Union Européenne est divisée en sous-ensembles.*
- *L'Europe politique est une idéalisation.*
- *Composé de pays continentaux par rapport à la Grande Bretagne. »*

(Williams et al. 2009)

Ce prototype étant posé, nous pourrons étudier dans le quatrième chapitre de cette thèse, concernant les analyses menées sur corpus sur le mot « Europe », quelles en sont les activations en 2007 et 2009 à travers les lignes de concordances en format KWIC du mot-clef « Europe », et s'il convient de le modifier en fonction des résultats fournis par l'analyse du corpus IntUne.

II.5- Conclusion.

A travers ce chapitre, nous avons pu appréhender la complexité du travail de définition de concepts qu'ils soient relativement simples et présents comme entités naturelles dans notre monde, ou qu'ils soient totalement construits par l'humain. L'Europe, en tant que

2- Cadre théorique, méthodologie employée.

construction géographique, politique, culturelle, etc. appartient à cette seconde catégorie, la définir représente en conséquence une gageure pourtant nécessaire si l'on souhaite pouvoir comparer la façon dont elle est comprise dans ses différents pays membres, par exemple (Wierzbicka 1996:237).

Le modèle des prototypes tel qu'il a été adapté par Hanks dans le domaine de la lexicographie (Hanks 1999, 2000) semble fournir une aide utile à la définition de ce concept. Sa flexibilité permet de rendre compte de la réalité mouvante et dynamique de la langue, et ce modèle rend également possible une comparaison à travers différentes langues (Williams et al. 2009). Nous avons pu élaborer un premier prototype sémantique pour le mot « Europe » à partir de la matière lexicographique déjà existante sur ce concept, la deuxième étape de notre analyse dans le quatrième chapitre nous conduira à vérifier l'activation de ce prototype dans la presse française à travers les quatre journaux de notre corpus, sur les deux périodes couvertes en 2007 et 2009.

Mais avant cela, nous allons expliciter précisément le cadre technique de cette thèse, la façon dont le corpus IntUne a réellement été construit, ainsi que ses principales caractéristiques. Nous allons également préciser les types de logiciels qui ont été employés pour interroger ce corpus ainsi que les manipulations qui ont été effectuées sur les corpus IntUne 2007 et 2009 à l'aide de ces logiciels. Nous expliquerons enfin comment les prototypes lexicographiques élaborés dans un premier temps ont été confrontés aux données du corpus, notamment via l'utilisation de calculs de collocations à l'aide du logiciel Xaira ou via la classification de lignes de concordance avec WordSmith.

3- Technique de construction et d'analyse du corpus IntUne.

III.0- Introduction.

Dans le chapitre précédent, nous avons dressé un cadre théorique et méthodologique dans lequel cette thèse s'inscrit en insistant sur l'importance de la tradition aristotélicienne dans le domaine lexicographique, mais également sur ses lacunes et sur les solutions qui y ont été apportées. L'approche retenue dans notre étude est le prototype sémantique tel que Hanks l'a appliqué. Néanmoins, si nous avons vu qu'il était possible de réaliser des prototypes sémantiques à partir d'une matière lexicographique déjà existante, il est essentiel de les confronter à l'usage réel de la langue, et dans le cadre de notre thèse, de les confronter à l'usage du mot étudié tel qu'il apparaît dans la presse française. Cette confrontation peut avoir lieu grâce à l'étude du corpus IntUne dont nous avons rapidement parlé lors de l'introduction. Mais avant de procéder aux analyses sur corpus, il convient de détailler plus précisément le corpus IntUne, son élaboration en plusieurs étapes, sa structuration, sa composition, ainsi que les techniques d'analyses qui ont été menées sur ce corpus. Ce sera l'objet de ce chapitre.

III.1- L'élaboration du corpus IntUne

Nous avons précisé dans l'introduction de cette thèse quel était le projet IntUne, et

3- Technique de construction et d'analyse du corpus IntUne.

comment plus précisément la tâche du groupe média se définissait à l'origine de ce projet (à travers la réalisation d'un corpus notamment). Nous allons maintenant voir comment la réalisation de ce projet s'est déroulée, quels choix principaux ont été effectués et quelles difficultés ont été rencontrées dans la conception du corpus. Dans un premier temps, nous nous pencherons sur la création d'un corpus pilote selon les choix effectués à l'origine du projet par le groupe média. Nous étudierons ensuite dans le détail la réalisation du corpus IntUne, les difficultés soulevées au cours de ce travail ainsi que la façon dont le corpus a été structuré. Enfin, nous donnerons une description plus générale du corpus IntUne en quelques chiffres.

III.1.1- Le corpus pilote.

La première tâche du groupe média a consisté en la création d'un corpus pilote dans le but de créer une méthodologie pour la conception du corpus IntUne. La coordination entre les différentes universités a également ainsi pu être améliorée pour davantage d'efficacité. Ce corpus pilote a été collecté du 21 au 26 novembre 2006 et il a été décidé par les membres du groupe média qu'il devait contenir des textes issus de presse écrite et des textes issus de journaux télévisés. Les textes sélectionnés pour la presse écrite étaient issus de deux journaux nationaux avec une orientation politique de droite et de gauche, de deux régionaux, et enfin d'un national spécialisé dans le domaine économique. Concernant les journaux télévisés, la collecte devait porter sur les principales éditions d'une chaîne publique et d'une chaîne privée. Tous les textes des journaux de presse écrite et des journaux télévisés ont été inclus dans le corpus, et non pas seulement les articles qui auraient concernés l'Union Européenne (Dugalès 2008). Le premier corpus de référence a été collecté entre février et avril 2007, tandis qu'un second corpus de référence, se basant sur le même type de source et les mêmes procédures de collecte, a été rassemblé au début de l'année 2009.

Nous allons maintenant voir comment le corpus IntUne a réellement été conçu, quelles méthodes pratiques ont été mises en place lors de sa collecte, quelles questions ont motivé les différents choix effectués par les quatre universités participant à ce groupe, etc. Les difficultés posées par l'élaboration d'un corpus comparable seront également abordées.

III.1.2- L'élaboration du corpus IntUne.

Le corpus IntUne se compose de textes issus de deux genres de médias : la presse écrite et la presse télévisée. Cela exclut de fait toute la presse radiophonique dont l'importance ne peut cependant être minorée. Une exploitation d'un corpus de ce dernier type de presse aurait été des plus intéressants (il suffit pour cela de voir les fréquentes mentions d'émissions politiques ou d'informations dans les textes du corpus IntUne), mais il aurait également été très coûteux à mettre en place. En effet, la constitution de corpus oraux demande beaucoup de main d'œuvre comme nous le verrons dans la seconde sous partie de ce paragraphe. Nous allons maintenant voir comment les corpus de presse écrite et de journaux télévisés ont été collectés, les choix qui ont été nécessaires à leur élaboration, ainsi que les difficultés posées dans la décision de ces choix. Nous verrons ensuite de quelle façon ils ont été structurés.

III.1.2.1- Le corpus de presse écrite.

Ce corpus se compose de deux ensembles, un premier corpus collecté en 2007, et un second collecté en 2009 selon les mêmes modes opératoires.

Journal retenu	France	Grande-Bretagne	Italie	Pologne
National conservateur	Le Figaro	The Telegraph (The Sun*)	Il Corriere della Serra	Nasz Dziennik
National Libéral	Libération	The Guardian (The Mirror*)	La Repubblica	Gazeta Woborcza
Régional conservateur	Sud Ouest*	The Western Mail	Il Giornale di Brescia	Kurier Szczecinski*
Régional Libéral	Ouest France*	The Scotsman	La gazetta del sud	Kurier Lubleski*

(Figure 1, journaux choisis par le groupe média pour le corpus presse.)

Les journaux entre parenthèse pour la colonne Grande Bretagne avaient été retenus lors du corpus pilote, puis retirés du corpus car n'ayant pas leur équivalent dans les autres pays de collecte du groupe média. Les journaux régionaux français n'ont pas d'orientation politique clairement spécifiée dans le corpus, puisque comme nous allons le voir ces journaux ont été choisis davantage en fonction de critères de coûts qu'en fonction de leur

3- Technique de construction et d'analyse du corpus IntUne.

orientation politique. Les journaux régionaux polonais présentent également une absence d'orientation politique clairement définie.

Ce corpus se veut un corpus comparable à travers la sélection de textes selon des critères précis, décidés lors de réunions entre les différents membres du groupe média IntUne. Néanmoins, tout corpus comparable pose le problème de savoir ce qui est réellement comparable. Dans la mesure, par exemple, où le statut de l'Écosse et celui du Sud Ouest sont sensiblement différents au sein de leurs États respectifs (Royaume-Uni et France), peut-on dire qu'un journal diffusé en Écosse (The Scotsman) est un journal régional, au même titre qu'un journal diffusé dans le Sud Ouest (Sud Ouest) ? L'Écosse est-elle une simple région, alors que le Royaume-Uni lui reconnaît le statut de nation constituante, ainsi que l'ONU, qu'elle dispose d'un parlement et d'un premier ministre (Alex Salmond). Peut-elle être considérée au même titre que le Sud Ouest en France, qui certes dispose d'une identité forte avec le pays basque, mais sans que celle-ci soit reconnue de façon institutionnelle ? Ces questions posent la question de la comparativité réelle du corpus.

D'autre part, l'établissement de critères permettant d'obtenir un corpus avec une structure la plus proche possible d'un pays à l'autre conduit à exclure certains médias des textes sélectionnés et donc certaines particularités nationales pourtant très significatives. C'est le cas notamment des tabloïds en Grande Bretagne. A titre d'exemple, le Mirror tire à environ 1 700 000 exemplaires, le Sun tire à environ 3 200 000 exemplaires, tandis que le Figaro en France tire à 360 000 exemplaires et Ouest France culmine avec 800 000 tirages⁶¹). Bien que très lus et traitant de sujets extrêmement divers au Royaume Uni, ce type de presse n'a pas pu entrer dans le corpus IntUne puisque cela n'existe ni en Italie, ni en France, ni en Pologne. Il est regrettable de se priver de l'analyse de ce type de journaux, mais pour une meilleure comparabilité du corpus, ce type de choix est nécessaire. De même, si en France, en Pologne et au Royaume Uni, il existe des journaux de spécialité financière largement consultés par les élites comme The Financial Time, Puls Biznesu ou Les Échos, ceux-ci n'ont pas leur équivalent en Italie par conséquent, le corpus IntUne a rejeté l'utilisation de ce type de journal. L'équipe de Lorient a néanmoins collecté les articles issus des Échos sur la période concernée et a ainsi constitué un corpus satellite afin de mener certaines études comparatives entre les quatre journaux faisant partie du corpus et les Échos pour étudier les différences de

61 [Http://www.observatoire-media.info/](http://www.observatoire-media.info/)

3- Technique de construction et d'analyse du corpus IntUne.

traitement de certains thèmes selon le lectorat principal de ces journaux.

Les articles des journaux choisis pour faire partie du corpus ont ensuite été collectés grâce à une base de données sur Internet, factiva⁶² sous un format .txt. Lors de la réalisation du corpus pilote, la base de données utilisée était Lexis Nexis, mais l'Université de Bretagne Sud a préféré les services de factiva pour une raison de coûts, ce qui a conduit au remplacement du régional « Le Télégramme » par « Ouest France ». Cela a posé certains problèmes, notamment quant à l'orientation politique de ces deux régionaux, sensiblement différente, alors que Le Télégramme avait été préalablement choisi notamment en fonction de son orientation politique conservatrice. Cette modification du support a aussi demandé au chercheur s'occupant de la conception du corpus presse de revoir toutes les procédures de travail mises au point lors de la création du corpus pilote. Ensuite, il s'est agi de vérifier à l'aide de la version papier du quotidien (qui est celle qui est la plus diffusée auprès des lecteurs, donc celle sur laquelle le corpus se basait, puisque, étant davantage lue, elle impacte davantage l'opinion publique) qu'aucun article n'était présent deux fois ou bien absent. Le texte était alors transformé d'un format .txt en un .xml requis pour pouvoir travailler avec une architecture de base TEI. Enfin, des modifications successives et répétées aux en-têtes et à certaines balises, concertées avec les quatre pays du groupe média ont été apportées afin de parvenir à un corpus comparable avec les autres pays du groupe média (Dugalès 2008).

Le corpus français a été de plus lemmatisé à l'aide de treetagger⁶³. La lemmatisation présente de nombreux avantages dans les études et analyses menées sur le corpus (Leech 2004). Ainsi, il est plus facile d'effectuer des recherches d'un même lemme (le lemme correspond à une entrée de dictionnaire), par exemple du mot Europe, quelque soit les formes graphiques prises par ce mot dans le corpus. Une fois le corpus lemmatisé, il devient possible de trouver toutes les occurrences du mot « Europe » en une seule requête, alors qu'auparavant, il était nécessaire d'effectuer des recherches supplémentaires telles que « d'Europe », « L'Europe ». Il est également plus aisé de mener des recherches affinées en fonction des catégories de mots, la lemmatisation permet par exemple de rechercher tous les adjectifs suivant le mot Europe dans le corpus.

62 Site recueillant différents ensembles d'information et le revendant dans des formats txt : <http://www.factiva.fr/fr>

63 Treetagger est un étiqueteur probabiliste qui n'est pas dédié à une langue particulière. Il se compose en effet d'un programme principal (fixe) et de fichiers de paramètres qui dépendent de la langue à tagger.

3- Technique de construction et d'analyse du corpus IntUne.

Le corpus 2009 a été construit selon le même mode opératoire. Les deux corpus ont été collectés du 05/02/2007 au 06/05/2007 pour le premier et du 02/02/2009 au 02/05/2009 pour le second. Ces deux périodes correspondent à des événements dont la présence est très marquée dans la presse française : la campagne pour l'élection présidentielle en 2007 et la crise économique internationale, ainsi que la campagne pour les élections législatives européennes en 2009.

III.1.2.2- Le corpus télévisé.

Après avoir eu différents échanges avec tous les pays membres du projet, il a été décidé de collecter tous les journaux télévisés nationaux sur une période allant du 12 février au 4 avril pour le corpus 2007, et du 09 février au 31 mars pour le corpus 2009. Là encore, ces périodes correspondent aux études menées par les autres groupes du projet IntUne et cela permet de voir si le discours médiatique influence ou non la perception des thèmes traités pour les élites, les experts et les masses. En outre, elles correspondent à la campagne électorale française pour le premier corpus, et à la crise économique internationale ainsi que la campagne pour les élections législatives européennes pour le second corpus comme nous l'avons déjà vu. Seuls les journaux diffusés en semaine ont été transcrits (hors samedi et dimanche). Lors de réunions du groupe média, les journaux du weekend ont ainsi été exclus, d'une part parce que la transcription du corpus oral était longue et coûteuse, et d'autre part parce qu'il a été jugé que les informations données lors des weekend contenaient pour la plupart des résultats sportifs et n'étaient pas des plus pertinentes dans le cadre de l'étude menée par le groupe média. Ce manque des journaux du weekend est néanmoins regrettable dans la mesure où l'émergence d'une identité au niveau local se fait souvent via l'identification à telle ou telle équipe de sport. Les journaux télévisés choisis par le groupe média peuvent se présenter comme suit :

Chaîne du JT	France	Grande Bretagne	Italie	Pologne
Chaîne d'état	France 3	BBC	TG1	TVP1
Chaîne privée	TF1	ITV	TG5	TVPolsat

(Figure 2, journaux télévisés choisis par le groupe média)

Le groupe média a choisi comme matériau télévisé l'édition principale du soir d'une

3- Technique de construction et d'analyse du corpus IntUne.

chaîne publique et d'une chaîne privée, selon le postulat que le fait d'être dépendant ou non de ressources privées pouvait influencer la ligne éditoriale et donc le contenu des informations qui sont diffusées sur ces chaînes. D'autre part, le choix de l'édition du soir est du au fait que c'est l'édition qui a la plus grande audience et qui impacte donc le plus l'opinion publique⁶⁴. Le choix d'une chaîne publique et d'une chaîne privée peut poser question, en effet, on peut se demander si une chaîne publique est nécessairement indépendante en comparaison d'une chaîne privée qui serait vue comme dépendante. A partir du moment où l'audience augmente, une chaîne, privée ou publique, pourra vendre plus cher ses espaces publicitaires, hors si le public estime qu'une information manque trop d'indépendance, il peut bouder la chaîne, une chaîne privée semble donc avoir tout intérêt à conserver une indépendance maximum. D'autre part, au niveau politique, le fait de dépendre de l'État peut inclure pour une chaîne publique une dépendance plus ou moins importante vis à vis de la mouvance politique dirigeante. En France, les journaux télévisés collectés ont été ceux de TF1 et de France 3. Tous les journaux du corpus 2007 ont été enregistrés sur support DVD, après que ce choix ait été fait par le groupe média car il semblait le plus satisfaisant d'un point de vue technique. En 2009, les journaux ont été acquis à partir d'une ligne de commande sous un environnement linux à partir du site internet des chaînes télévisées, puis stockés sur un simple disque dur externe, ce qui semblait à la fois plus fiable et moins coûteux en terme de temps et d'argent.

Un premier visionnage permettait de récupérer les informations nécessaires dans le corpus pour les différents locuteurs intervenant dans le journal, leurs noms, sexes, degré d'expertise sur le sujet abordé, métier, etc. Ces informations étaient prises en note, ainsi que les conditions de tournage du reportage (voix off, journaliste parlant face à la caméra) pour pouvoir être par la suite transcrites suivant une série de codes, entrés à l'aide de balises conformes aux normes de la TEI, harmonisées parmi les quatre pays de constitution du corpus. Nous reviendrons dans la sous-partie suivante sur les différentes balises utilisées dans ce corpus. A l'aide du logiciel libre de droit, audacity⁶⁵, le son de ces journaux a pu être capturé, puis la transcription s'effectuait avec le logiciel Transcriber⁶⁶.

64 <http://www.mediametrie.fr/> le journal télévisé rencontre une plus forte audience le soir que le midi, ce phénomène est d'autant plus accentué sur TF1 où l'on peut noter une différence de treize point entre l'audience moyenne de la journée, et celle du JT du soir, très regardé.

65 <http://audacity.sourceforge.net/> logiciel fonctionnant comme une sorte de magnétophone pour ordinateur.

66 <http://trans.sourceforge.net/en/presentation.php>

3- Technique de construction et d'analyse du corpus IntUne.

Ce logiciel a été une aide précieuse pour la transcription du matériau oral du corpus puisqu'il permet une saisie plus facile du texte, mais également de structurer le texte en trois divisions (sur lesquelles nous reviendrons plus bas), et enfin d'intégrer différentes informations concernant les différents locuteurs. Ce logiciel se présente sous la forme d'un traitement de texte qui peut se faire en temps réel avec le défilement d'une bande son (cf. annexe 5a). La saisie du texte oral défilant peut se faire beaucoup plus rapidement, notamment au moyen de différents raccourcis claviers permettant de mettre le fichier son en pause, ou de créer un tour de parole. Ce logiciel permet aussi d'attribuer des locuteurs aux tours de paroles créés (cf. annexe 5b) et de renseigner diverses informations sur ces locuteurs. La division du texte oral transcrit en différentes parties, respectant les différents moments pouvant être présents dans un journal télévisé, était également rendue plus aisée par ce logiciel qui permet comme nous l'avons déjà dit de structurer le texte en trois parties, les introductions de reportages, reportages et interviews pouvaient ainsi être créés dans la structure du fichier écrit dès sa transcription.

Une feuille de style de type xslt a permis de rendre les fichiers xml obtenus après transcription totalement compatibles avec les normes de la TEI (cf. annexe 6). Une relecture a permis d'ajouter aux textes .xml ainsi obtenus différentes balises afin de les structurer au mieux et d'avoir une plus grande harmonie entre les quatre corpus du groupe média. Les erreurs subsistant ont ainsi pu également être corrigées.

III.1.2.3- Structuration du corpus IntUne.

Le corpus est parfois décrit métaphoriquement selon l'image d'un « sac de mots » (Rastier 2005a), montrant qu'il s'agit certes d'une collection textuelle importante, mais dont la structure peut échapper au premier abord. Une des premières préconisations de Sinclair (1991) dans sa définition du corpus était d'en faire une entité structurée. C'est également le cas chez Rastier (*Enjeux épistémologiques de la linguistique de corpus*, 2005) :

« Un corpus est un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages et rassemblés : 1) de manière théorique réflexive en tenant compte des discours et des genres et 2) de manière pratique en vue d'une gamme d'applications. » (Rastier, 2005a : 32)

3- Technique de construction et d'analyse du corpus IntUne.

Le corpus IntUne 2007 a donc une structure lui permettant d'être interrogé et analysé de façon fine. Cette structure est basée sur un format .xml et suit les normes mises en places par la TEI (Text Encoding Initiative). La TEI est une réflexion née d'un besoin de numériser des textes sous un format qui soit le plus accessible possible, et qui donne une représentation riche d'un texte. Le balisage permet en effet de prendre en compte des informations quant à la source du texte, par exemple, ces méta-données permettent un meilleur archivage ainsi qu'un meilleur accès au texte, elles sont généralement stockées dans son en-tête. Des informations quant à la structure interne du texte peuvent être présentes dans le corps de ce texte à travers l'usage de balises entre symboles `<>` (début de l'action de la balise) et `</>` (fin de l'action de la balise). Le choix de la TEI rencontre plusieurs avantages. D'une part, ce format est totalement gratuit, de plus, il est déjà utilisé par une nombreuse communauté, notamment dans le domaine de la linguistique de corpus (le BNC dont nous avons parlé plus haut a été annoté selon des normes conformes à la TEI), et son usage dans le monde de la recherche est très répandu. Le groupe média IntUne a donc choisi d'utiliser cette norme pour coder et baliser les textes collectés au mieux, en s'efforçant de retranscrire avec le plus de précision possible les structures telles qu'elle existent dans ces différents textes. La TEI est par ailleurs un ensemble de normes fonctionnant au format .xml (Xtensible Markup Language) largement répandu⁶⁷.

Un exemple issu du corpus presse et du corpus télévisé est fourni dans les annexes 3 et 4. On y voit comment les différentes balises fonctionnent. Pour le corpus presse, il y a tout d'abord un en-tête, ou « tei header », dans cet en-tête, on retrouve diverses informations sur le texte, son titre, sa provenance, l'entité qui a réalisé sa transcription (Lorient IntUne Media Group), le corpus dans lequel se situe le texte (ici, le corpus IntUne pilote), on peut également y relever la date de publication du texte transcrit et des informations concernant sa diffusion, son orientation politique, etc. Toutes ces informations permettent non seulement de correctement archiver un texte, mais aussi de pouvoir l'interroger de façon plus efficace lorsque l'on mène une étude sur l'ensemble du corpus : le fait d'avoir dans cet en-tête les informations de diffusion ou d'orientation politique, notamment, permettent de n'interroger que les textes ayant telles ou telles orientations une fois le corpus constitué. La deuxième partie concerne le texte, elle se situe entre les balises `<text>` et `</text>`, les différentes

67 Pour plus d'information sur cet ensemble de normes pour le codage de textes numériques, consulter le site : <http://www.tei-c.org>

3- Technique de construction et d'analyse du corpus IntUne.

divisions de ce texte permettent de situer le texte transcrit dans l'organisation du journal, par exemple, la div2 permet ici de savoir que l'on se situe dans la une du journal Le Figaro, la division div3 permet de savoir que la rubrique retranscrite est ici l'éditorial du journal, les titres et sous titres sont également présentés en étant encadrés de balises, le corps du texte est encadré de balises <p>. Ce système de balises permet ainsi non seulement d'enregistrer les méta-données du texte transcrit, mais également de rendre sa structure, son organisation intelligibles pour ses lecteurs qu'ils soient informatiques ou humains.

Le corpus télévisé fonctionne selon le même type de principe (annexe 4). Il comporte tout d'abord un en-tête doté des mêmes types de méta-données, s'ajoute dans ces méta-données une liste des locuteurs intervenant dans ce texte, ainsi que leurs qualités. Le corps de texte se sépare lui aussi en trois divisions, la première division (div1) correspond à une édition de journal télévisé, la deuxième division (div2) correspond à un sujet traité dans le journal, et la troisième division (div3) correspond aux différents temps qui peuvent être présents dans un journal télévisé : titres, introduction d'un reportage par le présentateur, reportage, interview, reprise des titres à la fin du journal. L'extrait présenté ici est issu d'une des premières versions du corpus 2007, qui a par la suite été enrichi d'autres balises. Les balises dont ce corpus s'est enrichi à la suite de rencontres du groupe média sont notamment une balise <voiceover> ou <camera> permettant de savoir si le texte est produit en voix off ou en face de la caméra. Une balise <s> a également été ajoutée, permettant de découper le texte en phrases, ces choix réalisés à la suite de choix du groupe média ont permis un enrichissement des annotations et une plus grande harmonisation du corpus.

L'annexe 6 présente une des feuilles de style utilisées pour passer du fichier obtenu après une transcription sur transcriber au fichier xml conforme aux normes TEI et semblable dans sa structure aux autres corpus télévisés du groupe média. Son fonctionnement peut être explicité à travers les deux lignes de codes surlignées en jaune dans l'annexe. Ces deux lignes signifient que si dans le texte issu de transcriber, le programme mis en place dans cette fiche xslt trouvait une balise « male », alors, il fallait lui donner l'attribut sex=1, ce qui était conforme aux choix réalisés par l'ensemble du groupe média. La feuille xslt fonctionne selon ce principe de correspondances et a permis de réaliser l'essentiel des transformations nécessaires à l'harmonisation de façon automatique. Néanmoins, il a fallu par la suite relire le corpus manuellement pour amener les dernières corrections.

3- Technique de construction et d'analyse du corpus IntUne.

Cette structuration et les révisions, corrections multiples ont consommé une grande quantité de temps, mais cela s'est avéré nécessaire pour pouvoir mener des études et analyses fines du corpus à l'aide de logiciels d'extraction de texte compatibles avec les normes de la TEI, notamment XAIRA que nous avons déjà mentionné. D'autre part, ce travail a permis d'aboutir à un corpus se composant de textes de quatre langues, structuré de façon homogène, et dans la mesure où il permet de réaliser des études comparatives via Xaira, nous pouvons le qualifier de corpus comparable. Un autre logiciel d'extraction de texte est également employé au sein du groupe média, il s'agit de Wordsmith, ce logiciel édité par Mike Scott permet d'éditer des listes de fréquences issues du corpus, des concordances, et de rechercher les mots clefs de certains textes. Il ne requiert d'autre part aucune indexation du corpus utilisé, contrairement à Xaira. Cela rend son utilisation plus facile pour le novice, mais peut demander davantage de temps lors de recherches de concordances ou de collocations.

Xaira est un logiciel élaboré au sein de l'université d'Oxford afin de fonctionner avec le BNC (Burnard 2007). Il offre, outre les fonctions habituelles d'un concordancier, la possibilité de faire une recherche sur une partie seulement du corpus, puisqu'il prend en compte la structure TEI. Par exemple, il est possible de faire une recherche sur un phénomène linguistique au sein des lignes de titres des journaux télévisés uniquement. De cette façon, outre l'analyse quantitative habituellement menée dans le cadre de la linguistique de corpus, il est possible de procéder également à des études qualitatives, prenant appui sur des requêtes plus fines au sein du corpus. Ce logiciel permet également d'avoir des données purement statistiques, dont nous allons donner quelques exemples dans la section suivante.

III.1.3- Le corpus IntUne en quelques chiffres.

Voici quelques données statistiques nous renseignant sur la composition du corpus IntUne obtenues à partir du logiciel Xaira. Le tableau suivant nous renseigne sur le nombre de mots présents dans le corpus IntUne français 2007, et sur la proportion de chaque source dans le corpus. On y voit que la source qui compte le plus grand nombre de mots est Les Échos, suivi par le Figaro, Sud-Ouest, Libération et Ouest France. Cette répartition se retrouve la plupart du temps lorsqu'on recherche un mot clef particulier dans le corpus. Elle se retrouve également dans le corpus IntUne 2009. Le corpus télévisé occupe

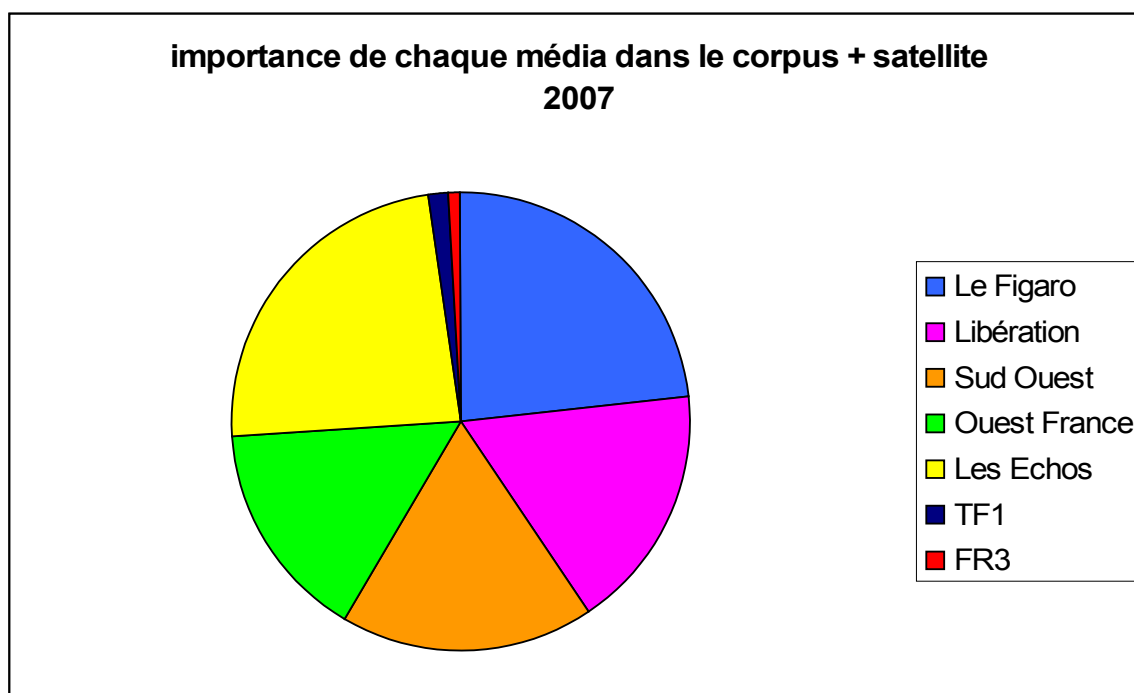
3- Technique de construction et d'analyse du corpus IntUne.

proportionnellement une place moindre, on s'aperçoit que les journaux de TF1 représentent beaucoup plus de mots que ceux de FR3 pourtant plus nombreux dans le corpus. Le corpus français IntUne 2007 comprend environ 20 millions de mots avec Les Échos, et 15 millions sans. En 2009, il comprend environ 18 millions de mots avec les Échos et 14 millions sans, le corpus télévisé n'a pas encore été transcrit pour le corpus 2009.

source	Nombre de textes	mots
Le Figaro	77	4861988
Libération	77	3602857
Les Échos	63	5015207
Sud-Ouest	77	3683377
Ouest France	77	3226744
TF1	32	251551
France 3	34	183741
Total	437	20825465

(Figure 3, répartition des mots du corpus IntUne 2007 français par média)

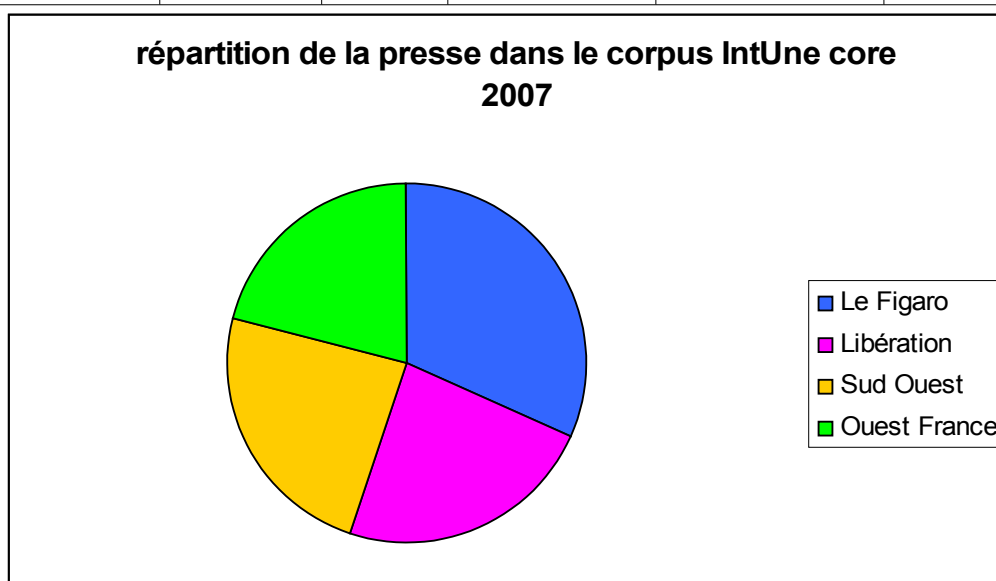
Le graphique suivant reprend cette répartition en fonction des médias dans le corpus IntUne français en 2007.



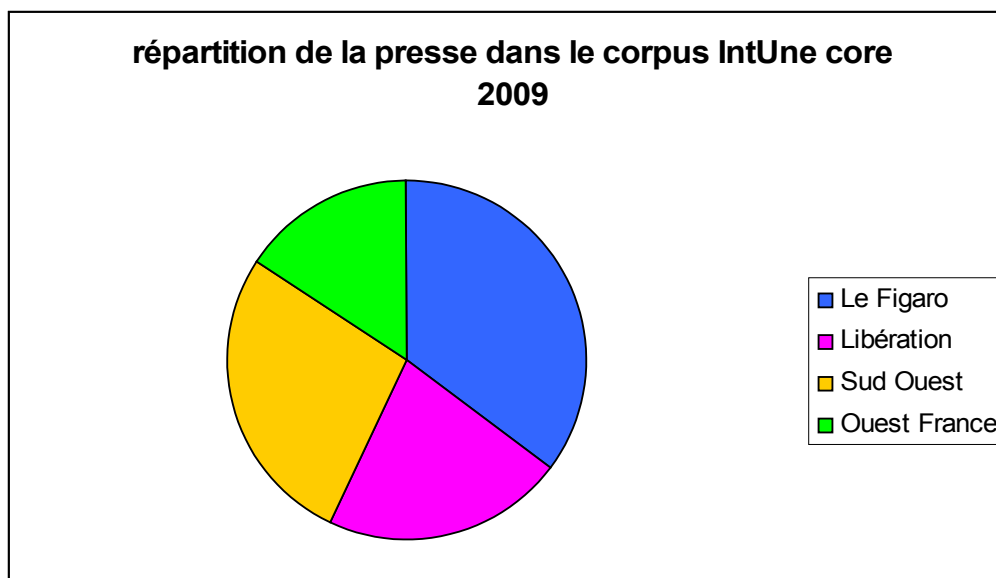
3- Technique de construction et d'analyse du corpus IntUne.

Ces données concernant le corpus 2007 montrent l'importance de la presse financière, le journal les Échos représente encore 4 473 616 mots dans le corpus 2009. Si on considère uniquement les médias pris en compte dans le corpus IntUne, tels que nous en avons parlé dans la partie 1.2, nous obtenons ainsi les statistiques suivantes :

IntUne Core 2007			IntUne Core 2009		
Source	Mots	Textes	Source	Mots	Textes
Le Figaro	4861988	77	Le Figaro	4858636	77
Libération	3602857	77	Libération	2973339	75
Sud-Ouest	3683377	77	Sud-Ouest	3786883	73
Ouest France	3226744	77	Ouest France	2166102	77
TF1	251551	32	TF1	-	-
France 3	183741	34	France 3	-	-
Total	15810258	376	Total	13784960	302



3- Technique de construction et d'analyse du corpus IntUne.



(Figures 4, 5 et 6 répartition des mots du corpus IntUne core 2007 et 2009 en fonction des médias)

III.2- Les analyses menées sur le corpus.

Une fois le corpus construit et fonctionnant, il s'agit de l'interroger, de l'analyser. Différentes possibilités existent dans ce domaine, différents logiciels, notamment, sont proposés dans l'étude de corpus. Deux logiciels ont été utilisés au cours de nos analyses : Xaira et Wordsmith dont nous avons parlé plus haut. Le choix de Xaira s'est imposé étant donné les capacités de ce logiciel à fonctionner sans difficultés majeures sur des corpus de taille importante, ainsi que sur son aptitude à gérer des structures complexes de textes. Le choix de Wordsmith a apporté un bon complément à l'utilisation de Xaira, notamment pour l'étude de lignes de concordance. Mais avant de préciser la technique d'utilisation de ces deux logiciels au cours de cette thèse, nous allons rapidement rappeler leur utilité en ce qu'ils permettent un fonctionnement de type corpus-driven, ainsi que l'identification de collocations.

Cette approche se veut dans un premier temps inductive ou de type corpus-driven (Tognini-Bonelli 2001). De cette façon, le linguiste ne cherche pas à imposer des patrons déjà existants, à faire se correspondre les données trouvées dans les ressources lexicographiques déjà existantes ou le premier prototype élaboré et les données extraites au cours de l'analyse du corpus. Il cherche plutôt à extraire de façon statistique, par exemple,

3- Technique de construction et d'analyse du corpus IntUne.

des combinaisons significantes de mots et à les étudier pour voir quelles relations se trouvent entre les mots. (Williams 2009). Dans cette thèse, si le mot étudié est « Europe », il sera ainsi procédé à la recherche de toutes les occurrences de ce mot dans le corpus, et non pas à la recherche de structures dont l'hypothèse aurait été faite dans un premier temps pour voir si elles y sont présentes ou non.

Firth avait une phrase fameuse quant à la façon dont on pouvait appréhender de façon plus adéquate le sens d'un mot : « *you shall know a word from the company it keeps* » (Firth 1957:11)⁶⁸. Cette phrase implique que c'est par le contexte du mot, non seulement les structures pouvant s'observer autour de ce mot mais aussi les différents autres mots que l'on peut y associer, que l'on parviendra à mieux comprendre celui que l'on étudie.

Les outils et techniques utilisés au cours du travail d'analyse de corpus dans cette thèse sont là avant tout pour permettre ce type de travail. Xaira dont il a déjà été question est un concordancier qui extrait des lignes des contextes au sein desquels les mots choisis apparaissent et qui calcule des collocations à l'aide de z-score de mutual informations pour le mot clef choisi. Wordsmith, également utilisé, permet notamment de classer les différentes lignes de concordances entre elles selon des choix effectués par le chercheur. Cela permet de rappeler que le but des études en linguistique de corpus n'est pas le développement d'outils, mais plus leur emploi et la façon dont on les utilise :

*« Le but n'est pas le développement des outils, mais l'emploi des outils pour regarder les mots en contexte à travers le mot-clé en contexte, KWIC »⁶⁹
(Williams 2006:155)*

L'étude de corpus pose le postulat que si des concepts sont reliés dans la réalité d'une communauté de discours (Swales 1990), alors ces relations pourront être atteintes à travers l'analyse du corpus :

« In so far as the central concepts relate to each other, similar relational patterns emerge in their surface constructs, words » (Williams 1998:151)⁷⁰

L'étude et le calcul des collocations des mots-clefs étudiés peut alors avoir un grand

68 « vous connaîtrez un mot à partir de la compagnie qu'il a ».

69 KWIC ou « Key Word In Context » est un format largement répandu parmi les divers concordancier, c'est celui qui est obtenu dans les deux concordanciers utilisés pour cette thèse. Il permet de voir apparaître le mot clef recherché au centre d'une ligne de contexte.

70 « A partir du moment où les concepts centraux sont reliés les uns aux autres, des structures relationnelles semblables émergent dans leurs constructions de surfaces, les mots. » (ma traduction).

3- Technique de construction et d'analyse du corpus IntUne.

intérêt. En effet, les relations collocationnelles peuvent être utilisées pour montrer ce qui est prototypique ou ce qui ne l'est pas au sein d'un corpus basé sur une communauté de discours (Williams 2002:45). Dans toute langue donnée, on peut retrouver des cadres conceptuels à travers des réseaux collocationnels qui seront prototypique dans le corpus étudié. L'analyse menée dans cette thèse se base donc sur ce type de calcul, notamment grâce au concours du logiciel Xaira qui permet pour tout mot clef d'en calculer les collocations, que ce soit par z-score ou par un système de Mutual Information score. Les calculs par Mutual Information score prennent en compte non seulement la fréquence de collocation pour deux termes au sein d'un corpus déterminé, mais également le fait de savoir si ces deux termes apparaissent généralement ensemble ou non. Les calculs par z-score reprennent cette idée d'une collocation plus signifiante quand elle est régulière mais inhabituelle avec des méthodes de comptage statistique plus complexes. La recherche de collocation permet de faire émerger un modèle prototypique du mot clef étudié. Elle a ensuite été complétée avec l'étude de lignes de concordances au format KWIC.

Une autre approche du contexte a en effet été menée suite à cette recherche des collocations : le mot-clé a été étudié par la lecture des différentes lignes de contextes au sein desquels il apparaît, chacune de ces lignes de concordance a ensuite été classée dans un contexte en fonction du domaine qui prédomine dans le contexte le plus proche du mot-clé. C'est à dire que si juste à côté du mot « Europe », on trouve les mots « ventes » ou « marché », on en induira que pour cette occurrence, le contexte dans lequel le mot-clé apparaît est « économique ». Cette démarche rencontre évidemment les obstacles principaux de la subjectivité du chercheur et de l'appel à des catégories (par exemple la catégorie « économique ») qui sont élaborées à partir de connaissances antérieures à l'étude sur corpus. Il est aussi problématique de devoir classer une occurrence dans une catégorie alors qu'en matière de signification, nous avons déjà vu que plutôt que de catégorie, il était à propos de considérer un continuum. La procédure qui a été appliquée dans le cas de ce classement consistait à se baser sur les mots qui étaient les plus proches du mot clef pour déterminer un contexte dominant, à la façon dont les prototypes fonctionnent. Bien que cette recherche se veuille inductive et essentiellement basée sur l'analyse contextualiste d'un corpus, il apparaît important de garder à l'esprit cette subjectivité du chercheur, cette introspection qui intervient toujours dans l'étude, mais sur laquelle, néanmoins, l'étude ne se fonde pas, puisque le modèle est élaboré à partir des collocations dans un premier temps, puis étayé en

3- Technique de construction et d'analyse du corpus IntUne.

fonction des découvertes réalisées dans l'analyse de la totalité des lignes de concordance.

Il existe également des formes d'unités récurrentes qui s'agglomèrent parfois autour d'un mot-clef, si au début ces unités sont le plus souvent de simples spécifications du mot-clef envisagé, il arrive qu'avec une grande fréquence d'usage, elles finissent par s'établir. (Williams 1998). On note ainsi parfois la présence de colligations avec certains mots-clefs qui peuvent nous renseigner sur le sens qui est majoritairement activé lorsque telle colligation apparaît. Williams note par exemple que le mot-clef « Europe » dans le corpus IntUne 2007 montre la formation de colligations qui font qu'en fonction de la colligation dont il est question, on ne parle pas de la même Europe (Williams 2008).

Le but de l'analyse inductive qui va suivre est de « typifier », à partir d'un calcul de collocations et de la lecture de nombreuses lignes de concordances, de donner la représentation dominante qui est faite d'un concept dans la presse française à travers différents mots clefs. Ainsi, les occurrences qui seront utilisées dans les chapitres à suivre ne seront que les plus « représentatives », les plus typiques de ce qui a été constaté par le chercheur au sein du corpus quand beaucoup de citations sont présentes avec un sens (ou un contexte) proche ou très proche.

III.3- Conclusion.

Le corpus étudié dans cette thèse est le produit d'un projet européen souhaitant étudier des thématiques politiques (au sens large du terme, des thématiques s'intéressant à la vie des citoyens). Il s'éloigne donc dans une certaine mesure des corpus constitués pour étudier le fonctionnement particulier d'une langue. Néanmoins, sa conception, sa structuration sont directement inspirées par des corpus de référence de la linguistique moderne comme notamment The Bank of English et le British National Corpus que nous avons évoqué dans le premier chapitre. Chacun des corpus français compte environ 15 millions de mots, sans compter le corpus satellite se composant du journal Les Échos, émanant par exemple de près de 8000 articles de presse et de journaux télévisés concernant le corpus 2007. Ils sont structurés selon une architecture suivant les contraintes de la TEI et disposent d'un système de balisage cohérent pour plusieurs pays dans le domaine de la presse écrite et télévisée.

3- Technique de construction et d'analyse du corpus IntUne.

Ce travail de choix quant à ce qui devait figurer ou non dans le corpus, quant à la façon dont il devait se présenter, ainsi que le travail d'élaboration pratique du corpus ont été des tâches longues et dévoratrices de temps, mais elles ont également permis d'affiner la conception de la recherche qui en découlait, et ont fourni un outil permettant au chercheur de pratiquer des analyses qualitatives basées sur l'analyse de discours, mais également plus quantitatives, basées alors sur une approche de type « corpus-driven ». Enfin, le fait d'avoir établi ces normes en plusieurs étapes, et au cours de plusieurs rencontres avec les équipes des autres pays participant au groupe Média a permis d'établir un corpus comparable au niveau de sa structure.

Le corpus une fois construit et fonctionnant, son analyse peut être mise en place, nous avons vu lors du chapitre précédent dans quel cadre théorique elle se situe, nous allons maintenant la présenter plus précisément. Le quatrième chapitre de cette thèse va ainsi montrer quelle « Europe » la presse française nous donne à voir, à travers l'élaboration d'un modèle suite au calcul de collocations pour ce mot clef et de la lecture des lignes de concordances dans lesquelles ce mot apparaît au format KWIC. Nous pourrions ainsi compléter ou modifier le premier prototype lexicographique élaboré pour ce mot dans le second chapitre, et vérifier quelle actualisation il rencontre dans la presse française.

4- « Europe ». Quelle représentation dans la presse française ?

IV.0- Introduction.

Cette thèse se donne pour but d'explorer un corpus de presse française avec des méthodes issues de la linguistique de corpus, afin de mieux appréhender la réalité de ce qu'est cette entité, l'« Europe », à travers quatre journaux quotidiens largement diffusés en France.

L'Europe est une entité à la réalité des plus complexes, et reposant sur une grande multiplicité de sens, ou d'acceptions. Par conséquent, savoir de quoi l'on parle lorsque l'on emploie ce terme semble problématique. Le dictionnaire Larousse en donne la définition suivante :

« une des cinq parties du monde, comprises entre l'océan Arctique au N., l'océan Atlantique à l'O., la Méditerranée et ses annexes, ainsi que, traditionnellement, la chaîne du Caucase au S., la mer Caspienne et l'Oural à l'Est. »(Larousse, 1992:1279)

Si au niveau géographique, cette définition précise que cette entité se situe entre les océans arctique, atlantique et la Méditerranée, un premier problème peut être soulevé par le

4- « Europe ». Quelle représentation dans la presse française ?

mot « annexes », jusqu'où, exactement, l'Europe va-t-elle ? La Mer Noire peut-elle être considérée comme une annexe de la Méditerranée ? Le même dictionnaire donne de cette mer la définition suivante :

« mer bordière de l'Atlantique, entre l'Europe méridionale, l'Afrique du Nord et l'Asie occidentale, couvrant 2,5 millions de km². Elle communique avec l'océan par le détroit de Gibraltar et avec la mer Rouge par le canal de Suez. (...) L'étranglement compris entre la Sicile et la Tunisie la divise entre deux bassins : la Méditerranée occidentale, avec son annexe, la mer Tyrrhénienne, et la Méditerranée orientale, plus ramifiée, avec ses dépendances (mer Ionienne, mer Adriatique et mer Egée). » (Ibid : 1465).

Selon cette définition, la mer Méditerranée exclue donc la mer Rouge et l'océan Atlantique avec lesquels elle communique. Elle se compose de diverses annexes qui sont citées, mais rien ne nous est dit concernant la mer Noire : Est-elle également une annexe ou bien communique-t-elle simplement avec cette mer via le détroit du Bosphore ? Cette question reste en suspens quant à la définition géographique de l'Europe. L'adverbe « traditionnellement » pour définir les autres limites géographiques de cette entité pose également problème en ce qu'il nuance la définition donnée, ces limites n'ont qu'une caution de tradition, elles ne sont pas affirmées sans nuance. Si à la suite de cette définition, on s'essaie à tracer les contours géographiques de l'Europe, on peut néanmoins définir une zone dans laquelle se trouvent plusieurs pays (cf Annexe 7).

Le *genus proximum* auquel est associée cette entité semble être dans la définition qu'en donne le Larousse « cinq parties du monde ». Là encore, la définition est plutôt vague : qu'est-ce qu'une partie du monde ? Si le dictionnaire s'appuie cette fois de façon implicite sur la « tradition », on peut supposer qu'il s'agit d'une paraphrase utilisée pour désigner le mot « continent » dont on a le plus souvent appris qu'il en existait bien cinq dans le monde (Williams 2008) : l'Amérique, l'Europe, l'Asie, l'Afrique et l'Océanie. Il est intéressant de noter que le Larousse donne du mot continent la définition suivante :

« Vaste étendue de terre qu'on peut parcourir sans traverser la mer. L'ancien continent : l'Europe, l'Asie et l'Afrique. - Le Nouveau continent : l'Amérique. » (Ibid:261).

On comprend mieux, du fait de cette définition, que ce mot ne soit pas employé de

4- « Europe ». Quelle représentation dans la presse française ?

façon explicite comme *genus proximus* pour définir l'Europe, puisqu'alors, de nombreux pays qui y sont dans un premier temps inclus (cf Annexe 7) n'en feraient plus partie.

Ainsi donc, bien que considérée uniquement au niveau géographique, cette entité semble poser un réel problème de définition. Mais lorsqu'on parle d'Europe, il n'est pas seulement question d'une entité géographique, d'une « partie du monde » dans laquelle certains pays n'auraient d'autre choix que de se situer. Il peut aussi s'agir de façon métonymique d'une création humaine, d'un ensemble politique qui est né en 1957 du Traité de Rome et est actuellement « l'Union Européenne ».

La réalité alors recouverte par ce mot est plus que variable : s'agit-il de tous les pays d'une zone définie tant bien que mal au niveau géographique ? De tous ceux qui appartiennent actuellement à l'Union Européenne, ou de ceux qui y ont appartenu depuis 1957 jusqu'en 2007 ? De tous ceux qui partagent une même monnaie européenne : l'euro ? Est-ce une zone qui partage des points communs au niveau culturel, religieux ou historique ? Ce sont à ces questions que l'élaboration d'un prototype sémantique tentera de répondre.

Le chapitre qui va suivre se propose de mettre en lumière les éléments que la presse française propose de façon diffuse pour répondre à ces questions, de décrire la représentation du concept « Europe » dans cette même presse. Différentes techniques explicitées dans le chapitre précédent seront utilisées à cette fin, permettant de mettre progressivement au point un modèle qui rende compte de la manière dont est activé le prototype du concept « Europe » dans la presse française. L'utilisation de deux corpus, l'un collecté en 2007 et l'autre en 2009, permettra également de procéder à des analyses comparatives à un niveau diachronique. Deux mots clefs seront utilisés dans cette étude « Europe », et « européen (ne, nes, s) », ce dernier venant en partie compléter l'étude menée sur le concept "Europe".

Tout d'abord, nous allons présenter une vision globale du mot « Europe » dans le corpus IntUne en 2007 et 2009. Après une analyse de la présence en termes quantitatifs de ce mot dans la presse française dans les deux périodes étudiées, la vision de la façon dont l'Europe est présente dans le corpus sera affinée grâce à une étude des collocations, et des divers réseaux sémantiques que celles-ci tissent autour du concept « Europe » au sein de la presse française. Cette étude des collocations nous permettra de mettre au point un modèle qui rende compte des principales acceptions, des principaux sens potentiels que le concept

4- « Europe ». *Quelle représentation dans la presse française ?*

étudié prend dans la presse française. Nous confronterons alors le prototype lexicographique du concept « Europe », cité précédemment à la fin du second chapitre, à cette analyse des collocations, le modifiant si besoin est, ou le complétant.

Dans une seconde partie, nous donnerons une présentation plus détaillée de l'activation du prototype de l'entité dans la presse française en 2007 puis en 2009. Nous verrons quels contextes sont associés à quels sens potentiels, et ainsi, quel portrait de l'Europe les médias français brossent au cours de ces deux périodes, en tachant d'en montrer l'évolution.

IV.1- L'Europe dans le corpus IntUne : vue globale.

La première approche du concept « Europe » dans le corpus s'est voulue exploratoire. En conséquence, nous avons tout d'abord recherché le mot « Europe » dans le corpus IntUne en 2007 et en 2009. Puis nous avons mené une étude des collocations par z-score et MI à l'aide du logiciel Xaira. Cette première prise de contact avec le concept dans le corpus IntUne a permis de faire apparaître un premier type d'informations quantitatives quant à la présence de ce mot dans le corpus, son évolution d'un corpus à l'autre, sa répartition dans les différents journaux français, etc. D'autre part, l'analyse des collocations a permis de voir quelles associations de sens ou de contextes revenaient le plus fréquemment et de façon signifiante avec le mot « Europe ». Dans un premier temps, nous allons ici présenter en quelques chiffres la façon dont le mot « Europe » est présent dans les médias du corpus IntUne, puis dans une seconde partie, nous procéderons à une analyse des collocations qui sont associées à ce concept. Enfin, dans une troisième partie, par la confrontation du prototype lexicographique avec les données issues de notre étude collocationnelle, nous élaborerons un prototype qui rende compte de la façon dont le mot « Europe » est présent dans la presse française en 2007 et en 2009.

IV.1.1- L'Europe en chiffres dans le corpus.

Une première observation rapide du mot « Europe » dans le corpus permet d'obtenir quelques informations sur sa représentation dans la presse française. Dans ce paragraphe, les données ont été obtenues pour les deux corpus à l'aide du logiciel Xaira. Dans le corpus

4- « Europe ». Quelle représentation dans la presse française ?

2007, ce mot est présent à 9650 reprises réparti dans 417 textes, se distribuant de la façon suivante dans les journaux :

<i>Europe dans le corpus IntUne en 2007.</i>					
Source	occurrences	total de mots	%	textes avec « Europe »	total des textes
Le Figaro	2617	4861988	0.054	77	77
Libération	1088	3602857	0.03	77	77
Les Échos	3269	5015207	0.065	63	63
Sud-Ouest	1228	3627131	0.034	76	76
Ouest France	1279	3226744	0.04	77	77
TF1	85	251551	0.034	23	32
France 3	84	183741	0.046	24	34
Total	9650	20769219	0.043	417	436

(Figure 7, répartition des occurrences de « Europe » dans les journaux du corpus IntUne 2007)

La première colonne de ce tableau concerne les différents journaux du corpus, la deuxième le nombre de fois où le mot Europe est présent en fonction du journal. La troisième colonne représente le nombre total de mots dans chaque journal, la quatrième colonne rapporte le nombre d'occurrences du mot « Europe » au total de mots de chaque journal et exprime ce rapport en pourcentage, enfin, les deux dernières colonnes donnent tout d'abord le nombre de textes où le mot « Europe » apparaît, et ensuite le nombre de textes total pour le journal dont il est question.

Il apparaît clairement que le journal mentionnant le plus ce terme est les Échos, suivi par le Figaro, Ouest France, Sud Ouest, et enfin Libération. Cette répartition donne lieu à différents constats. D'une part, c'est un journal financier qui semble accorder le plus d'importance au mot « Europe ». De plus, ce mot semble donc être abordé plus volontiers par un journal avec une ligne éditorialiste de « droite » que de « gauche » en période de campagne pour l'élection présidentielle, et ce, qu'il s'agisse de la présence du mot au niveau absolu, ou en rapport avec le total des mots présents pour chaque journal. Dans les deux cas, le quotidien national Les Échos est celui qui donne le plus d'importance au mot « Europe » et Libération, celui qui en donne la moins forte. Toujours au niveau relatif, le mot est davantage présent dans les journaux télévisés de France 3, néanmoins en présence absolue, TF1 donne

4- « Europe ». Quelle représentation dans la presse française ?

une occurrence supplémentaire d'Europe avec 85 citations contre 84 sur France 3.

Le corpus 2009 se compose de moins de textes (ce qu'on appelle texte consiste en des fichiers. Un fichier représente un jour de collecte pour un journal.) En conséquence, le mot « Europe » est moins présent en termes absolus, il se retrouve ainsi à 7050 reprises dans 362 textes. La répartition de ce mot est la suivante :

<i>Europe dans le corpus IntUne en 2009</i>					
Source	occurrences	Total de mots	%	textes avec « Europe »	total des textes
Le Figaro	2185	4858636	0.045	77	77
Libération	814	2973339	0.027	74	75
Les Échos	2468	4473616	0.055	63	63
Sud-Ouest	947	3786883	0.025	72	73
Ouest France	636	2146102	0.03	76	77
TF1	0	0		0	0
France 3	0	0		0	0
Total	7050	18238576	0.036	362	365

(Figure 8, répartition des occurrences du mot « Europe » dans le corpus IntUne 2009.)

Un premier constat entre ces deux corpus pourrait être que la représentation de l'Europe chute au niveau de l'importance de ce mot par rapport au total des mots composant le corpus : de 0,043 % des occurrences totales du corpus en 2007, on arrive à 0,036 % de ces occurrences en 2009. Le journal les Échos est toujours celui qui accorde le plus d'importance à l'entité avec 1372 occurrences soit 0,055 de ses occurrences. En termes relatifs, Le Figaro se situe en deuxième position avec 0,045 % des occurrences totales, tandis que le régional Ouest France arrive en troisième position, le mot « Europe » y représentant 0,030 % de l'ensemble du journal. Libération arrive cette fois en quatrième position, avant Sud Ouest.

Avec ces premières recherches dans le corpus, on s'aperçoit donc qu'entre 2007 et 2009, l'intérêt des journaux français semble moins se porter sur le sujet européen. En effet, le mot est moins présent que ce soit en valeur absolue ou relative, voire totalement occulté dans certains fichiers du corpus (un fichier représente un jour de collecte dans un journal). D'autre part, une tendance assez forte montre aussi que ce terme est plus facilement traité d'abord par le journal financier les Échos, puis par un journal ayant une ligne éditorialiste plutôt de

4- « Europe ». Quelle représentation dans la presse française ?

droite. Cependant, le journal de gauche a amélioré sa représentation du mot, en comparaison des autres journaux du corpus entre 2007 et 2009 (passant ainsi de la cinquième position à la quatrième). Cette tendance à la baisse peut s'expliquer par le fait qu'en 2007, le corpus a été collecté alors que la campagne présidentielle française battait son plein, « l'Europe » aurait alors constitué un des thèmes majeurs de cette élection, l'Europe, entité à la fois politique et géographique, aurait alors constitué un des thèmes majeurs de cette élection.

Néanmoins, il est étonnant que la campagne de 2009 pour l'élection du parlement européen n'ait pas eu ce même effet. Ce constat révèle que la presse française parle d'Europe avant tout dans un cadre national, le cadre européen n'étant que peu prétexte à la discussion de ce thème. En outre, en 2009, les médias ont été prolixes sur le thème de la crise économique internationale, les temps de crise économique étant souvent propices à un certain renfermement conduisant parfois à des mesures protectionnistes expliquant peut être l'ouverture moindre à l'Europe visible dans les médias du corpus IntUne en 2009.

Un autre indicateur de la présence de la thématique européenne dans les médias du corpus IntUne concerne les occurrences du lemme « européen » et ses déclinaisons en genre et en nombre. Ce lemme est présent en 12690 occurrences en 2007, se répartissant dans l'ensemble des 370 textes du corpus de presse. Le journal dans lequel il est le plus représenté est à nouveau les Échos, cette fois, l'écart avec les autres journaux du corpus est plus important, les Échos se situe en effet à 0,04 points au dessus de la moyenne du corpus. L'ordre habituel des journaux traitant du sujet européen reste le même que pour le mot « Europe », nous trouvons ce lemme en seconde position dans le Figaro, puis Ouest France, Libération et enfin Sud Ouest :

<i>Européen dans le corpus IntUne en 2007</i>					
Source	occurrences	total de mot	%	textes avec « Europe »	total des textes
Le Figaro	3304	4861988	0,068	77	77
Libération	1466	3602857	0,041	77	77
Les Échos	5017	5015207	0,1	63	63
Sud-Ouest	1293	3627131	0,036	76	76
Ouest France	1610	3226744	0,05	77	77
Total	12690	20333927	0,059	370	370

(Figure 9, répartition des occurrences du lemme « européen » dans le corpus IntUne 2007).

4- « Europe ». Quelle représentation dans la presse française ?

Si nous nous intéressons maintenant à la répartition du lemme dans le corpus en 2009, nous obtenons les résultats suivant à l'aide de Xaira :

<i>Européen dans le corpus IntUne en 2009</i>					
Source	occurrences	total de mot	%	textes avec « Europe »	total des textes
Le Figaro	3709	4858636	0,076	77	77
Libération	1666	2973339	0,056	75	75
Les Échos	3916	4473616	0,088	63	63
Sud-Ouest	1260	3786883	0,033	73	73
Ouest France	858	2146102	0,04	77	77
Total	11409	18238576	0,0586	365	365

(Figure 10, répartition des occurrences du lemme « européen » dans le corpus IntUne 2009).

Dans ce corpus, le lemme est toujours davantage présent dans Les Échos, cependant, la différence est moins importante entre Les Échos et les autres journaux du corpus. Le Figaro vient en seconde position, suivi de Libération, Ouest France et Sud Ouest. Une fois de plus le national de droite est celui qui traite le plus ce sujet, tandis que le régional Sud Ouest est celui qui en parle le moins.

Maintenant que nous avons vu en quelques chiffres quelle était la présence de la thématique européenne dans le corpus IntUne en 2007 et 2009, et quels journaux en parlaient le plus, après avoir vu le « qui » en parle avec quelle fréquence, nous allons nous intéresser au « qu'en dit-on », aux contextes dans lesquels cette thématique apparaît, à travers une étude des collocations des deux mots clefs « Europe » et « européen ».

IV.1.2- Quelle est la compagnie du concept Europe ?

IV.1.2.1- Le mot « Europe ».

Par le mot « compagnie », nous entendons ici les collocats qui sont associés au mot « Europe » dans le corpus IntUne. Dans ces paragraphes, les collocations seront calculées par z-score via le logiciel Xaira, elles ont été confrontées, toujours à l'aide de Xaira à des collocations MI calculées selon les mêmes modalités, pour voir si l'ensemble concordait. La fenêtre choisie pour ces collocations est de trois mots à gauche et trois mots à droite. Pour

4- « Europe ». Quelle représentation dans la presse française ?

dresser un tableau des collocations les plus significatives, les données obtenues dans Xaira sont enregistrées puis retravaillées à l'aide d'une feuille de calcul s-calc. Le fichier Xaira est ouvert via s-calc, sont alors retirés tous les collocats apparaissant avec une fréquence inférieure à trois (Stubbs 1995:180), puis les données sont à nouveau triées par ordre décroissant en fonction de leur z-score ou MI. On obtient alors les collocations les plus significatives pour le mot « Europe ». Parmi ces collocations, il arrive fréquemment que l'on trouve des virgules, prépositions ou autres mots grammaticaux. Dans les paragraphes qui vont suivre, le choix a été fait de les laisser de côté pour se concentrer sur les collocats qui ont un poids lexical. Pour le mot « Europe », cette procédure donne environ de 1500 (avec un calcul de MI) à 1800 collocats (pour un calcul de z-score). Nous n'avons pas étudié l'ensemble de ces collocats, nous nous sommes en effet concentrés sur les 100 plus significatives, ceux arrivant en tête du classement établi à l'aide de s-calc. L'annexe 8 présente un tableau de ces cent premiers collocats pour le mot « Europe » dans les corpus IntUne 2007 et 2009. Ces collocats présentent à la fois des similitudes d'une année à l'autre, et des différences que nous allons montrer. D'autre part, ils se regroupent selon certaines thématiques, permettant ainsi de compléter le premier prototype lexicographique élaboré pour le concept « Europe ». Le deuxième tableau de l'annexe 8 donne la légende du regroupement thématique des différents collocats, en les comptabilisant pour chaque année.

Les groupes de collocats peuvent brièvement être exposés comme suit pour 2007 et 2009 :

Thème	Collocats en 2007	Collocats en 2009
Différentes zones.	orientale, continentale, centrale, occidentale, ouest, nord, sud, méridionale, est.	orientale, centrale, occidentale, continentale, ouest, nord, sud, est
Sport	coupe, championnats, ettu, champion, cers, championne, vice-champion, championnat, rink-hockey, rugby, boycott, coupes, champions, anoeta, championnes	coupe, championnats, xxl, niederösterreich, champion, surfrider, quicksilver, championne, championnat, indoor, champions, vice-champion, coupes, record
Géographie	en, Seine-nord, partout, hors	en, Seine-nord, partout
Économie	airpost, ios, aew, streettracks, netjets, snc-lavalin, schott, msci, saint-groux, infiniti, ace, lyle, ixis, aig, rodamco, adobe, toyal, mutuel pharmaceuticals, ps3, gros-verheyde, vice-président	aew, émergente, carl-peter, chi-x, forster, 1664, stx, arca, gm, lycos, afi, schweppes, hauts-fourneaux, faugère, kronembourg, hudson, foundation, snc-lavalin, visa, colony, libérale, computer, filiales, opel, vice-président, lgv, endettés, implantés.
Culture	odéon-théâtre, 1/tv5, chrétiennes, tv5, libénautes, civilisations, chatignoux,	fogiel, civilisations, rendez-vous », Marc-Olivier, expressions.

4- « Europe ». Quelle représentation dans la presse française ?

	scouts, tv	
Autres zones	asie, états-unis, états-unis, moyen-orient, Amérique, Afrique, EMEA	états-unis, Asie, états-unis, moyen-orient, Amérique, Proche-Orient, Japon, Atlantique, Asie-Pacifique, Afrique
Défense	conventionnelles, antimissile, OSCE, FCE, bouclier	antimissile, défense, OSCE, défenses, antimissiles
Marque	1, express, val	1
Membres	pays, France, patries, nations, Budapest	pays
Symbole	vieille, toute, unie, passoire, pacifiée, sociale, désamour, protectrice	protège », vieille, froide, forte », unie, protège, entière, toute, solidaire, « grand
Politique, Union Européenne.	élargie, supranationale, référendum, relance, panne, restrictives, relancer	lipietz, devises, construire, Delors, liste.
Souhait		osons, voudrions, voulons
Écologie		écologie, ecologie
?	turquie	turquie

(Figure 11, regroupement thématique des cent premiers collocats de « Europe » dans le corpus IntUne 2007 et 2009.)

A travers ce tableau, ainsi que ceux de l'annexe 8, on voit clairement une cohérence entre les collocations qu'entretient le mot « Europe » dans la presse française en 2007 et celles qu'il entretient en 2009. On peut donc en déduire que « Europe » prend plusieurs sens dans cette presse.

Dans l'ordre des collocations, le premier sens qui apparaît est celui d'une entité divisée en plusieurs zones géographiques (Europe orientale, occidentale, de l'ouest, etc.), ce sens est également celui qui apparaît clairement au tout début des collocations pour l'année 2009, nous pouvons par conséquent penser qu'il s'agit là d'un sens potentiel fortement activé dans la presse française dès lors que l'on parle de l'Europe.

D'autre part, une forte présence de la thématique sportive se remarque lorsque ce mot est présent dans les journaux du corpus IntUne, on parle alors de coupe d'Europe, champion d'Europe, etc. Cette thématique sportive semble importante pour décrire ce qu'est l'Europe, et en quelque sorte constitutive voire harmonisante pour l'image européenne. Il peut donc s'agir ici d'un deuxième sens potentiel pour le mot Europe.

Une thématique géographique apparaît ensuite à l'aide de différentes prépositions ou adverbes qui marquent une notion de localisation. Cette thématique peut également se retrouver dans la mention d'organes de transports existant au sein de l'Europe. Ce troisième groupe de collocats peut donc laisser penser qu'un sens potentiel fréquemment activé dans la

4- « Europe ». *Quelle représentation dans la presse française ?*

presse française pour le concept « Europe » est celui d'une étendue géographique.

Un ensemble important parmi les collocations de « Europe » dans le corpus IntUne concerne le domaine économique, il peut alors s'agir de données macro-économiques, concernant un état ou un ensemble d'états, mais aussi de données micro-économiques, à la dimension d'une entreprise (notamment le nom de telle ou telle entreprise, sa division pour la partie « Europe » lorsqu'il s'agit d'une multinationale).

Certains collocats font apparaître un contexte plus culturel, au sens large du terme. Dans ce cadre, il peut s'agir à la fois de religions, de civilisations ou encore de médias ou de manifestations culturelles (théâtre de l'Odéon, télévision, etc.). Ce groupe de collocats peut alors correspondre à un sens potentiel « entité culturelle » pour le concept Europe dans la presse française.

Plusieurs toponymes apparaissent dans les collocations du mot « Europe » dans le corpus IntUne, que ce soit en 2007 ou en 2009. On peut à la fois trouver dans ce groupe des continents (Asie, Afrique, Amérique), mais également des États (États-unis), ou encore des zones non institutionnalisées (EMEA, Moyen-Orient). Ces différents collocats mettent en lumière les relations que l'Europe entretient avec d'autres zones plus ou moins définies du monde.

Le contexte de la défense armée est aussi présent à plusieurs reprises à travers ces collocations, il peut s'agir d'armes qui sont citées (bouclier antimissile, notamment), mais aussi d'organisations pour un ou plusieurs pays chargées de la sécurité armée de ces pays (OSCE qui signifie Organisation pour la Sécurité et la Coopération en Europe, par exemple), ou encore de décisions politiques dans ce domaine (à travers le collocat « conventionnelles » qui fait référence à un traité politique signé par différents partenaires).

Le concept Europe peut aussi être une partie d'une marque, comme cela est visible avec les collocats 1 ou express (Europe 1, France Europe Express), par exemple. Il est à noter que le collocat express n'apparaît plus parmi les cent premiers dans le corpus en 2009, cela étant dû à son lien avec le domaine politique particulièrement mis en relief lors de la campagne présidentielle (corpus 2007).

Différentes collocations font intervenir des noms représentant à la fois des lieux

4- « Europe ». *Quelle représentation dans la presse française ?*

constitutifs du concept "Europe" mais également des membres et partenaires de cette institution à l'instar des collocs « pays », « nation », « France » ou encore « Budapest », ces différentes entités montrent qu'un autre sens potentiel pour le concept étudié pourrait être une entité se composant de différents membres.

Une thématique intéressante qui apparaît à travers l'étude de ces collocations est celle de la symbolique. En effet, l'Europe est souvent associée à des termes qui concerneraient normalement davantage une mère. Dans un procédé métaphorique, elle est alors protectrice, ou forte, mais également vieille, voire frigide. D'autre part, elle peut être présentée, toujours selon ce principe métaphorique comme une communauté de divers membres qui est unie, solidaire, pacifiée, entière, etc. Elle est parfois tout simplement comparée à des objets de la vie courante, comme dans son association avec le collocat « passoire », toujours dans un phénomène d'image. L'Europe apparaît alors comme une entité qui est fortement symbolisée. Cette idée de symbolisation rejoint un autre groupe de collocs présents uniquement dans le corpus IntUne 2009, ce groupe est formé de verbes conjugués à la première personne du pluriel qui expriment tous un souhait, un vouloir irréalisé face à cette entité européenne. C'est donc, outre une entité symbolique, une entité qui suscite différents espoirs, d'autant plus en 2009.

Les collocations laissent aussi apparaître un contexte plus politique, il peut alors s'agir clairement de l'Union Européenne et de ses divers organes de gouvernance et de fonctionnement, mais aussi de différents hommes politiques qui évoluent au sein de cette entité, ou de membres de l'entité. Si l'institution Union Européenne est relativement présente via les collocations en 2007, on s'aperçoit que c'est moins le cas en 2009. D'autre part, en 2007, elle semble éprouver quelques difficultés à travers les collocs « relance », « relancer », « panne », « restrictives », difficultés qui sont associées dans ce contexte aux collocs « référendum », « supranationale », et « élargie ». En 2009, le rapport collocationnel entre Europe et Union Européenne est beaucoup moins explicite, il fait intervenir les collocs « construire », ou « liste », mais aucun ne se réfère explicitement et directement à l'Union Européenne.

Nous avons déjà vu que le corpus 2009 faisait intervenir un sens potentiel qui n'apparaissait pas en 2007 dans la presse française avec les verbes exprimant un souhait non réalisé. Une autre thématique parmi les collocations apparaît aussi dans le corpus en 2009, il

4- « Europe ». *Quelle représentation dans la presse française ?*

s'agit d'un contexte écologique, avec le mot « écologie » présent avec et sans accent. En ce qui concerne les évolutions au sein des collocations de « Europe » en 2009, on note également que le contexte de défense armée semble prendre de l'importance, puisqu'il arrive dans les quarante premières collocations avec plusieurs collocats, ce qui n'était pas le cas en 2007. Si l'on regarde toujours l'ordre d'importance des différentes collocations, l'on remarque que l'Union Européenne apparaît plus tard de façon explicite parmi les collocats de 2009, et qu'elle y est moins en concurrence directe avec d'autres collocats comme « nation », « patrie », ou « pays » comme en 2007, elle n'est pas non plus présentée aussi fortement comme étant « supranationale ». Comme nous venons de le voir, elle est davantage présente de façon imprécise, souvent symbolique, et parfois associée à des verbes traduisant un souhait non réalisé, ce qui n'était pas le cas en 2007.

Au niveau économique, on note une plus forte présence des collocats exprimant cette thématique en 2009, de plus, des collocats exprimant une réalité macro-économique apparaissent lors de cette deuxième année de collecte. Ces différences peuvent s'expliquer par le contexte de crise économique internationale qui est relaté dans les journaux du corpus IntUne en 2009. Les collocats qui parlent de civilisation, ou de religions semblent également connaître un comportement différent entre 2007 et 2009, ils sont en effet quasiment absents des cent premières collocations en 2009, et les collocats en rapport direct avec la religion ne sont plus présents.

Un dernier collocat n'a pas été regroupé avec les autres dans notre étude. Il s'agit du collocat « Turquie » qui apparaît en quarante-huitième position dans le corpus en 2009, il est présent également en 2007, mais seulement en cent-dixième position avec 17 occurrences. Ce collocat n'a pas été classé puisqu'il pouvait à la fois se situer avec les « voisins », les autres zones avec lesquelles l'entité Europe entretient des relations, mais aussi avec les « membres » de l'entité, puisqu'au niveau sportif, nous le verrons, la Turquie prend régulièrement part à des compétitions européennes.

Ce premier aperçu des collocations du mot « Europe » dans le corpus IntUne 2007 et 2009 nous permet de voir quels sens potentiels la presse française active lorsqu'elle parle de ce mot, ainsi que les évolutions qui semblent apparaître dans cette représentation entre 2007 et 2009. Cependant, l'image peut se compléter avec une étude des collocations du lemme « européen », c'est à dire le mot, ainsi que ses déclinaisons en genre et en nombre.

4- « Europe ». Quelle représentation dans la presse française ?

IV.1.2.2- Le lemme « européen ».

En procédant de la même façon que précédemment, nous obtenons les collocations du lemme « européen » dans le corpus. En retenant les 30 premiers collocats sans compter les mots non lexicaux, nous obtenons le résultat suivant :

2007				2009		
Rang	collocat	fréquence	z-score	collocat	fréquence	z-score
1	union	1378	350,3	union	1106	313,7
2	commission	752	211,5	élections	518	198
3	parlement	276	107,8	commission	699	180,3
4	avionneur	128	104,8	parlement	343	135,2
5	commissaire	187	104,1	commissaire	156	85,4
6	constitution	212	99,8	centrale	210	80,9
7	centrale	251	97,8	directive	68	69,2
8	directive	94	76,2	bce	88	67,8
9	député	186	73,4	banque	315	63,8
10	banque	318	69,3	07-juin	40	61,1
11	monnaie	100	61,4	listes	82	49,4
12	esa	22	56,6	juin	160	49,3
13	comité	178	52,7	député	128	47,8
14	spatiale	45	51,7	pays	381	46,8
15	construction	156	48,3	directives	22	44,1
16	devise	48	48	campagne	142	42,5
17	traité	108	46,1	bourses	46	39,1
18	eurostat	29	44,7	aux	649	38,5
19	pays	386	43,3	spatiale	24	35,6
20	bce	54	42,3	capitales	24	33,3
21	directives	26	40,6	les	2236	33,1
22	lunion	13	38,6	tournée	59	33,1
23	la	3504	38	raffineurs	6	32,5
24	l'	2824	38	condimentaires	4	32,5
25	cour	130	36,2	liste	99	30,4
26	communauté	99	35,8	diplomate	31	30,1
27	ministres	73	33	avionneur	19	29,4
28	voisins	56	32,3	l'	2227	28
29	référendum	63	32	places	63	27,4
30	constitutionnel	58	32	régulateurs	21	27,1

4- « Europe ». Quelle représentation dans la presse française ?

31	places	78	31	niveau	131	26,8
32	missilier	5	30,8	la	2788	26,6
33	bourlanges	9	30,5	défense	109	25,5

(Figure 12, collocations du lemme « européen » dans les corpus IntUne 2007 et 2009)

Un coup d'œil rapide à ces collocations permet d'effectuer là aussi des regroupements thématiques cohérents d'une année à l'autre qui donnent des informations sur la façon dont la presse française parle d'Europe. Lorsque quelque chose est « européen », cela recouvre dans la majorité des cas un organe de gouvernance de l'Union Européenne, « union », « commission », apparaissent tous deux en premières positions pour les collocations du mot « européen » dans le corpus IntUne, quelle que soit l'année concernée. En outre, on retrouve les mots « parlement », et « commissaire » qui rejoignent cette thématique des organes de gouvernance de l'Union Européenne. Cet ensemble est représenté dans le tableau ci-dessus par une couleur bleu clair. La thématique politique est également présente en 2009 à travers des collocations se rapportant davantage à la question de l'élection législative européenne. Elle est représentée dans le tableau par la couleur bleu foncé.

Une autre thématique concernant également l'Union Européenne est représentée dans le tableau avec les collocats colorés en vert clair, il s'agit alors de projets menés grâce et au sein de l'Union Européenne. Selon le corpus dans lequel on se situe, ces projets sont plus ou moins apparents, le corpus 2007 comporte ainsi les collocations les plus nombreuses. .

Le contexte économique est également présent dans ces corpus avec des collocats se rapportant essentiellement au monde boursier. Cette thématique est marquée par la couleur rouge dans le tableau. Avec les collocats "pays", "capitales", colorés en jaune; les membres sont aussi mentionnés, tout comme le contexte de défense armée à travers les mots "défense" et "missiliers".

Les collocations du lemme « européen » permettent ainsi de compléter l'image fournie par la première étude des collocations du mot « Europe ». Les mêmes groupes thématiques apparaissent, néanmoins, que ce qui est européen concerne essentiellement les organes de gouvernance de l'Union Européenne, tandis que l'Europe concerne un spectre plus large de thématiques. Dès lors, il est possible de reprendre le prototype sémantique lexicographique déjà donné pour le concept « Europe », et de le modifier pour qu'il puisse rendre compte de la façon dont l'Europe est présentée dans la presse française.

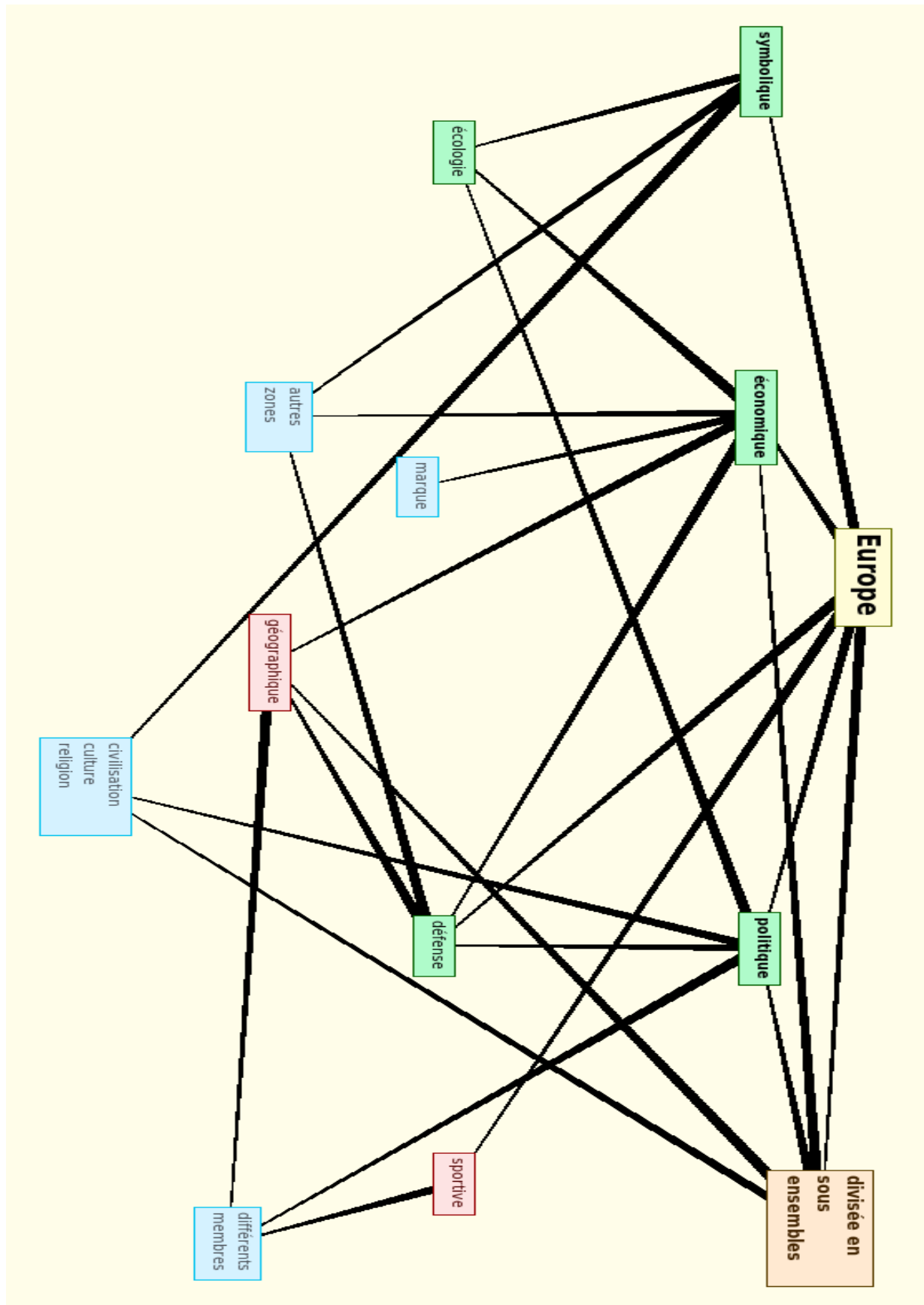
4- « Europe ». Quelle représentation dans la presse française ?

- « Une masse terrestre.
- Situé dans l'hémisphère Nord.
- Entouré d'îles plus ou moins grandes.
- Une péninsule d'Eurafrasié.
- Une péninsule d'Eurasie.
- Un continent.
- Situé entre l'Asie et l'océan Atlantique.
- Délimité à l'ouest par l'océan Atlantique.
- Délimité au Sud par la méditerranée.
- Délimité au Sud par la chaîne du Caucase.
- Délimité à l'Est par la mer Caspienne et l'Oural.
- Plus ou moins précisément délimitée
- Composée de plusieurs pays.
- Dont les pays sont regroupés dans des blocs géographiques.
- Dont les pays sont regroupés dans des blocs politiques.
- Partiellement composé d'une unité politique : l'Union Européenne.
- L'Union Européenne est divisée en sous-ensembles.
- L'Europe politique doit défendre ses frontières.
- L'Europe politique est une idéalisation.
- L'Europe est composée de pays continentaux par rapport à la Grande Bretagne.
- Une entité sportive.
- Une entité économique. »

Les sens potentiels grisés sont ici ceux qui sont présents dans les journaux du corpus IntUne, tandis que ceux soulignés n'apparaissent pas dans le premier prototype. D'autre part, on peut représenter le réseau collocationnel du concept « Europe » dans la presse française à l'aide du schéma suivant qui montre à quel point chaque collocation est tributaire d'une autre, et combien ces différents sens potentiels sont combinés les uns aux autres, comment les uns impliquent les autres. Les trois sens mis sur le même plan en vert, politique, économique et symbolique, entretiennent de plus une relation de proximité et d'interdépendance les uns avec les autres.

Maintenant qu'un réseau collocationnel pour le concept « Europe » a été mis en relief, et qu'un prototype a été élaboré rendant compte de la façon dont l'Europe est représentée dans la presse française, nous allons pouvoir analyser quelle actualisation de ce prototype est effectuée dans le corpus IntUne en 2007 et en 2009.

4- « Europe ». Quelle représentation dans la presse française ?



(Figure 13, réseau collocationnel du concept « Europe » dans la presse française.)

IV.2- Actualisation du prototype sémantique de « Europe » dans la presse française en 2007 et 2009.

Le prototype précédemment complété fait apparaître de multiples sens potentiels, dont certains sont actualisés avec une plus ou moins grande insistance dans la presse française. Dans la partie précédente, l'étude des collocations du mot « Europe » dans le corpus IntUne en 2007 et 2009 complétée par celle des collocations du lemme « européen » a permis de faire émerger les différentes acceptions que les médias donnaient au concept analysé. Maintenant que ces différents sens potentiels ont été relevés, nous allons pouvoir détailler les différentes actualisations qui en sont faites dans les journaux du corpus IntUne. D'autre part, nous allons également pouvoir analyser d'éventuelles évolutions dans l'actualisation de ces sens potentiels qui forment le prototype sémantique du concept « Europe » entre 2007 et 2009. Cette seconde partie va donc s'attacher à l'étude de cette actualisation de chacun des sens du prototype de « Europe » présent dans la presse française, sens potentiel par sens potentiel, en montrant à chaque fois quelle dimension est accentuée, quels rapports ces sens peuvent entretenir les uns avec les autres, et quelles évolutions on peut percevoir entre 2007 et 2009. Les sens potentiels seront étudiés selon l'ordre suivant :

- Une masse terrestre, plus ou moins précisément délimitée.
- Une entité dont les pays sont regroupés dans des blocs géographiques.
- Une entité économique.
- Une entité partiellement composée d'une union politique : l'Union Européenne.
- L'Europe politique est une idéalisation.
- L'Europe politique doit défendre ses frontières.
- Une entité composée de plusieurs pays.
- Une entité sportive.

Le choix de cet ordre répond à diverses relations entre ces sens, les uns impliquant plus ou moins directement les autres, ces relations seront plus précisément explicitées dans chacun des paragraphes les concernant.

Les remarques présentées pour chacun de ces sens seront par ailleurs accompagnées de lignes de concordances venant les étayer, présentées dans un format Key Word In

4- « Europe ». *Quelle représentation dans la presse française ?*

Context, et dont la source dans le corpus sera précisée par un code. Ce code donnera tout d'abord les deux premières lettres du média dont provient le journal (fi pour le Figaro, ec pour les Échos, li pour libération, su pour Sud Ouest, et ou pour Ouest France), il sera ensuite suivi de la date du journal dont sera extrait l'occurrence selon un format JJMMAA. Par ailleurs, l'évolution entre les corpus IntUne 2007 et 2009 ne sera proposée que lorsque l'analyse portera sur des locutions impliquant un contexte précis (par exemple « Europe orientale » ou « Europe de la défense »), en effet, l'analyse exhaustive de toutes les occurrences du mot Europe dans le corpus 2009 n'a pas pu être menée, ce corpus ayant été compilé quelques mois seulement avant la rédaction de cette thèse.

IV.2.1- L'Europe, une masse terrestre plus ou moins précisément délimitée.

Les collocations du corpus IntUne pour le mot « Europe » laissent apparaître une entité physique plus ou moins précisément délimitée avec des collocations comme « en », « partout » ou « hors ». La colligation « en Europe » est notamment très présente avec 2 886 occurrences. Cette colligation indique généralement un procès qui se déroule au sein, ou sur la masse terrestre appelée « Europe ». Elle fait donc appel à une vision concrète, physique ou encore réelle de l'Europe. Elle est essentiellement présente dans les journaux Les Échos (1353 occurrences, soit près de 46 % du total), et Le Figaro (853 occurrences, soit 29%). L'Europe, masse terrestre et physique, apparaît ainsi fréquemment comme une entité concrète, dont on peut tracer le contour. Néanmoins, nous allons voir dans les paragraphes suivants que si cette entité peut avoir une certaine réalité, conduisant à des réalisations ou des constats réels et physiques, elle peut également parfois être questionnée à un niveau plus abstrait, voire même évoluer et changer de contours.

IV.2.1.1- Une masse terrestre que l'on peut appréhender concrètement.

Différents contextes peuvent alors être combinés à cette évocation d'une masse terrestre où des choses se passent, il peut s'agir d'économie, dans ce cas, on montre notamment la fortune plus ou moins heureuse de telle entreprise ou tel gouvernement.

De façon plus signifiante, il peut aussi s'agir uniquement d'une entité dotée d'une réalité physique. Dans Le Figaro, cette réalité est dépeinte essentiellement à travers de

4- « Europe ». Quelle représentation dans la presse française ?

nouveaux transports que l'on réalise pour s'y déplacer d'un point à un autre, voire échanger avec d'autres zones. Ce contexte est également très présent dans le journal Les Échos, il marque ainsi l'entité terrestre comme étant le nœud où se déroulent sans difficultés des échanges et où les flux de populations vont et viennent sans cesse. C'est le contexte le plus important en termes de fréquence dans le corpus.

pour des trajets au sein des régions Asie et	Europe	. Les nouveaux appareils non seulement	fi060207
est de devenir une porte d'entrée en	Europe	commente René Brun, directeur de l'aéroport	fi010307
SAS déservent en direct au départ de l'	Europe	la capitale de l'état de Washington, berceau	ec210207
Le canal Seine-Nord	Europe	, qui doit être mis en service en 2013 est ac	ec160307

Néanmoins, un second contexte apparaît aussi dans le journal les Échos, celui d'une masse terrestre avec une vie climatique ou géologique propre, non conditionnée par différentes infrastructures humaines, mais ayant une réalité physique, naturelle. Ce deuxième contexte est plus largement repris dans les autres journaux du corpus.

La tempête qui avait frappé le nord de l'	Europe	À la mi-janvier	ec160207
le sommet de la plus haute montagne d'	Europe	. L'entreprise fondée il y a cent ans en Alle	ec020407
de son volcan, l'un des plus actifs d'	Europe	. Une coulée de lave a déjà atteint la mer	li280207
empêche les perturbations de passer de l'	Europe	centrale à la France.	su260407
Fortes chutes de neige au nord de l'	Europe	. D'importantes chutes de neige ont	ou090207

IV.2.1.2- Une masse terrestre aux contours flous et problématiques.

La délimitation de l'Europe, masse terrestre, par telle ou telle limite naturelle est parfois évoquée, mais semble avant toute chose être floue ou problématique. Le journal Libération n'aborde pas les problèmes posés par ces frontières, en évoquant de façon imagée des frontières naturelles de façon peu précise. Sud Ouest précise pour sa part que la réalité géographique de l'Europe n'est pas acquise, ses limites doivent être questionnées. D'une part des extensions sont envisagées par un biais économique, avec par exemple une présence de plus en plus importante de l'Europe en Afrique du Nord, d'autre part, cette extension peut se faire par le biais de nouveaux moyens de transports, notamment le percement d'un tunnel sous le détroit du Bosphore. A travers ces différentes occurrences les plus typiques qui présentent l'Europe comme une masse terrestre, on s'aperçoit qu'il s'agit dans la presse française d'une masse aux contours dont la délimitation pose problème, et d'autre part, que ces contours sont mouvants, qu'ils peuvent évoluer avec la création de nouvelles annexes, par exemple.

4- « Europe ». Quelle représentation dans la presse française ?

connus, comme les frontières de l'	Europe	, à Istanbul ou Odessa, il échappe adroitement	fi030507
la manche (...).C'est un peu le cap Horn de l'	Europe	. « Jusqu'à samedi, on est resté en station à	li100207
, ex-commissaire européen. Pour lui,	Europe	Doit maintenant « trouver ses limites géogra	ou190307
Point de passage obligé vers cette	Europe	Des 27 résolutions tournée vers l'est dont le	ou240307
		centre de gravité s'est déplacé	
arc se voit déjà comme une extension de l'	Europe	. Pourtant, il lui reste des obstacles à franch	ou020407
tranche d'un tunnel ferroviaire qui reliera l'	Europe	à l'Asie en passant sous le détroit du Bosph	ou050407

En somme, l'Europe masse terrestre dans la presse française de 2007 est plus ou moins délimitée physiquement et concrètement, elle semble néanmoins être définie par les mers qui la bordent de façon un peu plus stable : le détroit du Bosphore en est une limite, quand bien même un tunnel est en passe de rendre cette frontière naturelle aisément franchissable. La Manche, plus étonnamment semble également la délimiter au Nord Ouest, selon Libération qui la compare au Cap Horn que l'on retrouve au Sud du continent américain. Quoiqu'il en soit, il apparaît avant toute chose que ces frontières posent question, et qu'elles sont mouvantes, le centre de gravité de cette masse terrestre pouvant en effet se déplacer.

IV.2.1.3- Le problème de l'énergie.

L'Europe, masse terrestre est aussi liée de près à la question énergétique dans le corpus 2007. Ce point semble poser également de nombreux problèmes avec les frontières de l'Europe, puisqu'à l'intérieur de celles-ci, très peu de matières premières seraient disponibles. La tendance est alors à l'élargissement de ces frontières dans la presse française.

Ce dernier contexte dans lequel le mot « Europe » apparaît en tant qu'une masse terrestre montre donc ce double problème d'une entité qui dispose de peu de ressources sur son propre sol, en comparaison de ses besoins. D'autre part, il peut être question de tensions avec la Russie qui est l'un des principaux fournisseurs de l'Europe du point de vue énergétique, cette idée se retrouve plus particulièrement dans les journaux du Figaro et des Échos. Il peut en effet arriver que des décisions diplomatiques soient fortement influencées par une contrainte énergétique. La plus grande partie des occurrences de ce contexte proviennent du Figaro ou des Échos, il peut enfin s'agir d'un mouvement de concentration parmi les grands groupes énergétiques européens, afin d'affronter au mieux les enjeux soulevés par les deux problèmes précédents ; la France, via EDF semblant bien tirer son

4- « Europe ». Quelle représentation dans la presse française ?

épinglé de ce mouvement. Cette dernière acception du sens potentiel « masse terrestre aux frontières imprécises », s'associe donc très clairement à l'entité politique devant se défendre, mais aussi à une entité économique.

gaz turkmènes expliquent que les États-Unis et l'	Europe	ménagent le successeur de Niazov	fi100207
mer Caspienne pourra bientôt être distribué en	Europe	. Cette nouvelle infrastructure faciliter	fi160307
Seule certitude : l'	Europe	de l'énergie est plus que jamais en ple	fi040407
Depuis plusieurs mois, l'	Europe	de l'énergie est dans la ligne de mire	fi050307
Trois géants de l'	Europe	de l'énergie en lice	fi020307
Dans une	Europe	de l'énergie en pleine recomposition	ec040407
ne peut pas être vendu directement en	Europe	, car il n'y a pas de gazoduc en directi	fi170207
Et faire comprendre à l'	Europe	de l'ouest combien elle est dépendante de ses hy	ou160307
prix du gaz livré au reste de l'	Europe	déjà confrontée à une pression à la ha	su130307

IV.2.1.4- Quelle évolution en 2009 ?

Ces différents contextes ont connu diverses fortunes dans leurs évolutions. En 2009, la question énergétique semble beaucoup moins importante, la locution « Europe de l'énergie » est par exemple cinq fois moins présente qu'en 2007, cependant, il y est toujours question d'éviter au maximum la Russie, notamment, comme fournisseur d'énergie. D'autre part, la question de la délimitation précise des frontières en Europe n'est plus présente. Si cette idée de frontières de l'Europe apparaît parfois, c'est alors tout simplement dans un contexte de protectionnisme économique dans le cadre de la crise internationale.

L'idée de masse terrestre est celle qui se transmet le mieux, il est toujours question de différents transports permettant de s'y déplacer, de plus, on trouve toujours la colligation « en Europe » avec une fréquence de 2256 occurrences qui se répartissent tout comme en 2007 essentiellement entre les Échos (1053 occurrences) et le Figaro (675 occurrences), cette colligation est toujours massivement liée à l'évocation de processus économiques qui se déroulent sur cette masse physique.

L'Europe, en tant que masse terrestre est donc une entité réellement présente dans la presse française à travers les deux corpus IntUne. Cependant, ce sens potentiel est activé en montrant que l'Europe y est pour le moins complexe à définir, elle ne semble pas présenter de réalité fiable que l'on puisse édicter dans l'absolu, mais plutôt un visage mouvant et dynamique, évoluant selon les points de vue (historiques, économiques, politiques). Nous

4- « Europe ». Quelle représentation dans la presse française ?

allons maintenant nous intéresser plus précisément à l'Europe en tant qu'entité divisée en sous ensembles le plus souvent directement inspirés par leur dénomination de réalités géographiques.

IV.2.2- L'Europe, une entité dont les pays sont regroupés dans des blocs.

L'analyse des lignes de collocations du mot « Europe » a également fait apparaître clairement que l'Europe semblait être une entité se divisant en plusieurs blocs. Ces blocs ont la plupart du temps des dénominations géographiques (« Europe de l'Est », « Europe de l'Ouest », « Europe occidentale », « Europe centrale », « Europe orientale », etc.), mais nous pouvons nous demander dans quelle mesure ces dénominations ne cachent pas une réalité où ces blocs présenteraient une cohésion plus politique ou économique que géographique. Nous allons dans cette partie ainsi voir quelles sont les différentes « Europes » que l'on peut trouver dans la presse française, et ce qui est associé à chacune de ces entités, en traitant chaque "Europe" de façon diachronique à travers l'étude de sa représentation en 2007 et en 2009. Les collocations font apparaître les « Europes » suivantes : orientale, continentale, centrale, occidentale, de l'Ouest, du Nord, du Sud, méridionale et de l'Est. Nous allons les étudier selon une progression géographique qui ira de l'ouest vers l'est, puis du nord au sud. Enfin, nous nous interrogerons sur les rapports possibles entre ces diverses entités.

IV.2.2.1- L'Europe occidentale.

Cette locution est présente à la fois dans le corpus IntUne 2007 et 2009, elle connaît cependant un emploi plus important en 2009 avec 60 occurrences contre 42 en 2007. Nous allons maintenant proposer une étude de cette locution en 2007.

IV.2.2.1.1- L'Europe occidentale en 2007.

L'Europe occidentale apparaît donc à 42 reprises dans le corpus IntUne français 2007. Ces 42 occurrences sont majoritairement présentes dans le journal Les Échos (20 occurrences), Le Figaro (11 occurrences) et Sud Ouest (6 occurrences). Elles présentent avant tout une Europe occidentale qui a une forte homogénéité au niveau économique. Les résultats économiques de cette Europe occidentale sont alors plutôt mauvais :

4- « Europe ». Quelle représentation dans la presse française ?

La forte localisation de Michelin en	Europe occidentale	et ses effectifs excessifs lui font subir	ec050207
pprimé quelques 7000 emplois (...) en	Europe occidentale	. Et si l'on en croit la firme d'assuranc	su300407
de nouveaux débouchés hors de l'	Europe occidentale	où ils risquent d'être extrêmement li	su210407
la main d'œuvre est moins chère qu'en	Europe occidentale	, la direction de plastimmo met en	ou040407
richesses demeurent stables en	Europe occidentale	et au Japon.	li030407
La direction admet des sur-effectifs en	Europe occidentale	. Or, le président de PSA, Christian	fi230407
le secteur bancaire en	Europe occidentale	est pris d'une fébrilité inédite	fi230407
Delphi quittera-t-il l'	Europe occidentale	? C'est ce que craint un syndicat eu	fi160307
rix reste inférieur à ses concurrentes d'	Europe occidentale	. Il augmentera bien de quelques ce	fi260207
Progressé « que » de 7% en	Europe occidentale	. Samsung cède du terrain	ec050307

La plupart des occurrences qui montrent l'Europe occidentale en perte de vitesse, ou avec des résultats économiques peu satisfaisants est la conséquence d'une forte concurrence, ou d'un développement plus important d'une Europe de l'est. Cependant, quelques occurrences, beaucoup plus rares donnent de l'Europe occidentale une image moins négative :

croissance, donc, à côté d'une	Europe occidentale	où Fagor mise sur des innovations	ec090307
Falcons se vendent de plus en plus en	Europe occidentale	et orientale (Russie, Ukraine, etc)	su160207

L'Europe occidentale présente ainsi, essentiellement dans le corpus 2007, une homogénéité au niveau de la médiocrité de ses résultats économiques, mais elle semble également présenter un homogénéité culturelle, à travers une histoire commune, qu'il s'agisse de religion, ou de la domination de la France sur cette Europe à une époque de son histoire, voire de l'ambition ancienne de créer une alliance des pays la composant :

une grande alliance des pays d'	Europe occidentale	. Une idée qui l'amène à jeter en 1941	ec190307
la peuplée de toutes les nation d'	Europe occidentale	et sur l'armée la plus puissante du monde	fi130207
L'Eglise la plus puissante d'	Europe occidentale	. Mgr Bagnasco est, tout comme le card	fi080307

Cette Europe semble présenter également une forte homogénéité au niveau énergétique, les mêmes questions se posent dans ce domaine à l'ensemble des pays qui la composent, et il est à noter que dans ce contexte, tout comme dans le cadre économique, les problèmes qu'elle rencontre viennent essentiellement de l'est. La principale question pour cette entité est alors d'obtenir à un prix non « indu » (créé artificiellement par des ententes de « cartel » selon les occurrences trouvées dans la presse) les matières premières dont elle a besoin et qui viennent de ses voisins d'Europe orientale.

0km de long, vise à transporter vers l'	Europe occidentale	le pétrole de la Caspienne en contou	ec050407
---	--------------------	--------------------------------------	----------

4- « Europe ». Quelle représentation dans la presse française ?

lier dans 5 ans la Turquie aux pays d' Europe occidentale , en passant par les Balkans. Selon la ec060407

En 2007, l'Europe occidentale est donc un ensemble qui présente une forte homogénéité au niveau de ses problèmes économiques et énergétiques dus pour partie aux pays de l'est, et au niveau culturel. On peut maintenant étudier quelle est l'activation de cette Europe dans le corpus 2009.

IV.2.2.1.2- L'Europe occidentale en 2009.

Dans le second corpus IntUne, l'Europe occidentale est davantage présente, de 42 occurrences, on passe à 60 occurrences, alors que le mot Europe est moins présent dans ce corpus. Cette Europe est la plus présente dans Les Échos avec 27 occurrences, puis Le Figaro pour 18 mentions, et enfin dans Libération avec 9 occurrences. Chacun de deux régionaux fait figurer cette Europe à trois reprises. En 2009, l'Europe occidentale est donc davantage présente, notamment au plan national dans la presse française. L'image de cette Europe est toujours celle d'un ensemble qui présente une forte homogénéité, essentiellement au niveau économique. Dans ce cas, ce sont alors les graves difficultés de cet ensemble qui sont évoquées, mais il est à noter que ces difficultés trouvent la plupart de leurs sources dans un autre ensemble situé plus à l'Est, qui handicape l'Europe occidentale du fait de ses très mauvais résultats. Quand ce n'est pas par de mauvais résultats économiques, c'est par une concurrence déloyale de dumping social que cet ensemble handicape son voisin :

Face à de tels cadeaux, les pays d' Europe occidentale ont du mal à lutter à armes égales.	ec120209
avive les inquiétudes sur les banques d' Europe occidentale exposées aux économies d'Europe centrale, et de l'est, en plein marasme	ec180209
trogradation de notation de banques d' Europe occidentale du fait de leur exposition aux pays d	ec020309
modèles de petites citadines vendues en Europe occidentale à prix cassés proviennent d'usines s	ec040309
concernant la solidité des banques d' Europe occidentale et de leurs filiales de l'est. » Les exp	fi190209
aujourd'hui, dix ou douze banques d' Europe occidentale sont très fortement engagées en eur	fi270209

L'Europe occidentale semble néanmoins être davantage marquée par une bonne réussite économique, relativement à sa voisine centrale ou orientale, elle demeure surtout un marché important, contrairement aux marchés de l'est qui s'effondrent.

projets en blanc. Les pays d' Europe occidentale demeurent attractifs et l'activité pour	ec120209
2000 et 2008 de 36 à 39% des ventes en Europe occidentale . C'est la raison pour laquelle Christ	fi280309

Le fait que cette Europe soit notée comme clairement handicapée par sa voisine de l'est dans la presse française fait naître certaines tensions palpables et explicites entre les

4- « Europe ». Quelle représentation dans la presse française ?

deux Europe : l'une imputant à l'autre ses mauvais résultats économiques, l'autre lui reprochant de ne pas lui venir en aide. Ces tensions semblent poser problème au sein d'une troisième Europe présentant la synthèse de ces deux dernières, l'Union Européenne :

ars, en guise de coup de semonce à l'	Europe occidentale	. La Commission tente de contenir le	ec230209
nion en 2004 et 2007. Elle a appelé l'	Europe occidentale	à « protéger » le front de l'est et l'a	ec230209
ment public est important. Les pays d'	Europe occidentale	seront plus utiles à leurs voisins de l'	fi270209
ppelle à la solidarité économique de l'	Europe occidentale	à l'égard des pays de l'est et à son pa	fi040309

D'autre part, outre son homogénéité économique, cette Europe occidentale présente de nombreux facteurs de cohésion intérieure, qu'il s'agisse d'une représentation commune dans les plus hautes instances institutionnelles mondiales, de comportements sociaux semblables de ses populations, ou surtout d'une histoire commune de défense contre l'extérieur, avec le partage d'une même religion, et d'un patrimoine artistique commun. Elle est aussi présentée comme une entité réelle, concrète, servant d'aune physique à la lumière de laquelle on peut mieux appréhender la réalité physique d'une autre masse terrestre :

la Turquie qui occupe (.)un (.) sièges de membre			
non permanent, pour le groupe des « Etats d'	Europe occidentale	et autres états », s'est vu confier le	ec090409
sociales profondes des populations,	Europe occidentale	, Amérique, Alaska, Nouvelle-Guin	li260309
nion soviétique et ses alliés d'envahir l'	Europe occidentale	. La France est l'un des douze pays f	fi120309
centes d'André XXIII lorsqu'il dit que l'	Europe occidentale	n'a pas à garder les yeux rivés sur la	li190309
l' homme à taille humaine connues en	Europe occidentale	, exposés aujourd'hui au musée Fena	li250409
vouloir dépecer ce pays grand comme	Europe occidentale	. Le mois dernier, dans un entretien	li260309

Pour résumer, l'Europe occidentale présente toujours en 2009 l'image d'un ensemble animé par une forte cohésion renforcée depuis 2007. La presse se réfère pour arguer de cette cohésion à un passé commun de défense contre le bloc soviétique, ou encore à des racines religieuses et artistiques communes, c'est aussi une zone qui présente une réalité physique concrète, on peut la mesurer et s'en servir comme d'une aune pour mesurer d'autres masses terrestre, en cela, l'Europe occidentale n'est plus simplement abstraite. Cette évolution montre une Europe occidentale plus unie peut être en partie due à la crise économique qui sévit alors, cette partie de l'Europe cherche avant toute chose à se protéger, et notamment à se protéger de risques qui pourraient venir de sa voisine de l'est. L'Europe occidentale est donc une entité totalement différenciée de l'Europe orientale, lorsqu'elle est présente dans le corpus IntUne, et cette différenciation tend à croître entre 2007 et 2009.

4- « Europe ». Quelle représentation dans la presse française ?

IV.2.2.2- L'Europe de l'Ouest.

Cette deuxième « Europe » apparaît à 129 reprises dans le corpus IntUne en 2007, essentiellement citées par Les Échos (72 occurrences) et Le Figaro (33 occurrences). Les autres journaux du corpus qui citent cette Europe sont Ouest France (15 occurrences), Sud Ouest (6 occurrences) et enfin Libération (3 occurrences). Nous allons maintenant analyser les lignes de concordances de cette expression afin de voir si elle dispose d'une homogénéité, tout comme l'Europe occidentale ou non. Nous commencerons par le corpus 2007 puis analyserons le corpus 2009, afin de voir s'il y a éventuellement eu une évolution dans la représentation de cette entité pour la presse française.

IV.2.2.2.1- L'Europe de l'Ouest en 2007.

Le premier contexte en termes d'importance est cette fois encore économique. La majorité des occurrences dépeint certes un portrait flatteur au niveau économique de ce sous-ensemble, d'autant plus lorsqu'il s'agit de valeurs boursières régulièrement en hausse. Néanmoins, de nombreuses difficultés apparaissent dans divers domaines. Ainsi, les marchés contribuant à la croissance de l'Europe de l'Ouest ne sont plus intérieurs, et ses emplois se détruisent rapidement. L'ensemble peut aussi être politique, dans ce cas, il s'agit essentiellement de parler de liberté, de droits de l'homme ou de démocratie, ou encore de montrer que la zone « Europe de l'ouest » semble être opposée, ou tout du moins agacer une autre entité fréquemment présente : « Moscou ». L'immigration est également présente lorsqu'il s'agit de cette Europe, cette zone recevant alors de nombreux immigrants avec plus ou moins de difficultés. Elle s'oppose parfois à l'Europe de l'Est dans ce mouvement migratoire. L'énergie pose le problème d'un lien sensible entre cet ensemble et son voisin plus à l'est, en effet, il s'agit le plus souvent d'un manque de ressources au sein de la zone « Europe de l'Ouest » qui dépend alors notamment de l' « Europe de l'Est » :

la croissance se situe bel et bien hors d' Europe de l'Ouest .	fi140407
mais les salariés d' Europe de l'Ouest n'en voient pas les retombées.	ec050207
117 000 emplois qui ont disparu en Europe de l'Ouest entre 2000 et 2006	su300407
reculent sur le marché domestique de l' Europe de l'Ouest et ne sont pas compensées par le	ou240407
libertés dont jouissaient déjà d'autres pays d' Europe de l'Ouest : nous ne pouvions lire les ouvra	su230307
Le mot « démocratie » évoque, en Europe de l'Ouest , les droits de l'homme ; chez les	li170207
Ses voyages à Washington et en Europe de l'Ouest , aussi nombreux que ses visites en Russie ont pu agacer Moscou	fi040407

4- « Europe ». Quelle représentation dans la presse française ?

à l'immigration illégale, les pays d' Europe de l'Ouest	utilisent en effet la régularisation	fi200307
généralisation de ses pratiques, l'arrivée en Europe de l'Ouest	d'une vague d'immigration famili	fi200307
pays de transit pour le gaz russe exporté en Europe de l'Ouest	.	fi040407
Et de faire comprendre à l' Europe de l'Ouest	combien elle est dépendante de	ou160307

L'Europe de l'Ouest présente donc ici une homogénéité, étant essentiellement présente dans un contexte économique, on voit qu'elle subit des difficultés dans ce domaine, notamment à travers la destruction de nombreux emplois. Cependant, elle jouit d'une image plutôt favorable dans d'autres domaines comme la politique, où on l'associe à des concepts positifs comme celui de liberté. Cette Europe s'oppose à plusieurs reprises à son pendant géographique « l'Europe de l'est », que ce soit au niveau économique, diplomatique ou énergétique.

IV.2.2.2.2- L'Europe de l'Ouest en 2009

L'Europe de l'Ouest apparaît en 92 reprises dans le corpus IntUne 2009. La plupart des citations proviennent du journal Les Échos (54 occurrences) et Le Figaro (24 occurrences). Dans une vaste majorité, il s'agit en 2009 d'une zone économique, mais à la différence de ce que l'on pouvait voir en 2007, cette zone bien résister à la crise. De plus, la lecture des lignes de concordances montre aussi que la crise devrait se terminer dans environ un an, et que les entreprises d'Europe de l'Ouest doivent se préparer à cette reprise pour gagner en performance, la perspective adoptée par les journaux français est alors positive. Enfin, de nombreuses occurrences montrent que l'Europe de l'Est demande le soutien de cette zone, impliquant ainsi que celle-ci ait davantage de possibilités. C'est aussi parfois une zone qui a une histoire, cette histoire est alors faite d'oppositions avec une « Europe de l'Est » et de savoir. L'image de l'Europe de l'Ouest est en conséquence dans ce contexte plutôt positive.

limiter les risques en réaffectant à l' Europe de l'ouest	, plus sûre et plus rémunératrice	ec270209
Ses ventes ont augmenté de 8,3% en Europe de l'ouest	en février alors que le marché	fi160309
Robert Zoellick appelle l' Europe de l'ouest	à aider l'est parce qu'une nouvelle divisi	fi250209
Est devenu du XVIe au XVIIIe en Europe de l'ouest	, une « science du corps ». Elle a engend	li170309
Nombre de mes amis se rendaient en Europe de l'ouest	, moi je n'en avais pas envie. Ces voyages étaient comme des grâces accordées par l'État communiste.	li120309

L'Europe de l'Ouest connaît donc en 2009 une évolution positive au niveau économique par rapport à 2007. Celle-ci résiste plutôt bien à la crise, d'autant plus si on la

4- « Europe ». Quelle représentation dans la presse française ?

compare à sa voisine qui devient même dangereuse dont les difficultés croissent sans cesse. D'autre part, c'est un ensemble qui a une histoire commune, faite de savoir, de science, mais aussi d'un rejet passé de l'idéologie communiste. Ce bloc semble donc ne pas réellement disposer de points communs avec l'Europe de l'Est, tel qu'il est présenté dans la presse française en 2007 et 2009. L'Europe continentale va maintenant être analysée, toujours selon l'idée d'une progression géographique allant de l'ouest vers l'est.

IV.2.2.3- L'Europe continentale.

Cette acception qui se base sur une dénomination géographique désigne l'Europe dont on exclue les îles britanniques et l'Irlande. La locution apparaît à 54 reprises dans le corpus français IntUne 2007, contre 31 en 2009.

IV.2.2.3.1- l'Europe continentale en 2007

En 2007, cette locution est présente dans une vaste majorité des cas dans le journal Les Échos (45 des 54 occurrences), on peut en déduire que c'est une expression essentiellement utilisée par les élites. De plus, lorsqu'on parle d'Europe continentale, il s'agit en conséquence le plus souvent d'un contexte économique. Les autres occurrences relatives au contexte économique donnent des résultats particuliers pour cette zone. Il est intéressant de noter que dans ce cas, si les résultats économiques sont généralement bons, ils le sont beaucoup moins en comparaison de l'Europe non continentale, c'est-à-dire le Royaume-Uni et l'Irlande. Enfin deux occurrences se réfèrent à un comportement culturel proche pour les pays de cette zone.

britanniques du secteur à celles d'	Europe continentale	, dont les valorisations sont tendues	ec070307
La France a été le pays le plus actif en	Europe continentale	en 2006 avec 215 LBO réalisés	ec080307
robuste sur deux marchés clefs, l'	Europe continentale	(+7%) et l'Amérique du Nord	ec140307
producteur irlandais (...) prend pied en	Europe continentale	.	ec190407
son milieu, défendue par la pensée d'	Europe continentale	, et celle d'un homme prédéterminé	li070407

L'étude nous renvoie donc l'image d'une Europe dont le comportement économique semble plutôt bon. Toutefois, une ambiguïté est révélée par la presse française à l'égard de cet ensemble, en effet, il semble effectuer de très nombreuses opérations de LBO (Leverage Buy Out), or, ce type de transactions peut être risqué pour le marché qui les accueille, puisqu'il s'agit de recevoir des capitaux étrangers : cela montre l'attractivité de la zone en

4- « Europe ». Quelle représentation dans la presse française ?

question, mais cela montre aussi que cette zone est peu à peu mise sous un contrôle étranger. Cette problématique n'existe pas pour les îles britanniques. La différenciation culturelle entre les deux zones (îlienne et continentale), ne semble pas avoir une importance déterminante, puisqu'elle n'occupe qu'une seule occurrence.

IV.2.2.3.2- L'Europe continentale en 2009.

Dans le corpus IntUne 2009, cet ensemble est moins présent avec 31 occurrences, mais son usage s'est davantage réparti entre les différents journaux : Les Échos demeure le principal journal le mentionnant avec 17 occurrences, suivi de près par Le Figaro avec 13 occurrences. On peut donc en déduire que cette expression passe davantage dans le langage des masses en 2009. D'autre part, l'évolution de la répartition des occurrences de cet ensemble s'accompagne d'une évolution de son image qui s'améliore sensiblement, et d'une diminution de la présence du contexte économique. En outre, lorsqu'il s'agit de performance économique, il est explicitement dit que la zone est plus performante que les îles britanniques. Il est aussi question de différences de fonctionnement majeures entre une Europe continentale et les îles britanniques au niveau de l'Union Européenne ou de la façon de régler la crise économique sévissant alors au niveau mondial.

celière, on estime que la plupart des pays de l' Europe continentale , notamment la France et l'Allemagne, fi210209
sont d'accord (..). La difficulté sera de
rallier (..) le premier ministre britanniq

e-Uni (..), leurs performances sont aujourd'hui
bien plus mauvaises que celles de l' Europe continentale (..) outre leur spécificité géographique fi020309

FMI, l'ont répété à l'envi: la relance de l' Europe continentale est insuffisante. fi300309

Il est de l'intérêt de l' Europe continentale tout autant que de la place de Londres ec090409

Plus qu'un simple sous ensemble désigné pour des raisons de commodités géographiques, l'Europe continentale semble dans la presse français réellement bénéficier d'une identité autonome : comportements semblables, décisions communes. Cette identité s'affirme de façon notable dans le corpus 2009, où elle n'est jamais aussi forte que lorsqu'elle s'oppose au Royaume-Uni ou aux États-Unis. Ce sont les journaux lus par les masses qui mettent le plus en avant l'affirmation de l'identité continentale, alors que Les Échos rapproche ce bloc de celui des îles britanniques en 2009.

4- « Europe ». Quelle représentation dans la presse française ?

IV.2.2.4- L'Europe centrale.

Nous venons de voir que l'Europe continentale voyait une évolution de sa représentation dans la presse française entre 2007 et 2009, qui coïncidait notamment avec une évolution des journaux dans lesquels les occurrences la mentionnant apparaissaient. L'ensemble « Europe centrale » semble rencontrer le même phénomène dans le corpus IntUne. Sa présence totale dans ce corpus ainsi que sa répartition dans les différents journaux peuvent se résumer comme suit :

2007		2009	
Les Échos	87	Les Échos	97
Le Figaro	35	Le Figaro	66
Libération	7	Libération	18
Sud Ouest	10	Sud Ouest	5
Ouest France	7	Ouest France	4
Total	146	Total	190

(Figure 14, répartition de « Europe centrale » dans le corpus IntUne)

On constate une progression de la représentation de l'ensemble en 2009, essentiellement dans les deux journaux nationaux. Une question est alors de savoir si cette progression se ressentira dans la représentation que fait la presse française de l'Europe centrale.

IV.2.2.4.1- L'Europe centrale en 2007.

Différents contextes peuvent être associés à cette « Europe centrale » qui nous permettent de préciser cette entité. Deux contextes émergent essentiellement : celui d'une économie dynamique et celui de la défense. Une comparaison utilisée pour décrire le comportement économique de cette zone est notamment éloquent, puisque la construction des usines y est comparée à la prolifération de champignons. En ce qui concerne le contexte de défense, c'est la crise entre « Moscou » et les pays du « centre » de l'Europe qui est mentionnée à travers l'installation par les États-Unis d'un bouclier antimissiles.

Cette zone a également dans le corpus 2007 une histoire violente commune, entre révolutions populaires (considérées comme étant à la source de la création d'un esprit européen) et persécutions des juifs. L'environnement, tout comme les droits de l'homme

4- « Europe ». Quelle représentation dans la presse française ?

semblent y être bafoués au nom d'un développement économique rapide, ce qui est décrié dans la presse française. Ainsi, la justice locale semble peu amène à réprimer les instigateurs d'une réelle traite humaine, tandis que dans le même temps, elle impose une forte ségrégation aux roms. Ses ressources énergétiques sont par ailleurs présentées comme importantes, mais l'instabilité politique des gouvernements de cette zone fait remettre en cause leur accès. Enfin, la presse met en relief un certain euro-scepticisme de ces pays. On relève par ailleurs deux constructions très fréquentes pour cet ensemble qui sont « pays d'Europe centrale » et « Europe centrale et orientale », la première apparaît à 21 reprises, la seconde à 39.

Usines ont poussé comme des champignons dans les	pays d' Europe centrale et orientale.	su030407
a été le vent de changement qui a balayé l'	Europe centrale et orientale dans les années 1980	ou240307
forêts les plus riches en plantes et animaux d'	Europe centrale . Si la Pologne n'obéit pas	fi010307
Institutions ont moins la cote dans les pays d'	Europe centrale excepté pour la Slovénie	fi260207
L'activité de la banque en	Europe centrale a toutefois fortement progressé	ec230207
ces étrangères, 29,8% Étaient originaires des pays d'	Europe centrale et orientale (...). Ainsi, la prostitutio	su130307

On relève par ailleurs une construction très fréquente pour cet ensemble qui est « Europe centrale et orientale », en 39 occurrences. Ces formules montrent explicitement que cet ensemble est très fortement associé à l'Europe orientale, en ayant un comportement similaire à ce deuxième groupe.

Pour préciser les pays qui composent cette zone, on peut relever toutes les occurrences qui listent différents états ou nom de lieux en faisant partie. L'image de l'entité reste malgré tout très floue. Pour reprendre le fonctionnement des prototypes selon E. Rosch, on pourrait dire que la Pologne en est le représentant prototypique. Selon une classification plus proche du fonctionnement par « air de ressemblance » de Wittgenstein, on pourrait y inclure la Slovaquie, la Pologne, les pays Baltes, la Roumanie, l' « ex-Yougoslavie » (=Croatie, Serbie, Montenegro, Bosnie Herzégovine, Macédoine), la Hongrie, la Bulgarie, la Moldavie et l'Albanie. L'Autriche pose problème, puisque dans la première occurrence, ici, elle est exclue de cet ensemble, alors que dans la deuxième, au contraire, elle y est incluse :

l'Irlande, la Suisse, la Grèce, l'Autriche et l'	Europe centrale	fi130207
Les pays d' Europe centrale , l'Autriche et la Slovaquie,	notamment	fi190407
ologne s'est prononcée en faveur de l'installation en	Europe centrale d'un bouclier américain antimissiles	li160207

4- « Europe ». Quelle représentation dans la presse française ?

Pays d' Europe centrale : Pologne, pays Baltes, ex-Yougoslavi	li030307
Varsovie en est d'ailleurs une variante unique en Europe centrale .	li280207
Hongrie a réussi à enrayer (..), fait rare en Europe centrale .	li210307
d' Europe centrale et orientale ont indiqué (..) à Budapest	fi030407
en Europe centrale et orientale (Bulgarie, Roumanie, Moldavie, Albanie)	su130307

Pour résumer la présentation de cet ensemble dans la presse française en 2007, il s'agit avant tout d'un lieu où la situation économique est excellente et dynamique, dotée de nombreuses matières premières, mais où les droits de l'homme et la législation environnementale, par exemple ne sont pas respectés. C'est également une zone qui s'est construite comme européenne lorsqu'elle a refusé l'idéologie communiste, mais qui marque un sentiment euro-sceptique.

IV.2.2.4.2- L'Europe centrale en 2009.

Dans le corpus IntUne 2009, il y a une progression au niveau de la présence de ce terme (cf. figure 14). Mais une différence de taille intervient, en effet, si dans le corpus 2007, on notait les performances exceptionnelles des économies de l'Europe centrale, leurs nombreux recrutements dans le domaine automobile, notamment, il est davantage question en 2009 d'un possible effondrement de leur système financier, et de leurs grandes difficultés pouvant créer des problèmes aux banques occidentales qui s'y seraient risquées. Le second contexte qui apparaît était également déjà présent dans le corpus 2007, il s'agit de la défense et du problème posé par la volonté d'installer un bouclier antimissile de la part des américains.

Cependant, les positions semblent moins radicales qu'en 2007, dans ce contexte défensif même si un conflit s'est produit en Géorgie. Le contexte historique est toujours présent et traite essentiellement de l'histoire violente et communiste de cette partie de l'Europe. Deux contextes qui n'étaient pas présents en 2007 apparaissent en 2009, il s'agit de la politique et surtout de la question de l'intégration, de la réelle union en Union européenne. Au niveau de la politique, il s'agit de l'élection d'un nouveau gouvernement en Slovaquie et en Pologne. Concernant la question quant à la supposée union ou non au sein de l'Union Européenne, il est noté que si l' « Union n'est pas constituée de blocs », il est néanmoins important voire « urgent » d'y donner davantage de cohésion :

4- « Europe ». Quelle représentation dans la presse française ?

Les économies d' Europe centrale vacillent et menacent d'entraîner dans leur fi240209
 L' Europe centrale n'est-elle pas trop dépendant de l'industrie fi270209
 en Géorgie -menaces de types classique en Europe centrale et orientale- fi180209
 revenir sur le déploiement d'un bouclier en Europe centrale . fi040309
 mer de cette triste et malheureuse plaine d' Europe centrale , parcourue depuis toujours par des li120309
 armées
 ni dans celui de l'UE en tant que telle, que l' Europe centrale et l'Europe de l'Est se sentent abandonn fi040309
 pas constituée de blocs. Mais certains pays d' Europe centrale et orientale ont de sérieuses difficultés dont li190309

Cette Europe centrale n'est pas une simple entité dotée d'une réalité géographique, il s'agit d'un ensemble partageant de nombreux comportements, de nombreuses racines communes avec une identité proche. Il est à noter alors qu'entre 2007 et 2009, la vision qui est donnée par la presse française de cet ensemble est sensiblement différente : de pays au développement économique impressionnant, au réel dynamisme et décidant de se passer de l'influence politique de la Russie, on passe à un statut de pays portant un risque fort au niveau économique et financier, dont l'histoire marque une proximité continue avec la Russie, et qui pourraient ne pas être aussi « solides » dans leur identité européenne que les pays d'Europe occidentale. Cette évolution coïncide comme pour l'Europe occidentale avec un glissement de l'utilisation de l'expression « Europe centrale » vers les journaux nationaux les plus lus du corpus, notamment par les masses. L'Europe centrale tend à s'éloigner davantage de l'Europe occidentale dans le corpus 2009.

IV.2.2.5- L'Europe de l'Est.

Ce groupe apparaît dans les deux corpus IntUne en de nombreuses occurrences, il est ainsi présent à 198 reprises en 2007 contre 321 en 2009. On note une progression de sa présence d'une année à l'autre, toutefois, contrairement aux blocs précédents, ici, la progression ne correspond pas à un changement sensible de sa répartition dans les journaux du corpus :

2007		2009	
Les Échos	95	Les Échos	188
Le Figaro	37	Le Figaro	68
Libération	15	Libération	27
Ouest France	24	Ouest France	14
Sud Ouest	24	Sud Ouest	24

4- « Europe ». Quelle représentation dans la presse française ?

TF1	2	TF1	0
France 3	1	France 3	0
Total	198	Total	321

(Figure 15, répartition de « Europe de l'Est » dans le corpus IntUne)

IV.2.2.5.1- L'Europe de l'Est en 2007.

Considérant l'Europe de l'Est, une analyse des lignes de concordances montre que cette entité apparaît en premier lieu dans un contexte économique, elle prend alors la forme d'un bloc économique aux intérêts communs. Cette économie est très largement dynamique et porteuse, que ce soit au niveau concurrentiel où elle est très bien située par rapport à d'autres pays (Europe occidentale, notamment) ou en termes de marchés qui sont très porteurs. Dans une proportion moindre, elle apparaît de façon moins positive, il y est question d'aide au développement qui lui est fournie par l'Europe occidentale au même titre que l'Afrique, de son développement qui n'est pas encore effectué, ou du fait que finalement, cette zone n'est pas un calcul si rentable en terme de délocalisation.

D'autre part, c'est une zone qui est fortement marquée par les trafics en tous genres : ventes d'armes illégales, marché de la drogue, de la prostitution, trafic alimentaire constituent une partie des difficultés auxquelles elle est confrontée. Cela concerne plus de 10 % des occurrences d'Europe de l'Est dans le corpus IntUne 2007, hors Les Échos. L'Europe de l'Est pose également le problème de l'immigration dans la presse française : la zone est alors vue à la fois comme très pauvre, puisqu'elle est notamment comparée à l'Afrique, mais aussi comme dangereuse, disposant d'une main d'œuvre qualifiée concurrençant sérieusement la main d'œuvre occidentale.. Au niveau culturel, cet ensemble est idéal, il génère une grande richesse artistique parfois même qualifiée de « magique ». Enfin, au niveau politique, l'influence communiste semble encore peser énormément sur cette zone, et poser de fait quelques problèmes avec l'Europe occidentale (représentée par l'Union Européenne) :

raittance par la Florentaise à des entreprises d' Europe de l'est	ou050507
expédie du matériel (..) en Afrique, en Asie, en Europe de l'est	su170407
gique de renvoyer les matériels en Asie ou en Europe de l'est est une vraie catastrophe	ou170307
Ces jeunes femmes venues d' Europe de l'est se retrouvent prostituées dans toute l'Eur	su170307
Nous venons d' Europe de l'est Et la vie n'est pas facile pour nous	su170207
Immigration : une vague venue d' Europe de l'est	fi260207

4- « Europe ». Quelle représentation dans la presse française ?

Vive les plombiers-zingueurs de l' Europe de l'est ! ».	li170407
De Bretagne, s'Irlande ou d' Europe de l'est . Roland Conq fait rimer virtuosité et mu	ou130307
En revanche, Lénine, l' Europe de l'est Chers à Martine, la présidentielle, l'actua	ou100407
Le Kremlin continue de considérer l' Europe de l'est comme une zone d'influence historique de	li270407
Le fabricant Suédois a fortement progressé en Europe de l'est , avec une hausse de plus de 41% des ven	ec050207

L'Europe de l'Est, plus qu'une commodité linguistique pour désigner un lieu sans identité particulière est donc une locution qui nous renvoie à une entité qui a une existence propre, entrant parfois en conflit avec d'autres sous ensembles de l'Europe (occidentaux, notamment). Cette entité bénéficie d'une image très favorable lorsqu'on l'aborde sous un jour économique dans le corpus IntUne 2007, ou dans un domaine culturel. Néanmoins, elle semble poser de nombreux problèmes, que ce soit au niveau d'une forte migration de ses populations, des nombreux trafics qui s'y déploient, ou d'une menace politique encore présente de façon implicite dans les médias français en 2007.

IV.2.2.5.2- L'Europe de l'Est en 2009

La locution « Europe de l'Est » est plus présente dans le corpus 2009 avec 321 occurrences. Les lignes de concordances montrent dans la presse française alors un bloc qui a toujours une forte réalité économique, avec la nuance qu'il s'agit non plus en 2009 d'un climat euphorique, mais plutôt d'un secteur morose. L'économie des pays de ce bloc est ainsi montrée en crise dans différents cas de figure. Cela peut être leurs entreprises qui chutent et menacent par là même les entreprises occidentales. Il peut aussi s'agir de forts risques pour leur système financier qui est menacé à court terme, notamment par les pressions répétées des pays de cette zone sur leurs monnaies pour gagner en compétitivité. Des aides sont alors requises par ces pays vis-à-vis de l'Union Européenne, un effort de solidarité est demandé, en particulier auprès des pays de l'Europe de l'Ouest, qui est accusée par sa voisine de l'Est de repli protectionniste. Un autre risque se profile de fait, celui d'un éclatement de l'Union Européenne en plusieurs blocs : la solidarité allant de l'un vers l'autre. Cette solution apparaît dans la presse français comme ne pouvant pas fonctionner.

La dimension historique et culturelle de ce bloc est toujours présente, et recouvre les mêmes réalités qu'en 2007 : grande richesse culturelle, passé violent fait de communisme et de persécutions des juifs. La dimension sociétale est elle aussi cohérente avec la représentation de 2007, les trafics posent toujours des problèmes importants, et ils sont

4- « Europe ». Quelle représentation dans la presse française ?

renforcés dans ce corpus avec une présence de plus en plus importante des mafias. Une nouvelle dimension est apportée en 2009 au portrait de l'Europe de l'Est via l'émergence du contexte de la défense, la possible montée des affrontements au sein de cette zone est alors mise en relief par la presse française. Enfin, les journaux du corpus notent que l'intégration de cet ensemble à la zone euro semble poser davantage de difficultés que prévues, nonobstant, l'Union Européenne ne stoppe pas ce processus d'intégration.

presque deux ans a fini par contaminer l' Europe de l'est , passée en un trimestre d'une croissance	ec050209
oumanie n'est toutefois pas le seul pays d' Europe de l'est à souffrir de la sorte, loin de là	ec130209
Les devises d' Europe de l'est sont prises dans une spirale baissière	ec180209
exposition des banques occidentales à l' Europe de l'est , en plein marasme économique. La monnaie	ec180209
lisés par des bandes organisées venues d' Europe de l'est .	fi110409
Extermination des juifs et des Tziganes en Europe de l'est .	fi160409
ouest est liée à la fin du communisme en Europe de l'est . Ils travaillaient dans les usines et les kolkhoze	su270409
L'Italie angoissée par l' Europe de l'est .	fi050309
En outre, les nouveaux états membres d' Europe de l'est , qui n'ont ni la force économique, ni la stab	fi050309
quelques semaines, plusieurs pays d' Europe de l'est , confrontés à une tourmente économique qui	li020309
l' Europe de l'est ne forme pas un bloc, mais a des problèmes sp	li020309
En accueillant quatre compagnies d' Europe de l'est	li250309
En Europe de l'est , la mafia calabraise faire la loi. Lors de l'ouver	fi230309
des centaines de femmes d' Europe de l'est et les a forcées à se prostituer.	fi100309
affaire du bouclier antimissile américain en Europe de l'est a montré qu'on n'était pas très loin d'une repris	su070209
pour faciliter l'intégration de pays de l' Europe de l'est .	fi050309
fixée : accélérer l'intégration politique de l' Europe de l'est et des pays Baltes.	fi250209

IV.2.2.6- L'Europe orientale.

Cet ensemble est nettement moins présent dans les journaux du corpus IntUne, on le retrouve en 24 occurrences dans le corpus 2007 et 11 occurrences dans le corpus 2009.

La dimension principale que revêt cet ensemble dans le corpus 2007 est économique. L'avenir est alors florissant pour cette zone qui attire de multiples capitaux, et qui remporte de bons résultats pour les entreprises qui y sont implantées. Le journal Libération insiste également sur le comportement culturel ou historique de cette zone en montrant que celui-ci tend à la rapprocher de plus en plus de l'Europe occidentale.

Issu d'une famille originaire d' Europe orientale , Idan Raichel est un claviériste	li240207
Le microcrédit (..) est intéressant (...) en Europe orientale ou postconflits, comme en	li210307

4- « Europe ». Quelle représentation dans la presse française ?

Du temps du rideau de fer, le seul pays d' Europe orientale à se voir accorder l'organisation d'un euro li190407

Alors que la demande est forte en Europe orientale , devons-nous nous concentrer sur ec120207

Dans le corpus IntUne 2009, on relève 11 occurrences de cet ensemble, les lignes de concordances font alors apparaître un bloc où l'économie est en difficulté, ce qui tend à créer des tensions au sein de l'Union Européenne. D'autre part, l'histoire mouvementée de cette zone est mise en relief.

Inquiétudes sur la situation des pays d' Europe orientale . L'euro est tombé sous 1,26 dollars ec180209

de l'euro dont (..) est absente une partie de l' Europe orientale . fi040209

L'histoire est (..) passée à côté de mot. Ici, en Europe orientale ou centrale, c'est une situation exceptionn li120309

Entre les deux corpus IntUne, on remarque que cet ensemble est de moins en moins présent, en 2009, il semble d'autre part s'éloigner davantage de l'Europe occidentale, à travers des difficultés économiques qui l'excluent notamment de sommets européens, en outre, la presse française souligne le passé violent de cette zone.

IV.2.2.7- L'Europe du Nord.

Ce septième bloc géographique apparaît à 63 reprises en 2007 puis à 84 reprises en 2009. Sa répartition dans les différents journaux du corpus peut se résumer comme suit :

2007		2009	
Les Échos	40	Les Échos	25
Le Figaro	15	Le Figaro	19
Libération	8	Libération	3
Sud Ouest	10	Sud Ouest	12
Ouest France	8	Ouest France	4
Total	84	Total	63

(Figure 16, répartition de « Europe du Nord » dans le corpus IntUne)

C'est surtout la moindre citation de ce groupe dans le journal Les Échos en 2009 qui fait qu'il est moins présent, les autres journaux du corpus continuant à le présenter dans des proportions sensiblement semblables. Nous allons voir si cette progression entre les deux corpus IntUne s'accompagne d'une évolution de la représentation de cet ensemble dans la presse française.

L'Europe du Nord est avant toute chose dans le corpus 2007 un modèle à suivre, et cela quels que soient les domaines évoqués : développement durable, économie, modèle

4- « Europe ». Quelle représentation dans la presse française ?

social, éducation... Seul le journal Ouest France fait apparaître un léger bémol en une occurrence quand il explique que des prises de civelles bretonnes auraient du être utilisées pour repeupler les cours d'eau d'Europe du Nord à bas prix par décision de la Commission européenne. Au niveau économique, c'est également un marché porteur. Du point de vue défensif, c'est un allié fidèle de l'OTAN, qui aide le monde occidental à lutter contre le terrorisme, et culturellement, c'est une zone très riche.

Modèle social-démocrate, que l'on retrouve en Europe du Nord . fi200307
Le regard se tourne plutôt vers l' Europe du Nord où l'on est souvent fier de payer l'impôt, d li080207
de civelles pour repeupler les cours d'eau d' Europe du Nord ou140407
L'assistanat n'a pas la côte dans les pays d' Europe du Nord , où la morale et l'égalitarisme demeurent ec170307

Quand elle apparaît dans le corpus 2009, elle demeure ce modèle que la presse française encense. Son économie, son système de santé et d'éducation, son comportement environnemental est toujours un modèle à suivre dans les journaux du corpus, c'est également une zone où la crise économique se fait légèrement moins sentir que dans les autres pays européens.

leurs lacunes importantes face à nos voisins d' Europe du Nord . fi140209
La progression atteignant même 40% en Europe du Nord . Thomas Cook observe également une ec200309
Une méthode qui vient des États-Unis et d' Europe du Nord . Nous sommes allés nous former en su300309
une semaine sans qu'un nouveau nom filtre d' Europe du Nord Petit eldorado folk qui nous semble famili li290409

Cette Europe a ainsi pour principale acception dans la presse française celle d'un modèle à suivre, il est notable que ce modèle est d'autant plus présent, mis en valeur selon que l'on se situe dans une période d'élection ou non. Il est dès lors indéniable que loin d'être une simple dénomination géographique, la locution « Europe du Nord » recouvre une valeur sémantique forte dans la presse française que l'on se situe en 2007 ou en 2009, à savoir celle d'être un idéal à atteindre.

IV.2.2.8- L'Europe du Sud.

Cette Europe apparaît dans les mêmes proportions entre 2007 (77 occurrences) et 2009 (66 occurrences). Néanmoins, sa répartition dans les journaux du corpus évolue au bénéfice du Figaro en 2009.

En 2007, elle présente divers visages : vie culturelle riche, zone économiquement dynamique dans laquelle de nombreuses multinationales sont implantées (il est intéressant de

4- « Europe ». Quelle représentation dans la presse française ?

noter que dans ce cas, la Turquie fait partie de la zone en question), mais aussi zone fortement menacée par l'immigration, se situant en première ligne pour affronter ce problème. Parfois même, il est souligné que les migrants qui posent problèmes à l'ensemble de l'Europe peuvent provenir de ce sous ensemble. Enfin, elle présente au niveau culturel et historique une homogénéité certaine, partageant une conviction forte sur l'existence de racines communes pour cette Europe, racines pouvant être étendues à l'ensemble de l'Europe.

bitieuse de création d'agences en France et en Europe du sud a connu une explosion de son ec040407
Cette idée suscite plus d'élan dans les pays de l' Europe du sud , conscients que c'est dans le passé grec fi290307
vient d'être promu directeur général pour l' Europe du sud zone incluant la France, (...) et la Turquie fi030307
Un quart d'entre eux, en général d' Europe du sud et du Maghreb, optent pour des allers-retours li210207

Quand elle apparaît dans la presse française en 2009, elle est essentiellement une zone économique, dont le dynamisme s'est quelque peu essoufflé, et présente de nombreuses failles, surtout au niveau de son système financier, elle présente aussi une homogénéité culturelle que la presse française transmet à travers une image sportive.

La lente dérive financière de l' Europe du sud . ec160209
Ne suffit pas à compenser l'effondrement de l' Europe du sud . PSA Peugeot-Citroën est en retrait de ec160309
dégradation de la notation des pays d' Europe du sud et par les déboires de l'Irlande fi130209
Milan AC qui tenait le rôle de représentant de l' Europe du sud , cerné par Chelsea, Manchester et Liverpool fi170409

Cette Europe du Sud, au contraire de l'Europe du Nord semble dans le corpus IntUne être un ensemble qui ne parvient pas à se poser comme un modèle dans la durée, et dont l'économie met en péril ses voisins européens en 2009. Elle présente par ailleurs une homogénéité à la fois économique et culturelle.

IV.2.2.9- Une Europe unifiée ?

Cette revue des différentes « Europes » qui coexistent plus ou moins bien montre que ce sont des groupes ou blocs qui possèdent des réalités bien ancrées. Outre leur forte représentation en termes de fréquence absolue, ces ensembles présentent à chaque fois dans le corpus IntUne de fortes ressemblances et une grande homogénéité. De plus, il apparaît fréquemment que les uns soient présentés de façon opposée à leur pendant géographique. L'Europe de l'Est est économiquement dangereuse pour l'Europe de l'Ouest, qu'il s'agisse d'une forte concurrence en 2007 ou d'une déstabilisation du système monétaire en 2009. L'Europe du Nord est un modèle que ne parvient pas à égaler l'Europe du Sud, que ce soit au

4- « Europe ». Quelle représentation dans la presse française ?

niveau économique ou même sportif.

D'autre part, certaines de ces entités apparaissent avec des contours flous, c'est le cas des groupes « Europe de l'est », « Europe centrale », et « Europe orientale », nous avons vu que donner une liste des pays les composant précisément n'était pas chose aisée, certaines occurrences se contredisant : tantôt, ces groupes sont apparentés aux anciens pays communistes, tantôt, à des pays encore sous influence russe, tantôt à des pays délibérément atlantistes.

Cette première analyse d'un sens potentiel du mot « Europe » comme une entité divisée en blocs géographiques montre que ce concept est une réalité complexe et multiple, se détachant clairement en des sous ensembles qui, s'ils ont une dénomination géographique, sont davantage sous-tendus par leur passé, leur fonctionnement économique ou leur culture propre. L'Europe n'apparaît pas comme une entité unie de ce point de vue. Il serait maintenant intéressant de voir comment cette entité est présentée lorsqu'elle est désignée comme un tout. Dans un premier temps, la dimension économique de cette entité sera analysée.

IV.2.3- Une entité économique.

Lorsque le mot « Europe » est présent dans le corpus IntUne 2007, il apparaît en de multiples occurrences avec une acception d'entité économique. L'Europe peut être présentée dans ce cas comme ayant de bons résultats, ce qui représente la majorité des occurrences de ce sens potentiel, mais elle peut également rencontrer quelques difficultés. D'autre part, la presse française la présente comme menant une politique d'innovation et de formation. Elle est aussi montrée comme disposant d'une économie agricole importante ou enfin, comme négociant d'une seule voix à un niveau mondial. Nous allons donc voir dans les paragraphes à suivre quelles actualisations le sens potentiel d'« entité économique » prend-il dans la presse française en 2007.

IV.2.3.1- De bons résultats.

Tous journaux confondus, l'Europe apparaît comme obtenant de bons résultats dans la majorité des occurrences de ce sens potentiel. La plupart des journaux du corpus laissent

4- « Europe ». Quelle représentation dans la presse française ?

une place importante à la présentation de l'entité sous ce jour favorable. Il s'agit alors notamment d'un ensemble qui représente un marché important et porteur dans lequel de nombreux investissements sont réalisés. Ses entreprises montrent des indicateurs dans le vert, réalisent de bonnes performances.

Il peut néanmoins aussi s'agir d'une entité qui est déchirée entre une Europe de l'ouest où l'économie ne serait pas au beau fixe au profit d'une Europe de l'est qui lui prendrait son tissu industriel. Si les pays de l'Europe de l'ouest ou la France sont souvent cités comme ayant un bon développement ou une bonne « résistance » économique en de nombreuses occurrences, l'Europe de l'est, semble être porteuse d'encore davantage de promesses au niveau économique : ses pays sont notamment souvent qualifiés d'« émergents ». Celle-ci présente à la fois d'excellentes perspectives en terme de recrutement (au détriment d'autres pays de l'Europe de l'ouest), mais aussi de marché où il est possible de vendre ses produits. Libération est le seul journal qui montre une « Europe » davantage en difficulté qu'auréolée de succès dans le domaine économique. Néanmoins, il dresse quelques occurrences un portrait plus rassurant de l'entité. Les domaines où cette économie est porteuse sont alors le plus souvent ceux de l'internet et du transport par avion. D'autre part, cet ensemble serait en meilleure santé économique que les États-Unis, et parviendrait même à conserver des emplois dans ses secteurs industriels.

automobile a fini en tête des hausse sectorielles en Europe , tirée par DaimlerChrysler	fi200207
fermeture de son usine du Nord pour produire en Europe de l'est.	fi200307
ntissement économique des Etats-Unis et de la bonne santé de l' Europe	li280407
Progression beaucoup plus spectaculaire en Europe de l'est, au Royaume Uni (+24%)	su220207
Semi-remorques venus de toute l' Europe attendent pour charger des com	su200207
Une voiture neuve sur 20 est un 4x4. En Europe c'est même sensiblement plus.	ou240307
Pour les poids lourds, la demande a progressé de 11 % en Europe .	ec050207

IV.2.3.2- Des difficultés.

Toujours en tant qu'entité économique, l'Europe semble néanmoins éprouver certaines difficultés dans ce domaine avec une proportion moindre. L'inquiétude concerne notamment le monde de la bourse, des problèmes de suppressions d'emplois. Au niveau de l'avenir, les journaux du corpus notent également certaines inquiétudes à l'égard du secteur automobile. Dans une approche plus structurelle, les difficultés essentielles que rencontre

4- « Europe ». Quelle représentation dans la presse française ?

cette entité sont de lourds déséquilibres budgétaires et commerciaux, ou des difficultés à être innovante dans le domaine des nouvelles technologies. La presse française montre que les difficultés de cette entité sont alors surtout structurelles, et en adoptant des comportements plus valides économiquement, comme d'autres zones, telles les États Unis, elles pourraient être résolues. Seul le journal Libération donne de l'Europe économique une vision majoritairement pessimiste. La zone est « mature » et peu dynamique. Il est alors question de délocalisations, de suppressions de postes, mais aussi d'inégalités qui vont croissantes au sein de cet ensemble.

plan prévoit la suppression de 10 000 emplois en Europe , ainsi que la cession de six usines	fi100307
Europe : la dette bondit de 25% en cinq ans	fi200207
écart des rémunération est passé de 80 à 411 et l' Europe s'aligne sur ce croissant grand écart	li290307
USA, grâce à une main d'œuvre payée à des niveaux inférieurs à l' Europe , sont moins touchés.	ou280307
L'Amérique (..) est restée dynamique alors que l' Europe (+1,8%) fait encore moins bien	ec040507

IV.2.3.3- Une zone qui innove et forme.

L'Europe est ensuite montrée par la presse française comme une entité dans laquelle les citoyens sont formés dans le but de pouvoir ensuite s'insérer dans le monde économique, et elle a également une préoccupation, qu'elle soit privée ou publique, en termes de recherche et de développement. Ainsi, il sera dans le corpus question d'innovation : projets européens comme Galileo, ou encore projet du CNES, mais aussi initiatives plus privées. En règle générale, ces projets sont décrits comme porteurs, mais non encore existants, ou bien comme ayant des difficultés. Sur le plan privé, ce sont souvent des budgets de recherche qui disparaissent. D'autre part, il peut être question de recherche et de formation universitaire, la plupart du temps, les universités européennes et françaises sont montrées sous un jour plutôt favorable, avec quelques initiatives jugées encourageantes, mais on note surtout le besoin pour ces universités de s'améliorer. Ce qui apparaît fortement dans ce sens pour « Europe » est que selon l'orientation politique du journal qui le présente, la recherche privée sera décrite comme plus ou moins dangereuse, de même, la recherche publique sera elle aussi impactée par une image plus ou moins bénéfique :

L' Europe spatiale doit sortir de sa crise existentielle à l'heure où la Russie reno	fi280207
En Europe , les MBA sont beaucoup plus internationaux qu'aux États-Unis	fi050307
Galileo, le projet phare de l' Europe industrielle et spatiale, conçu pour rivaliser avec le GPS américain (G	li230307

4- « Europe ». Quelle représentation dans la presse française ?

Laboratoires de recherche d' Europe spécialisés en traduction automatique, le Geta (..) hériter direct de l' li100207
Innovation : l' Europe comble son retard avec les États-Unis Ec230207

IV.2.3.4- Une économie agricole importante.

L'entité « Europe » lorsqu'elle est économique, a une économie dans le secteur primaire qui revêt une importance particulière dans le corpus IntUne. La grande majorité de cette acception de l'Europe se trouve dans les colonnes du journal Ouest France. La réalité présente alors le plus souvent une forme de paradoxe, de déchirement et d'hésitation entre deux visions du rapport Europe / Agriculture. D'une part, l'Europe, et surtout « Bruxelles », quasiment personnifié en étant souvent le sujet de verbes d'action, sont des entités néfastes à l'agriculture européenne : elles la « sacrifient », s'y intéressent peu ou alors pour s'y « attaquer » ; des « contentieux » sont alors à régler entre ces parties.

D'autre part, une image moins problématique peut apparaître, dans des proportions moins importantes, du rapport agriculture / Europe, cette fois, l'entité aide l'agriculture, permet une augmentation de revenu, souhaite maximiser durablement la production, et veille aux stocks et ressources de pêche etc. Cependant, les termes utilisés pour donner une bonne image sont moins forts que les termes employés pour montrer qu'une véritable bataille est en jeu. Enfin, dans ces occurrences plus favorables, toutes ne se réfèrent pas explicitement à l'Union Européenne tandis que c'est toujours le cas lorsque l'image est défavorable. L'orientation politique du journal donne également une image de ce rapport agriculture/Europe sensiblement différente : les agriculteurs profitent ou bien sont écrasés par l'entité.

ssion européenne jette un pavé dans la marre de l' Europe bleue, la politique commune de la pêche. Elle a ou270207
L' Europe s'attaque aux rejets de poissons morts ou270307
revenu par actif agricole a progressé de 3,8% en Europe , en 2006 ou300307
perçoivent des subventions colossales de l' Europe . li100307
S'ils reçoivent des primes, c'est parce qu' l' Europe leur impose des prix de vente au-dessous des fi160307
terroir si riche et si emblématique (..) de l' Europe . su240207

IV.2.3.5- Une entité négociant d'une voix au niveau mondial

La dernière acception en termes de fréquence dans laquelle le mot « Europe » est considéré en tant qu'entité économique est celle d'une entité qui s'exprime d'une voix lors de

4- « Europe ». Quelle représentation dans la presse française ?

négociations sur les règles économiques entre grandes régions mondiales dans diverses occasions. Il s'agit alors le plus souvent de négociations menées à l'OMC (Organisation Mondiale du Commerce), ou de remontrances adressées à d'autres zones jugées moins vertueuses économiquement : la Chine par rapport à la fabrication de contrefaçons, le Japon ou les États-Unis pour leur politique monétaire. Néanmoins, les régions « Europe » et « États-Unis » semblent également se rapprocher au niveau économique.

Japonaise suscite (..) une grande irritation en Europe , la réponse du G7 constitue un fi120207
Les États-Unis et l' Europe veulent renforcer leurs liens éco fi240407
Washington impose 116 millions (..) de pénalités annuelles à l' Europe . ou280307

IV.2.3.6- Pour conclure.

Le sens potentiel « entité économique », lorsqu'il est activé dans le corpus IntUne présente une réalité complexe se composant de diverses facettes. Si l'Europe est une zone essentiellement en bonne santé économique, un certain schisme se crée au sein de cette entité, ce qui semblait déjà apparaître avec l'existence d'une multitude d'« Europes » différentes analysées dans la partie 2.2 de ce chapitre. De plus, l'activation de ce sens dans la presse française en 2007 montre que cette bonne santé économique est assez régulièrement relativisée par des résultats peu satisfaisants. Ses systèmes d'innovation et de formation semblent fournir le même bilan mitigé : entre réussites et difficultés structurelles lourdes.

Son économie agricole, base historique et symbolique de la construction européenne à travers la Politique Commune Agricole (PAC) connaît quant à elle de graves difficultés, seule la qualité de ses terroirs semble lui assurer une réussite relative. Enfin, elle est capable de faire preuve d'unité, mais il s'agit alors d'affronter un monde très concurrentiel voire même agressif, et l'Europe est alors bien souvent présentée comme la plus faible de la partie.

IV.2.4- Une entité partiellement composée d'une Union politique : l'Union Européenne.

Le mot « Europe » est très présent dans un contexte où l'on parle en fait de l'Union Européenne : c'est le contexte le plus important en termes de fréquence dans lequel ce mot apparaît au sein du corpus IntUne 2007. Un des sens potentiels que l'on pourrait donc donner

4- « Europe ». Quelle représentation dans la presse française ?

pour le mot « Europe » au niveau de son activation dans la presse française en 2007 est : « union politique ». Cette union politique est présentée de diverses façons dans la presse française, on verra dans une première partie que ses difficultés sont évoquées dans une large majorité des occurrences, puis, nous noterons cependant la présence dans le corpus IntUne de quelques réalisations positives pour cette entité. Enfin, nous verrons qu'il s'agit d'une union qui se base sur un passé culturel riche et commun.

IV.2.4.1- L'Union européenne confrontée à de nombreuses difficultés.

Cette « Europe » politique semble rencontrer de nombreux problèmes. Un des problèmes principaux de cette entité concerne ses difficultés institutionnelles qui tardent à être réglées à la suite de l'élargissement de l'Union Européenne à vingt-sept pays en 2007. Les mots que l'on note le plus fréquemment dans ce contexte sont par exemple : « crise », « blocage », et on insiste fortement sur le besoin d'une « relance ». Dans le corpus IntUne 2007, il apparaît que c'est l'élargissement qui est source de la plupart de ces maux institutionnels. Outre des problèmes institutionnels, les différents journaux du corpus IntUne 2007 mettent également en avant dans l'Union Européenne un grave déficit démocratique. L'Union Européenne doit d'autre part affronter dans la presse française une inefficacité patente dans ses projets industriels communs, qu'il s'agisse d'Airbus, dont le plan de restructuration conduisant à la suppression d'une dizaine de milliers d'emplois dans toute l'Europe est amplement commenté, du projet Galileo qui rencontre maints obstacles à sa mise en service, ou encore de la mise en place d'un nouveau moteur de recherche destiné à concurrencer l'américain « google », mais qui semble avoir une efficacité relative.

L'Union Européenne est enfin en de nombreuses occurrences une entité qui rencontre diverses difficultés : manque d'efficacité dans sa politique de communication, absence criante de solidarité entre ses différentes régions, incurie de la part de sa gouvernance à gérer les problèmes qui s'y présentent, voire des paradoxes réels entre la volonté de l'Union Européenne et les actes qu'elle réalise, notamment dans le domaine des droits de l'homme.

cette capitale européenne : comment sortir l' Europe de son impasse institutionnelle ? fi220207
Coups de poignard en série contre l' Europe . li090307
Je pense qu'il veut aussi dire que l' Europe ne peut pas continuer à s'agrandir sans une réfo su230307

4- « Europe ». Quelle représentation dans la presse française ?

Ce n'est pas à cause de lui que l' Europe	ne fait plus preuve d'aucune volonté	su260207
français () qui est responsable de la crise de l' Europe	, c'est la crise de l'Europe.	ou230207
Deux ans ont passé après le « non » français et l' Europe	est encore en panne.	ou160307
Le problème de l' Europe	, c'est qu'elle se fait à l'insu des citoyens	fi220207
mépris pour l'électeur fourvoyé dans le(..) non à l' Europe	pour qu'une parole tombant d'en haut puisse en	li050207
Marika 25 ans : L' Europe	est trop centralisée et trop administrative	li210307
La politique monétaire de l' Europe	est considérée comme l'une des causes des diff	fi060307
A tâtons : Europe	Les mésaventures de Galileo valent symbole :	li240407
de l'afssa, et surtout, rien n'a bougé au niveau de l' Europe	.	su140207
C'est contraire aux valeurs de l' Europe	mais il n'est pas sûr que le Parlement puisse fa	li160207
La solidarité entre régions d' Europe	n'est guère populaire	ou120207
st admise, comme si de rien n'était, au Conseil de Europe	, garant des droits de l'homme	ou240407
	l'	
L' Europe	ne peut être le produit d'une somme d'impuissa	ou300107

D'autre part, le corpus IntUne 2007 montre aussi l'Europe comme étant une entité politique qui entretient un rapport pour le moins complexe avec la France. Il semble alors que la France ne puisse exister sans l'Union Européenne et inversement, bien que ce rapport soit problématique. La France est aussi souvent vue comme une sorte de « poids lourd » en Europe où elle tiendrait une place importante, bien que cette prépondérance soit remise en question par le non au référendum sur la constitution. La campagne présidentielle française reste néanmoins suivie avec intérêt dans toute l'Europe.

D'autre part, il est souvent noté que la France aurait des prétentions impérialistes au niveau de l'Europe, voulant en faire une « France en grand », ces prétentions seraient alors des réminiscences de la période gaulliste. Il est notamment question dans la presse française en 2007 d'une France qui est un pays essentiel pour l'Union Européenne, c'est LE pays qui l'a créée, le pays qui fait qu'elle peut également avancer ou stagner, or, un certain malaise ou désaccord est noté : la France n'a plus le prestige qu'elle devrait avoir en Union Européenne, l'Union est donc en crise. D'autre part, la presse donne également à la France l'image d'un pays s'enlisant dans des idées obsolètes en complet décalage avec la réalité européenne. Ce sont ainsi de nombreux paradoxes entre ce que voudrait la France par rapport à l'Union Européenne et son comportement effectif qui sont dépeints dans le corpus IntUne 2007: on aime, mais on honnit deux parties d'un même tout.

Si la France a besoin de l' Europe	il est au moins tout aussi vrai que l'Europe a besoin	fi190307
cette campagne, comme si l'avenir de l' Europe	en dépendait.	fi200307

4- « Europe ». Quelle représentation dans la presse française ?

Et l' Europe sans la France ne serait qu'un grand continent amp fi240307
Un certain malaise vis-à-vis de l' Europe est perceptible dans tous les pays, mais la France
elle, n'a jamais vraiment clarifié son rapport à l' Europe rêvant des avantages d'une Europe fédérale sans li240307
onomique comme elle voulut régner sur toute l' Europe . li190407
ous reste à cesser de croire que le monde et l' Europe vont s'adapter à notre « modèle » obsolète mais su050307
Aimer l' Europe mais honnir Bruxelles ou170307

Cette entité politique est aussi présentée par la presse française comme un ensemble qui est doté de diverses règles. Ces règles peuvent concerner notamment le libre échange, ou bien l'environnement et la qualité de l'eau. La quasi totalité des occurrences se trouve alors dans le journal Ouest France où c'est une Commission européenne autoritaire, punitive, contraignante et peu encline au dialogue qui est décrite : l'institution « s'attaque » alors à la France, toujours dans le secteur primaire qui est essentiel dans la région de parution du quotidien.

La France a été contrainte par l' Europe de réduire sa capacité de pêche de 3% au prix de ou150207
L' Europe vient d'adresser un ultimatum cinglant à la France ou310307
Ascenseurs : amende de 992 millions d'euros : Europe . » fi210207

IV.2.4.2- Une union dotée d'une image favorable.

Néanmoins, ce contexte dominant fort peu favorable à l'Union Européenne au sein de la presse française est nuancé par la mise en valeur de diverses réalisations positives de cette création politique. Ainsi, certaines occurrences donnent une vision plus amène de cet ensemble politique. Avant toute chose, la réussite la plus mise en avant par les différents quotidiens du corpus IntUne réside dans la pacification de cette zone qui apparaît comme une réalisation extraordinaire, un pari pour lequel le succès n'était pas chose facile, cette zone ayant connu une histoire des plus belliqueuses depuis de nombreux siècles. Dans ce cas, cette entité politique est vecteur de paix non seulement à l'intérieur de ses frontières, mais également à l'extérieur, faisant rayonner son idéal pacifique, et créant des passerelles entre les diverses zones du monde, aidant à leur coopération mutuelle. D'autre part, cette construction, faite de solidarité et de coopération, est mentionnée comme ayant apporté de nombreux bienfaits à ses citoyens, notamment concernant leur pouvoir d'achat et leur possibilité de consommer davantage. Certains des habitants se montrent en conséquence satisfaits de l'Union Européenne.

A un niveau matériel, l'Europe peut aussi être présente sous la forme d'une

4- « Europe ». Quelle représentation dans la presse française ?

généreuse donatrice, permettant à des projets la plupart du temps locaux et donc très présents dans la vie des citoyens (réalisation de collèges, d'infrastructures routières, échanges scolaires) de se réaliser. Cette dimension est présente dans tous les journaux du corpus.

L' Europe	se félicite d'avoir surmonté les divisions d'un co	fi220307
Au long des cinquante dernières années, l' Europe	a été une force génératrice de liberté et de sol	su230307
L' Europe	, c'est la paix, la paix non seulement pour l'Eur	ou240307
au-dessus de l'Atlantique entre les Etats-Unis et l' Europe	, mais aussi aujourd'hui entre la Russie et l'Eur	ou050407
négocié collectivement en Europe	, chaque ménage français réalise ainsi quelque	ou280307
Les Polonais satisfaits d'être dans l' Europe	. Élargissement.	fi040507
Avec l' Europe	, la consommation a progressé	su090207
Le premier financeur de ce projet est l' Europe	, via le Federt, avec près de 40% de l'addition to	su270207

IV.2.4.3- Une entité politique qui se base sur une culture commune.

Cette union politique entretient un lien étroit avec la base culturelle de ses pays membres, supposée commune dans une large part. Par culturel, est ici inclus tout ce qui concerne le domaine historique, linguistique, artistique ou religieux. L'entité « Europe », dans une acception culturelle peut se séparer en quatre sous ensembles, selon l'évocation de critères bien précis : la vie artistique, historique, religieuse ou encore linguistique. Nous allons maintenant voir quelle représentation la presse française donne-t-elle de la base culturelle de l'ensemble européen.

IV.2.4.3.1- Une vie artistique rassurante, entre homogénéité stabilisatrice et diversité enrichissante.

La vie artistique en Europe en donne à voir une facette plutôt favorable, en effet, l'essentiel des occurrences se situant dans ce domaine font de l'Europe une zone où l'on trouve une grande homogénéité culturelle. L'Europe y apparaît comme s'étant construite sur des idées communes, par exemple suite au mouvement philosophique des Lumières au XVIIème siècle. Parfois, l'Europe apparaît comme une entité culturelle qui donne à voir une diversité entre différents sous ensembles européens. Mais dans ce cas, la diversité n'est pas source d'incompréhension ou de concurrence, il s'agit davantage de quelque chose d'enrichissant, qui est vu comme une chance.

Palais Farnèse, mais avec Volpato, c'est toute l' Europe	qui s'y met	fi150207
Renaissance des cités d' Europe	; une association sensibilisée à la valorisation et	su140307
Toute l' Europe	célèbre les 400 ans de l'opéra ce samedi	ou170207

4- « Europe ». Quelle représentation dans la presse française ?

Je pense à l' Europe des Lumières	ou240307
fragments s'accumulent, plus le film se révèle. Europe de l'est, cela m'évoque une atmosphère magique	fi070207
Sonorité française d'un vernis très vieille Europe , avec des cordes chaleureuses et un pupitre	fi100207
Varsovie en est d'ailleurs une variante unique en Europe centrale.	li280207
de Chagall et en écoutant la musique juive d' Europe de l'est.	su100207

IV.2.4.3.2- Une Histoire commune ancienne, alternant entre violence et coopération.

Dans le corpus IntUne 2007, l'Europe est aussi une entité qui partage une histoire commune depuis de nombreux siècles. Cette histoire montre soit des bases culturelles communes, soit de violents affrontements. Ainsi, l'accent peut être porté sur le passé commun et les racines gréco-romaines de cette entité, sur les deux guerres mondiales qui ont été déclenchées dans cette zone au vingtième siècle, ou sur le passé communiste de toute une partie de l'ensemble. Parfois, ce passé commun peut se construire avec la Turquie, l'intégrant alors à l'histoire européenne. Les journaux du corpus montrent aussi cette zone comme partageant une histoire de coopération, à l'occasion du cinquantenaire du Traité de Rome marquant le début de la construction européenne. D'après le portrait historique que la presse française fait de l'Europe, nous pouvons nous poser la question des pays qui en font partie, peut-on considérer en effet la Russie ou la Turquie comme des pays européens ?

Pour avoir affirmé, lors d'un colloque, que l' Europe débutait à la fin de l'Empire romain quand la chré	fi220207
À l'université de Harvard s'adressait à une Europe ruinée par la guerre	fi260307
« socialisme de caserne » qui sévissait à l'est de Europe .	fi040507
Le génocide -...- n'est revenu en Europe sur le devant de la scène que pendant les con	li270207
bole de l'orgueil, elle pourrait aussi symboliser l' Europe , Bosphore inclus	li280307
mandé à Dieu de réconcilier les deux parties de l' Europe et de mon coeur, se souvient-il dans une inte	su280407
Si l' Europe n'est pas une histoire d'amour, elle a été une his	ou030307
Car partout en Europe on bataille, on étriipe, on massacre (c'est la guer	ou300307

Ce contexte historique fait donc certes apparaître un passé commun, mais montre également d'une part, que les pays intégrant ce passé peuvent être multiples et variés, et d'autre part, que ce passé est considérablement marqué de déchirures, de conflits, l'équilibre au sein de cet ensemble semblant fragile.

IV.2.4.3.3- Une entité marquée par les religions.

Le mot « Europe » apparaît également comme marqué par les religions, que cette influence soit problématique ou non. Ce sont les trois religions monothéistes qui sont alors

4- « Europe ». Quelle représentation dans la presse française ?

évoquées.

La religion chrétienne intervient notamment dans le cadre d'un débat sur le devoir de l'inscription des « racines chrétiennes de l'Europe » dans la constitution européenne. Cette question suscite en effet un vif débat entre personnes pensant que l'Europe s'est constituée ou non autour de cette religion. D'autre part, un déclin de cette pratique chrétienne est observé en Europe. La religion musulmane semble porter davantage de problèmes quand elle est associée au mot « Europe ». Cette fois, la question de l'influence mutuelle des deux entités n'est plus envisagée, c'est surtout une problématique d'affrontement qui est en jeu, que ce soit au niveau de la non reconnaissance par l'Union Européenne de certains faits de guerre à l'encontre de musulmans, ou de remontrances faites à des pays dont les gouvernants s'inspirent de cette religion. De plus, l'entrée de la Turquie, dont la population est à majorité musulmane, dans l'Union Européenne pose de nombreux problèmes.

La religion juïque est beaucoup moins présente, il y est alors question des persécutions subies par les juifs tout au long du vingtième siècle, mais aussi de la proximité culturelle qui relie Israël à l'Europe. Il peut encore être question d'un livre antisémite financé par l'Union Européenne qui fait scandale ou de la bonne intégration des juifs en Europe. Cette religion jouit globalement d'une vision beaucoup plus consensuelle et favorable dans l'ensemble du corpus 2007. C'est également la moins présente en terme de fréquence.

C'est au christianisme que l' Europe	doit sa civilisation morale et, dans toute l'Eur	fi100407
Cependant, pour Pau Veyne, l' Europe	dans laquelle nous vivons doit moins au li050407	christianisme qu'à tous ceux qui l'ont combattu
sur la manière dont la France et l' Europe	ont vécu la sortie de la religion tout en restant h	su110407
Compostelle qui a été sacré 1er itinéraire culturel		
européen par le conseil de l' Europe	le 23 octobre 1987	ou140207
censure avait, dans la foulée, parcouru l' Europe	, contre laquelle l'hebdomadaire satirique ente	fi230307
L' Europe	a été construite sur un principe islamophobe	li280207
que la Turquie, c'est l'Asie mineure, ça n'est pas l' Europe	.	li030507
Les totalitarismes, religieux (..) veulent imposer		
leur loi jusqu'aux portes de notre Europe	?	ou230407
communauté juive de France est la seule en Europe	à avoir autant grandi après la guerre, sur les li	fi290307
nombre, au fil des persécutions. Ils gagnent l' Europe	du Nord (...) où ils sont bien intégrés	ou270207

IV.2.4.3.4- Des influences linguistiques communes.

La réalité linguistique de l'Europe est aussi évoquée en 13 occurrences dans notre

4- « Europe ». Quelle représentation dans la presse française ?

corpus. Trois journaux associent le mot « Europe » à ce contexte. L'entité est alors présentée dans les deux journaux régionaux comme ayant une volonté de politique commune dans ce domaine en souhaitant sauvegarder les langues régionales, la France étant plutôt désignée comme un mauvais élève. L'Europe peut aussi être considérée comme ayant une culture qui s'est fondée sur des bases linguistiques communes, le Figaro insiste alors sur l'importance qu'a pu revêtir le français dans cet ensemble.

français commença de rayonner dans les cours d' Europe , acquérant un statut de langue diplomat	fi110407
La France est l'un des rares pays d' Europe à ne pas avoir ratifié la charte européen	su090407
français (..) n'a cessé de se ressourcer aux langues d' Europe	ou100307

IV.2.4.4- Une entité politique à géométrie variable.

L'Europe est aussi une entité qui peut être précisée en fonction d'un nombre, ce nombre désignant alors de façon assez simple les États membres de l'Union Européenne. Cependant on s'aperçoit qu'une confusion apparaît alors, en effet, si depuis 2007, l'Union européenne se compose bel et bien de vingt-sept États membres⁷¹, plusieurs nombres apparaissent dans la presse française pour préciser cette entité, il est alors complexe de savoir de quelle Europe on parle...

Les deux « Europes » les plus présentes dans le corpus à égalité sont « L'Europe des quinze » et « l'Europe des vingt sept ». Il devient donc compliqué de savoir si l'Union Européenne se compose de pays dont la forte majorité se situe en Europe de l'Ouest (= « Europe des quinze ») ou s'il s'agit d'un mélange de l'Europe de l'Est et de l'Europe de l'Ouest (= « Europe des vingt sept ». L'Europe des quinze est en fait une sorte d'étalon auquel on peut comparer le comportement de la France dans notre corpus, c'est aussi une zone qui a un problème de coût de la main d'œuvre trop élevé. On peut en déduire qu'en France, la presse parle beaucoup de cette Europe, qui institutionnellement n'existe plus, car elle a une plus grande proximité avec la France, l'Europe des vingt sept présentant une réalité beaucoup moins proche :

La France se situe dans la moyenne de l' Europe des Quinze, comme le rappelle un rapport du li040507

L'Europe des vingt sept, quant à elle, présente une image beaucoup moins

71 Allemagne, Autriche, Belgique, Bulgarie, Chypre, Danemark, Espagne, Estonie, Finlande, France, Grèce, Hongrie, Irlande, Italie, Lettonie, Lituanie, Luxembourg, Malte, Pays-Bas, Pologne, Portugal, Roumanie, Royaume-Uni, Slovaquie, Slovénie, Suède, République Tchèque.

4- « Europe ». Quelle représentation dans la presse française ?

homogène et rassurante, en effet, elle est au mieux existante dans un contexte jugé exceptionnel grâce à un accord sur l'action contre le réchauffement climatique, au pire victime d'un schisme interne. Il est également noté que son centre de gravité se situe à l'Est.

Le point de passage obligé vers cette Europe des 27 résoluement tournée vers l'est dont le cent ou240307
doute (...) y compris ceux qui viennent de (...) l' Europe des Vingt-Sept. fi200307
En unissant l' Europe des Vingt-Sept dans la lutte contre le changeme fi150307
référendum en 2009, ce qui signifie que l' Europe des Vingt-Sept serait encore paralysée pendant ou240307

L'Europe des six, enfin, est présente à trois reprises. Dans ces 3 occurrences, elle est une occasion de rappeler le prestige passé de la France, et la place que ce pays a eu dans la création de l'Union Européenne :

Cette Europe des Six, dominée politiquement par la France et économiquement par l'Allemagne, cette su230307

IV.2.4.5- Pour conclure.

Ainsi donc, l'Europe dans le corpus IntUne 2007 est très fréquemment comprise sous l'acception d'une entité politique. Ce sens potentiel d'une « entité partiellement composée d'une union politique » semble couvrir une réalité contrastée : cette entité politique connaît de nombreuses difficultés, quand bien même des avantages peuvent y être trouvés dans une moindre proportion. Elle alterne entre bénéfiques pour ses citoyens et perte de souveraineté au niveau national.

Cette union politique s'est construite sur une base culturelle commune, mais nous venons de voir que dans bien des domaines, cette base culturelle pouvait être floue, que parfois même, il était complexe d'établir au juste quels pays appartenaient au continent européen. La presse française semble ne trouver de compromis que pour exclure la religion musulmane de ce passé européen commun. Cette entité, bien que supranationale est largement montrée dans la presse française comme entretenant un lien des plus complexes avec la France : lien très fort puisque les fondateurs de l'Union Européenne, ceux qui en ont imaginé l'idée étaient pour beaucoup des hommes politiques français. En revanche, ce lien se distend dangereusement du fait non pas d'une crise française, mais d'une crise européenne dont les sources pourraient plus vraisemblablement se trouver, selon les médias français, dans l'élargissement à vingt sept pays en 2007. La dernière sous partie de la présentation de cette entité politique montre d'ailleurs assez explicitement ce sensible désamour de la presse française pour l'Europe des vingt sept, puisqu'elle continue de faire référence à une Europe

4- « Europe ». *Quelle représentation dans la presse française ?*

composée de quinze États membres en de nombreuses occurrences, voire parfois de l'Europe à l'époque de la signature du Traité de Rome.

IV.2.5- L'Europe politique est une idéalisation.

L'Union Européenne, nous venons de le voir connaît une fortune mitigée dans la presse française en 2007. Parfois, des réussites sont mentionnées, ou des souhaits vis à vis de cette entité. En fait, dans la majorité des cas, ces réussites ou ces descriptions favorables de l'Union Européenne sont des idéalizations. Nous allons maintenant donner quelques exemples de cette idéalisation à travers son évocation dans des constructions linguistiques, nous verrons également que certains thèmes, comme l'action environnementale, peuvent être prétextes à cette idéalisation de l'Europe politique. Enfin, l'utilisation de cette représentation de l'Europe sera abordée dans le cadre de la campagne présidentielle française.

IV.2.5.1- Une idéalisation présente à travers différentes constructions linguistiques.

Si l'Europe est parfois présentée sous un jour favorable dans le corpus IntUne 2007, il est important de noter qu'elle est alors fréquemment idéalisée. Outre les réseaux sémantiques utilisés qui la situent d'emblée dans une thématique noble, mais abstraite (« valeurs », « rêve », « idéalisée »), cette entité apparaît dans des énoncés au futur, dont la réalité est absente. D'autre part, certaines constructions montrent clairement que l'on n'est pas dans le domaine de la relation de faits, notons ainsi la métaphore d'une Europe maternelle et protectrice ou encore la construction « que l'Europe soit » parfois utilisée, qui montre un parallélisme frappant avec l'incantation divine du livre de la genèse.

Un autre trait linguistique montre parfois une idéalisation de l'Europe, il s'agit de l'emploi de ce mot dans différents noms de marque. La conception d'un nom de marque suppose que l'on prête des qualités positives au concept que représente ce nom et que l'on espère qu'une partie de ces qualités ressurgiront sur la marque créée sur cette base. La présence du mot « Europe » comme partie d'une marque est importante dans le corpus IntUne 2007. La grande majorité des occurrences de cette Europe / marque concerne la radio « Europe 1 », ainsi que l'émission télévisée « France Europe Express ». Cela s'explique notamment par le fait que l'émission en question et la radio concernée ont une forte

4- « Europe ». Quelle représentation dans la presse française ?

dimension politique, et donnent de nombreuses interviews des différents candidats à l'élection présidentielle. D'autre part, le mot « Europe » peut aussi être utilisé pour dénommer des places dans diverses villes, ou des théâtres, la fonction symbolique est ici plus forte : un théâtre ou une place de l'Europe existe, c'est donc une manifestation concrète d'une entité parfois trop abstraite.

L'utilisation d'adjectifs à la suite du mot « Europe » est aussi une façon d'idéaliser cette entité. L'adjectif « entière » apparaît ainsi à de nombreuses reprises comme qualificatif du mot « Europe » dans le corpus IntUne. Contrairement à la plupart des autres compléments apportés au mot « Europe », ce terme donne de l'entité une vision unie, d'un tout compact, agissant comme un seul corps. Les domaines dans lesquels cette entité se manifeste sont économiques, culturels ou encore politiques, dans tous les cas, cet ensemble est envisagé de façon métaphorique sous la forme d'une sorte de public (ou de marché dans un contexte économique) qui vit des événements. Le qualificatif « forte » montre aussi une entité idéale, que l'on appelle de ses vœux. Dans le corpus 2007, toutes les occurrences révèlent le souhait que cette entité se crée, sans que la procédure de cet avènement ne soit davantage précisée. Elle est également fréquemment considérée comme étant la seule façon d'accéder à une Europe plus sociale.

valeurs dont elle se réclame pourront essaimer et d'autant plus qu'elles s'appuieront sur une	Europe	volontariste	su260207
Ils carressent le rêve d'une	Europe	qui damnerait le pion aux États-Unis, mais quan	fi140307
Ainsi, l'	Europe	réapparaîtra-t-elle comme un nouveau modèle	li270207
qui fera de l'	Europe	une protection pour les français et non pas le	li030507
Une	Europe	idéalisée pour faire rêver d'un monde meilleur,	ou240307
Hubert Védrine à Lorient « Que l'	Europe	soit une puissance ».	ou070407
quartier du Grand-Parc, place de l'	Europe	à Bordeaux	su070207
constituera aux yeux de l'	Europe	entière une rupture avec la prétendue	fi280407
L'Union Européenne voudrait le généraliser à l'	Europe	entière.	fi280307
avenir commun et que seule une	Europe	forte de ses racines pourra se permettre de	fi290307
On a donc besoin de faire une	Europe	forte et en bonne santé, pour qu'elle le soit	li300407
et le monde n'ont jamais eu tant besoin d'une	Europe	forte qui défende les règles au niveau de l'ON	su230307

IV.2.5.2- L'idéalisation à travers le thème de l'engagement environnemental

L'Union Européenne est très souvent associée à une action environnementale dans

4- « Europe ». Quelle représentation dans la presse française ?

le corpus IntUne 2007. C'est alors un domaine dans lequel l'Union s'illustre et prend une place de leader « naturel » au niveau mondial. C'est aussi une question qui lui permet d'agir de façon unie, et elle joue ici le pion notamment aux États-Unis jugés très mauvais élèves. Cette présentation de l'Union, parvenant à agir dans un domaine de façon efficace est en fait très souvent une façon de présenter l'entité comme idéale, ayant une action novatrice et unique au niveau planétaire, ce qui n'est pas toujours le cas dans la réalité.

Une telle taxe était défendue par l' Europe , mais elle s'est heurtée à la position américain li020307

Dans ce domaine, l' Europe montre le bon exemple. su070407

IV.2.5.3- Un sujet de l'élection présidentielle française.

Un autre contexte dans lequel le mot « Europe » est largement idéalisé est celui de la campagne présidentielle française. L'Europe est un des sujets, un des thèmes importants de cette campagne. Bien que le mot soit rarement précisé, il représente le plus souvent de façon implicite l'Union Européenne. Les divers candidats à l'élection sont alors fréquemment tentés de donner pour cette entité une vision « idéale », d'une part, leur propos se situe dans un futur non réel, d'autre part, il s'agit de séduire un nombre maximum d'électeurs. L'Europe, qui est dans ce cadre considérée comme une entité lointaine et abstraite est ainsi régulièrement idéalisée par les différents candidats.

L'utilisation de nombreux conditionnels, ou de modaux est en conséquence fréquente chez les hommes politiques qui s'expriment sur ce thème. D'autre part, l'hyperbole est un procédé fréquent, les moindres déclarations concernant l'échelle planétaire. Des mots comme « doit », « voulu », « rêve », contribuent à renforcer cette idée d'une Europe idéalisée dans le discours politique. La tournure « Une Europe qui... » est également présente 28 fois dans le corpus IntUne, le plus souvent utilisée par les candidats à l'élection présidentielle et représentant une Europe allégorique, largement idéalisée. L'utilisation d'adjectifs à la suite du mot « Europe » est aussi une façon de l'encenser pour les candidats à l'élection présidentielle. Parmi les qualificatifs qui reviennent le plus souvent, on trouve notamment les mots « sociale », « politique » ou « solidaire » qui recouvrent des réalités du domaine du vœux.

L' Europe doit protéger, a-t-il expliqué fi070307

L'amitié entre l' Europe et les États-Unis est une nécessité pour l'équilibre du monde fi010307

4- « Europe ». Quelle représentation dans la presse française ?

Ceux, obstinés, qui ont voulu l'	Europe	a dit le candidat UDF qui se bat pour que l'Europe ne soit pa	fi130207
Bayrou rêve encore d'une	Europe	française, mais il y a belle lurette qu'une majorité des 27 pa	fi230207
rejetant une	Europe	qui ne serait qu'une zone de libre-échange adossée à l'Otan	ou010307
Pour elle de plaider pour une	Europe	qui protège des délocalisations	fi290307
d'après guerre, mais une	Europe	qui apporte des solutions	fi100407
réussir à reconstruire une	Europe	qui réussisse à lutter contre le chômage	li300407
Il faut une	Europe	qui réussisse à lutter contre les délocalisations	li300407
Je veux une	Europe	qui protège ses citoyens et qui les aide	ou010307
d'une Europe politique, d'une	Europe	qui garantisse notre modèle social	su120307
Buffet à Berlin pour défendre l'	Europe	sociale.	ou100307
Ainsi, l'	Europe	politique, donc sociale, a de grandes chances	fi050207

L'idéalisation de l'Europe politique rencontre un fort succès dans le corpus IntUne 2007. Cette idéalisation peut s'effectuer selon différents procédés que nous venons d'analyser (propositions relatives, conditionnel, futur, métaphore, hyperbole, etc.) et elle rencontre souvent le discours politique, très présent au moment de la collecte du corpus. Si elle montre une volonté d'améliorer la représentation de cette entité politique, elle pose néanmoins le problème de lui conférer une essence irréaliste et lointaine. Nous venons de voir dans les parties IV.2.4 et IV.2.5 que l'Europe était une entité politique, et qu'elle était largement idéalisée dans la presse française, notamment via le discours d'hommes politiques à son sujet. Cette entité politique existe néanmoins concrètement, et cette existence suppose entre autres la présence de frontières. Nous allons maintenant voir comment cette question des frontières est abordée dans le corpus IntUne 2007.

IV.2.6- L'Europe politique doit défendre ses frontières.

Le mot « Europe » apparaît très fréquemment dans le corpus IntUne 2007 dans le sens d'une entité qui devrait défendre ses frontières. Ces frontières se rapportent à l'existence d'une Europe, entité politique. Elles peuvent alors faire surgir des problèmes de conservation, qu'il s'agisse des agressions de voisins belliqueux, de menaces terroristes, de flux de population difficiles à gérer voire de virus. Dans une première partie, nous verrons les tensions internationales que cette entité doit gérer au mieux pour préserver ses frontières. Dans une deuxième partie, nous aborderons les problèmes de flux migratoires, enfin, une troisième partie étudiera la défense de ces frontières contre des virus.

4- « Europe ». Quelle représentation dans la presse française ?

IV.2.6.1- Une entité au sein de tensions internationales.

Les différents journaux du corpus abordent tous des tensions au niveau international entre les différentes grandes régions du monde dont l'Europe fait partie. Différents portraits sont brossés selon l'orientation politique du journal qui parle de cette Europe défendant ses frontières. Dans *Le Figaro*, le problème dominant concerne des relations conflictuelles avec la Russie, le plus souvent dénommée par une métonymie (« Moscou »). Ces difficultés sont principalement dues à l'adhésion de certains anciens pays du bloc communiste à l'OTAN, et au fait que les États-Unis souhaitent installer un bouclier antimissile dans ces pays. Le conflit peut aussi concerner l'élection d'un gouvernant pro-russe en Biélorussie, Alexandre Loukachenko est ainsi considéré par l'État major américain comme un « dictateur ». Le contexte ukrainien tendu entre pro-russes et pro-européens (Union Européenne) est aussi évoqué dans ce cadre. En outre, on insiste sur la dépréciation de l'image de la Russie en Europe pour diverses raisons.

Ce journal présente d'autre part l'Europe comme une entité organisée, présentant diverses structures dans le but de se défendre, essentiellement l'OTAN et une « Europe de la défense ». En ce qui concerne l'OTAN, deux visions paradoxales en sont données, cette organisation semblant à la fois intervenir davantage en Europe (installation d'une base en Italie), mais également y diminuer son influence au profit d'autres régions mondiales (Maghreb). Quoiqu'il en soit, *Le Figaro* ne remet jamais en cause la proximité entre Europe et États-Unis. L'« Europe de la défense » est dans ce journal une entité en projet, quelque chose que l'on souhaite, voire que l'on idéalise, qui dans tous les cas est nécessaire, mais pour l'instant quasi inexistante. Ce contexte fait aussi apparaître l'idée d'une modification de l'ordre géopolitique mondial, on sous entend alors que l'Europe peut être en position de défier le leadership américain.

Le journal *Libération* reprend également majoritairement l'idée d'une entité connaissant des tensions avec certains de ses voisins, et surtout la Russie, dans le cadre de l'installation d'un bouclier antimissiles en Pologne par les États-Unis. L'image de la Russie est d'autant plus ternie par l'évocation de droits de l'homme bafoués en Tchétchénie. L'hégémonie américaine y est aussi fortement dénoncée, cela représentant la deuxième idée principale en termes de fréquence dans laquelle on trouve l'entité européenne comme devant

4- « Europe ». Quelle représentation dans la presse française ?

défendre ses frontières dans le journal Libération. Enfin, à la différence du Figaro, ce journal note que l'Union Européenne est également fortement engagée dans d'autres conflits, qu'il s'agisse de l'Afrique ou du Moyen-Orient. Les journaux régionaux, quant à eux, n'apportent pas une vision supplémentaire de cette question.

Alexandre Loukachenko, le dernier dictateur d' Europe	(..) a été réélu (..) dans un scrutin entaché	fi050207
de voir ses relations avec les États-Unis et l' Europe	se détériorer, Moscou n'hésite plus à donner de	fi250407
Mais l' Europe	doit montrer que (..) elle est prête à mettre sur la table le problème de la légitimité de l'ordre mondial	fi280407
Cet organisme, qui dépend du Conseil de l' Europe	s'est ému de la persistance de la torture en Tchét	li140307
rak, su Guantanamo et sur l'action de la CIA en Europe	.	li120207
Tripoli exige que l' Europe	y participe de manière plus conséquente, ce que l'UE refuse	li090407
L' Europe	reste sourde à la pression américaine	su090207
En s'installant à l'est de l' Europe	, les États-Unis font également prévaloir leurs prét	ou210207
discours belliqueux contre les États-Unis et l' Europe	.	ou280307
L' Europe	a du mal à assurer la maîtrise de sa propre sécurité	ou210407

Ce contexte d'une entité qui doit défendre ses frontières au sein de tensions internationales est aussi activé à travers l'utilisation d'une locution « Europe de la défense » qui revient à de nombreuses reprises dans le corpus IntUne 2007. L'Europe de la défense, est alors avant tout une idéalisation qui est souhaitée par les candidats à l'élection présidentielle française, sans exister au moment de la collecte du corpus, elle est le plus souvent traitée au conditionnel ou à l'interrogative. D'autre part, elle semble poser un partenariat privilégié avec le Royaume Uni.

Cette locution connaît par ailleurs une forte progression entre 2007 et 2009. Il est alors essentiellement question du retour de la France au sein de l'OTAN ce qui peut être vu soit comme une difficulté, soit comme un pas en avant pour former cette « Europe de la défense ». Mais elle reste toujours ardemment souhaitée, notamment parce qu'elle permettrait à la France de retrouver un certain prestige et un pouvoir décisionnel. Cependant, des difficultés subsistent dans sa mise en œuvre et son fonctionnement réel n'est que très peu mentionné.

Lorsqu'il s'agit des rapports entre OTAN et « Europe de la défense », deux points de vue peuvent être développés dans la presse française. Un point de vue consiste à penser que ces deux entités sont antinomiques, et ne peuvent pas coexister ou se développer sans que ce

4- « Europe ». Quelle représentation dans la presse française ?

soit au détriment l'une de l'autre. L'OTAN forme alors une sorte de puissance américaine qui est concurrente d'une Europe de la défense. Néanmoins, le rapport OTAN / Europe de la défense peut être plus favorable, cette dernière entité étant appelée des vœux de gouvernants américains. D'autre part, le retour de la France au sein de l'OTAN est présenté comme une condition nécessaire sinon suffisante à la création de cette entité. Enfin, de nombreux pays partenaires au sein de l'Union Européenne semblent trouver que l'OTAN peut être une base pour la constitution d'une Europe de la défense. Quel que soit son rapport à l'OTAN, cette entité est souhaitée et largement idéalisée. Elle crée le consensus, mais n'existe que dans un futur hypothétique. Un des atouts donnés à cette entité est le fait que sa réalisation permettrait de redonner davantage de prestige et de pouvoir à la France. Cependant, de nombreuses difficultés surviennent dans le processus concret de création de cette entité, et cette réalisation apparaît à la période de collecte du corpus comme étant fortement compromise. Seules 3 occurrences extraites du Figaro citent cette entité comme étant déjà existante, et la présentent comme ayant un fonctionnement parfaitement complémentaire de celui de l'OTAN.

politique sans l'Allemagne, on ne peut faire l'	Europe	de la défense sans les britanniques.	ou070407
rien ne serait plus vain que de faire de l'	Europe	de la défense un moyen d'alléger les charges	fi050307
France renonce à tout espoir de construire une	Europe	de la défense indépendante. Prétendre que l'on	fi170309
certes pas une condition suffisante au succès de l'	Europe	de la défense, mais une condition nécessaire	fi210209
	L' Europe	de la défense : une chance à saisir	fi240309
enjeux sont nombreux : l'avenir économique et			
démographique de notre continent, l'	Europe	de la défense.	li190309
Alliance et de favoriser le développement d'une	Europe	de la défense.	ou120309
L'avenir incertain de l'	Europe	de la défense	li110309
	L' Europe	de la défense est aujourd'hui une réalité : elle	fi170209

Une fois de plus, on voit clairement que l'orientation politique des journaux du corpus influence grandement la façon dont ils activent le sens « entité politique qui doit défendre ses frontières » pour le concept « Europe ».

IV.2.6.2- Le problème de l'immigration.

Le corpus IntUne 2007 montre à travers l'analyse des lignes de concordances que l'Europe est une entité qui connaît une réelle réflexion au niveau des flux de population, dont le but serait sinon de renforcer, du moins de conserver une certaine imperméabilité à ses

4- « Europe ». Quelle représentation dans la presse française ?

frontières. Dans le journal Le Figaro, ces flux migratoires sont relevés comme étant situés essentiellement au sud de l'entité et comme posant de réels problèmes. Nonobstant, les flux économiques qu'ils engendrent sont cette fois abordés sous un angle plutôt positif, sorte d'aide au développement déguisée. Enfin, l'immigration peut aussi se faire à l'intérieur de la zone européenne, dans ce cas, les immigrés vont le plus souvent de « l'Est » vers « l'Ouest », de façon plus ou moins encadrée. Ouest France reprend cette idée d'une immigration massive, et posant de nombreux problèmes à l'Europe, sorte de fléau que l'entité cherche à endiguer par tous les moyens.

Le journal Libération adopte un angle de vue radicalement différent dans ce problème migratoire en se positionnant systématiquement au niveau des migrants, et en expliquant les causes de leur migration. La responsabilité de ce problème apparaît alors comme étant européen, tandis que l'entité est comparée à une forteresse. Sud Ouest reprend cet angle d'approche, notant que la condition des migrants africains est autrement plus critique que celle des migrants d'Europe de l'Est. Il apparaît également dans les colonnes de ce journal que loin d'être un problème, l'immigration va dans les années à venir être un besoin pour l'Europe, étant donnée sa faible démographie.

aucune restriction à l'arrivée de travailleurs de l' Europe de l'est lors de l'élargissement, le	fi260207
uros) par la banque mondiale, les fonds venant d' Europe représentent plus de trois fois l'aide au dév	fi220307
il n'est pas alarmiste de dire que la menace vise l' Europe du sud, et plus particulièrement la France.	fi130407
roblème de l'immigration clandestine à destination de l' Europe .	ou090207
placé en première ligne, avec l'Italie, de la forteresse Europe .	li160207
res qui aident (...)leurs fils à partir, Bové explique que l' Europe et la France ont une grande responsabilité	li260207
Comme il semble le comprendre et le dire, l' Europe , (...), va avoir besoin d'immigrés.	su060407

IV.2.6.3- La menace terroriste.

En tant qu'entité politique devant défendre ses frontières, le corpus IntUne 2007 montre en outre que l'« Europe » est menacée par le terrorisme. En règle générale, ce terrorisme est le fait de groupes islamiques intégristes, et de façon beaucoup plus minoritaire, il peut s'agir de régions demandant leur indépendance à un état dont elles feraient partie.

Dans Le Figaro, il est question de terrorisme islamique uniquement. Différentes situations sont évoquées, des attentats très meurtriers comme celui de Madrid sont rappelés,

4- « Europe ». Quelle représentation dans la presse française ?

mais une réflexion de fond est également menée sur ce problème. La question est de savoir comment ce terrorisme a pu se développer et où. Un rapport d'Euro Pol intervient dans ce contexte et insiste sur la menace du terrorisme islamique qui pèse sur l'Europe. Les autres types de terrorisme pouvant exister en Europe ne sont jamais évoqués dans ce journal. Les deux régionaux du corpus traitent de la même façon uniquement du problème du terrorisme islamique comme constituant une réelle menace pour la sécurité européenne.

L'Europe est aussi présentée dans Libération comme étant une entité qui connaît des difficultés avec le terrorisme. Une fois de plus le type de terrorisme qui est cité concerne essentiellement des groupes d'islamistes radicaux qui menacent l'Europe, mais cette fois, sont aussi menacés des pays qui se situent en dehors des frontières de l'entité et qui sont les premières victimes de ce terrorisme. D'autre part, ce journal montre l'Europe, comme un ensemble connaissant des attermoissements internes en particulier lors des manifestations d'autonomie voire d'indépendance de certaines régions. Il montre également l'aide apportée par l'Union Européenne dans ces processus séparatistes à l'instar de certaines régions du Royaume Uni:

dres est (.) le centre de l'extrémisme islamique en	Europe	, le hub des réseaux européens de la terreur	fi050207
du plus grave attentat islamiste jamais commis en	Europe	a débuté hier à Madrid	fi160207
informations plus fluides entre états membres, l'	Europe	reste un lieu privilégié de recrutement, de finan	fi120407
es sur les menaces qu'Al-Qaeda fait peser sur l'	Europe	.	li130407
pas vues d'un mauvais oeil, car l'	Europe	des régions est plus maniable qu'une Europe	li140407

IV.2.6.4- Protéger les frontières de problèmes sanitaires.

Dans un contexte plus particulier, l'Europe est également une entité qui doit protéger ses frontières d'un virus. Il s'agit alors dans les occurrences concernées du virus de la grippe aviaire, de type H5N1. Diverses mesures sont alors prises pour le contenir, le pays européen le plus touché par ce virus semblant être le Royaume Uni. Il est intéressant de voir que parmi les deux régionaux, Sud Ouest est le quotidien se préoccupant le plus de ce problème, cela peut s'expliquer par le fait que la région de diffusion du journal a une tradition d'élevage de volaille en plein air, directement menacée par la diffusion du virus.

ché, l'an dernier, quand le virus s'était répandu en	Europe	. »	ou050207
D'autres mesures sont à l'étude en	Europe	où quinze pays ont été touchés par le virus, l'an	ou050207
ilance face à la résurgence de la grippe aviaire en	Europe	. Il a invité le gouvernement à veiller aux mesure	li080207

4- « Europe ». *Quelle représentation dans la presse française ?*

idémie humaine pourrait aussi bien se produire en Europe qu'en Afrique a précisé, à précisé le porte parole su050207

Quand la presse française active le sens d'une entité qui doit défendre ses frontières, les agresseurs peuvent ainsi être multiples : terroristes (le plus souvent liés à la religion musulmane), virus, immigrants (venant du sud ou de l'est). Mais l'agresseur qui rencontre le plus grand nombre d'occurrences dans ce contexte est sans aucun doute « Moscou » via l'affaire des boucliers antimissiles.

IV.2.7- Une entité composée de plusieurs pays.

L'Europe, dans le corpus IntUne 2007 apparaît aussi comme étant une entité qui se compose de plusieurs pays. Ces pays peuvent adopter des comportements singuliers au niveau économique, ou dans diverses autres situations. Ce sont également des pays dans lesquels des personnes habitent, organisées en sociétés, et dans lesquels divers trafics peuvent avoir lieu. L'actualisation par la presse de ce sens potentiel « une entité composée de plusieurs pays » va ainsi être étudiée à travers l'analyse de ces différents contextes où ce sens potentiel peut apparaître. Dans une première partie, nous verrons que ces différents pays peuvent être montrés comme ayant des comportements différents au niveau économique. Dans une seconde partie, leurs comportements dans d'autres domaines tels que la presse française les souligne seront analysés. Enfin, nous verrons comment des sociétés prennent place dans ces différents pays, et comment notamment elles peuvent être plus ou moins sujettes à différents trafics.

IV.2.7.1- Des comportement différents au niveau économique.

Lorsque l'Europe est présentée comme une entité composée de plusieurs pays ayant des comportements différents au niveau économique, cet ensemble prend le visage d'une zone concurrentielle où les divers pays cités sont plus ou moins bons élèves par rapport à une moyenne européenne. La France est un de ces pays adoptant un comportement singulier. Selon l'orientation politique du journal du corpus dans lequel l'activation de ce sens potentiel se fait, le contexte peut sensiblement différer. Ainsi, dans Le Figaro, elle est essentiellement représentée comme étant un pays handicapé au niveau européen par une forte fiscalité ou une mentalité suspicieuse par rapport à la réussite, même si dans une moindre proportion, elle semble aussi bénéficier de quelques avantages (économie agricole puissante, bonnes

4- « Europe ». Quelle représentation dans la presse française ?

infrastructures, main d'œuvre qualifiée).

Le journal Libération accentue le fait que ce pays dispose d'une chaîne télévisée très rentable, bien qu'étant de médiocre qualité, le fort chômage des jeunes et ce journal relativise fortement l'importante ponction fiscale qui y serait réalisée. Le journal Ouest France donne quant à lui une vision de la France par rapport au reste de l'Europe qui oscille entre grandes performances et handicaps lourds. Ainsi, si les entreprises françaises réalisent de véritables miracles autant en termes d'innovation que d'exportation, et si les investissements globaux sont importants sur ce territoire, au niveau structurel, ce pays paraît malgré tout bien mal loti pour faire face à l'économie globalisée, et il ne travaille pas suffisamment.

ait alourdir l'impôt sur les sociétés le plus élevé d'	Europe .	fi160207
La France se situe dans la moyenne de l'	Europe es quinze, comme le rappelle un rapport du	li040507
et notre croissance est l'une des plus faible d'	Europe .	ou010307
Et nous sommes les seuls à la maîtriser en	Europe .	ou190407

L'Allemagne est aussi très présente dans ce corpus comme ayant un comportement économique la distinguant des autres pays de l'entité « Europe ». Dans Le Figaro, elle est le plus souvent représentée sous les traits du bon élève, on parle de « locomotive de l'Europe », notamment, pour décrire un État dont l'économie se porte bien malgré une réunification qui a été dure à réaliser dans les années 1990. Dans le journal Sud Ouest, c'est le comportement paradoxal de ce pays qui est noté : bien que produisant des voitures fortement polluées, elle s'inscrit par ailleurs dans une sorte de leadership européen dans le domaine environnemental. Ouest France donne à voir un pays très performant au niveau industriel, du fait, entre autres, de salariés très laborieux.

la croissance y est plus dynamique qu'ailleurs en	Europe (4,5%), le chômage, plus faible (moins de 6%)	fi060407
tion de véhicules (..) parmi les plus « sales » d'	Europe .	su140207
premier fabricant d'ordinateurs d'	Europe augmente le temps de travail de ses salariés sans hausse de salaire.	ou060207

L'Italie, dans les colonnes du Figaro est dépeinte de façon plutôt alarmiste : pays avec la plus forte dette, tarifs dans les télécommunications trop élevés, etc. Le journal Libération confirme largement cette vision du pays, en notant que les frais bancaires y sont les plus élevés d'Europe. L'image donnée par Sud Ouest de ce pays n'est guère plus favorable puisqu'il peut alors y être question d'un scandale financier, ou encore d'une place de leader obtenue au niveau européen dans le domaine de la contrefaçon.

4- « Europe ». Quelle représentation dans la presse française ?

Italie avec 200 milliards, soit 23% de tous les États d' Europe fi200207
L'Italie est incontestablement le leader en Europe pour le trafic de contrefaçons ou220307

La Grande Bretagne cumule les superlatifs dans Le Figaro. Elle est le meilleur élève en Europe, celui que l'on envie, devançant même la puissance américaine. Le journal Libération donne également de ce pays une image plutôt positive, puisqu'il s'agit de montrer à quel point il a su passer d'un statut économique peu enviable au statut très dynamique qu'il connaît au moment de la collecte du corpus, notons au passage que dans une des occurrences, le Royaume Uni y est traité linguistiquement comme s'il ne faisait pas partie de l'Europe. Le journal Sud Ouest reprend cette image favorable en donnant l'exemple d'une vente qui s'y est déroulée, attirant des acquéreurs du monde entier pour un montant record. Enfin, le journal Ouest France pose ce pays comme une alternative au modèle nordique qui est viable, pérenne et stable d'un point de vue économique, contrairement à la France.

Londres, si tant qu'il appartienne en propre à l' Europe est le plus puissant de la terre, en pas fi240307
L' Europe (..) avait un taux de croissance supérieur à celui li100307
de l'Angleterre
emaine de ventes à Londres, du jamais vu en Europe ! su100207
excellents élèves du moment en Europe : la Grande-Bretagne et l'Allemagne fi250407

L'Espagne est souvent comparée aux bons élèves de l'Europe pour la bonne santé et le dynamisme de son économie, portée notamment par le boom du secteur immobilier. Ce pays est même parfois comparé au modèle nordique, présentant là encore une alternative viable qui pourrait être suivie par la France.

L'Irlande est également montrée comme très dynamique à travers la comparaison des prix de l'immobilier dans sa capitale et ceux de la France. Une autre comparaison très favorable à ce pays se fait sur l'étalon du Luxembourg, puisqu'elle apparaît alors comme le deuxième pays le plus riche de l'entité européenne. La Pologne présente aussi un portrait mitigé, oscillant entre un marché porteur et dynamique, et un pays qui ne parvient pas à retenir sa main d'œuvre.

L'ensemble « Europe du Nord », est clairement considéré comme un modèle, modèle qui serait préféré à un autre, anglo-saxon, modèle également qui inspirerait les candidats à l'élection présidentielle. Ce comportement vertueux est même parfois considéré comme étant inscrit dans le code génétique de ses habitants.

4- « Europe ». Quelle représentation dans la presse française ?

La Roumanie est citée une seule fois comme ne devant pas devenir une « poubelle » pour les autres pays européens, ce qui la situe dans un ensemble tiers mondiste. En outre, un pays des Balkans dont on ne précise pas le nom est dépeint comme étant très pauvre, mais comme postulant pourtant à l'entrée dans l'Union Européenne.

Bilbao a(..) les valeurs économiques de l'	Europe	du Nord	su160209
ergante de Dublin est devenue la plus chère d'	Europe	, éclipsant l'avenue parisienne.	fi230307
Irlande(..) est le deuxième pays le plus riche d'	Europe	après le Luxembourg	ou190307
Pologne(..) une croissance de 12% qui n'était que de 3% en	Europe	de l'ouest.	fi250307
Après le départ d'un à deux millions de polonais vers l'	Europe	de l'ouest, la Pologne se retrouve	ou160207
Vous citez les démocraties du nord de l'	Europe	, c'est ce qu'ils font	li030507
Les Scandinaves ont ça dans les gènes (...) d'	Europe	. Pour survivre, c'est une nécessité	ou100407
La Roumanie ne doit pas devenir la poubelle automobile de l'	Europe	, martèle le premier ministre	li040407
pour ce pays des Balkans qui(..) est l'un des plus pauvres d'	Europe	.	su190207

Ces différents pays cités dans un contexte économique comme faisant partie de l'entité « Europe » nous montrent que pour la presse française, au niveau économique, l'Europe semble avant toute chose être composée de pays issus de sa partie occidentale (France, Royaume-Uni, Allemagne, Irlande, Italie, Espagne), tandis que peu de pays de l'Est sont cités (Pologne, Roumanie), et le plus souvent, dans des contextes peu favorables.

IV.2.7.2- Des comportements différents dans d'autres situations.

Les comportements de différents pays composant l'entité « Europe » sont décrits dans le corpus IntUne 2007 dans des domaines autres qu'économiques : il peut alors aussi bien s'agir de la politique sociale, ou des mœurs observées dans un pays en particulier de l'Europe. Le premier pays dont il est question est logiquement la France. Une fois de plus, les actualisations de ce sens varient en fonction des différents journaux. Dans Le Figaro, le portrait dressé de ce pays est plutôt paradoxal. En effet, l'accent est mis sur le fait que la France fait souvent cavalier seul au sein de l'Union, ou qu'elle a des dysfonctionnements au sein de cette même institution. C'est un pays où la liberté d'expression est en danger, où de nombreuses voitures sont brûlées chaque année, qui a peur de la mondialisation, et dont la police est mal classée. Pourtant, c'est aussi un pays qui a un système social efficace, excluant peu de personnes, un pays où le taux de fécondité se porte bien, qui est riche et dynamique au niveau culturel, dont le modèle laïque pourrait devenir l'inspiration d'une Europe plus tolérante. Le journal Ouest France dépeint aussi une France oscillant entre réussites (que ce

4- « Europe ». Quelle représentation dans la presse française ?

soit en termes de démographie ou de santé) et échecs (les exemples donnés concernent alors le budget de la justice ou le fonctionnement de la médecine générale).

Libération donne également de ce pays une image contrastée, mais selon des angles différents : les difficultés par rapport aux autres pays européens concernent alors l'université et la justice qui ont moins de moyens, la faible espérance de vie des ouvriers, un fort euro-scepticisme, une montée des discriminations et du racisme, la peur de l'économie de marché, et l'isolement au sein de cet ensemble. Dans une proportion bien moindre, ce journal montre aussi la France comme jouissant de quelques atouts par rapport à ses voisins dans des domaines comme la santé ou la culture. Le journal Sud Ouest brosse une peinture plus favorable à ce pays, très bien positionné au niveau des droits de l'homme, de la démographie, l'image en est un peu plus nuancée au niveau de la santé ou de l'emploi des seniors, enfin, elle possède de nombreux atouts tels que des villes dynamiques au niveau culturel, ou une riche tradition gastronomique.

Pourquoi la France semble-t-elle être la seule en Europe à souffrir d'un tel dysfonctionnement ?	fi230207
solète, on le voit : c'est peut être même l'avenir d'une Europe tolérante et pacifique.	fi030407
vie d'un ouvrier français est parmi les plus basses d' Europe .	li270307
baissé de 4% en France, l'un des pays leaders en Europe .	li100207
emploi des 55-64 ans en France, l'un des plus bas en Europe .	su190307
brevet de meilleur apprenti cuisinier de France et d' Europe .	su200307
budget de la justice en France ne soit plus le 29è d' Europe , derrière l'Arménie ou l'Estonie	ou140207

Le Royaume-Uni est aussi décrit dans des domaines autres qu'économiques comme ayant un comportement particulier au sein de l'Europe. Pour le journal Le Figaro, il s'agit essentiellement du bilan de Tony Blair, qu'il soit économique ou social au niveau intérieur, mais aussi de sa politique extérieure, par rapport à la guerre en Irak, ou à l'influence qu'il a eu sur l'Union Européenne. Le journal Libération donne de ce pays une image positive avec une seule occurrence concernant les bons résultats des chercheurs britanniques. Cette image positive est reprise par Ouest France où cette fois, c'est au niveau des habitudes culturelles que le Royaume-Uni se distingue favorablement de ses voisins européens.

terni le bilan européen du premier ministre, l' Europe d'aujourd'hui est plus libérale, anglophone et	fi020507
gne apparait aujourd'hui en avance sur le reste de l' Europe .	li260307
de fréquentations encore inférieur aux autres pays d' Europe . En Angleterre, c'est quasi culturel !	ou040407

Les journaux du corpus décrivent également le comportement singulier de la Pologne dans un contexte unique, celui des lois de ce pays sur l'avortement. Il est alors

4- « Europe ». Quelle représentation dans la presse française ?

associé au Portugal et à l'Irlande. Le contexte est largement défavorable dans le corpus IntUne 2007 français à ces trois pays. Pour le Portugal, la situation est même comparée à une survivance de l'ère dictatoriale. L'Irlande s'illustre par ailleurs par l'alcoolisation massive de ses jeunes, ce qui ne contribue pas à en donner une image plus favorable.

Malte, la Pologne et l'Irlande l'un des quatre pays d' Europe où l'avortement est quasi impossible ? li100207

Les Irlandais sont les champions d' Europe du binge drinking, ou biture express, p li150307

L'Allemagne apparaît comme un pays au comportement singulier dans Le Figaro, Libération et Sud Ouest. Le contexte varie selon le journal qui en parle : pour Le Figaro, le pays est porteur d'un réel dynamisme au niveau de la construction européenne, idée reprise dans Sud Ouest où ce pays est alors associé à l'Espagne et à l'Italie. Libération s'attarde davantage sur une fécondité très faible, et un constat quelque peu moins favorable.

L'Allemagne est devenue la chance d'avenir de l' Europe Merkel, contrairement aux dirigeants français li140207
gne connaît un des taux de natalité les plus faibles d' Europe et a enregistré un nouveau recul démogr li310307

Les pays situés dans le Nord de l'Europe comme les Pays-Bas ou la Norvège bénéficient globalement d'un contexte favorable lorsqu'ils se singularisent en Europe par un comportement particulier. Les Pays-Bas effectuent en effet des démarches d'avant garde dans le domaine de l'architecture ou de l'environnement. La Norvège est un modèle d'insertion pour les enfants handicapés, et au niveau politique, est capable d'élaborer des consensus constructifs plus facilement que les pays dits « latins » dont la France ferait alors partie. Ces pays sont donc des modèles à suivre dans différents domaines : environnement, santé, politique, etc. C'est dans cette zone qu'il fait le mieux vivre en Europe.

dans le pays le plus densément peuplé d' Europe . Des projets d'avant-garde ont été signés par li090207

Dans les démocraties du nord de l' Europe , 100% des enfants ayant un handicap sont sco li030507
où il fait le mieux vivre sont ceux du nord de l' Europe , Pays-Bas, Suède et Danemark en tête. ou150207

Pour conclure, arrêtons-nous un instant sur une occurrence qui traite de la Russie par rapport à l'Europe. En comparant la Russie à l'Europe comme un bloc, Libération en fait effectivement une entité extérieure qui a ici un fonctionnement différent au niveau politique :

Russie, il n'y a pas d'hommes politiques comme en Europe , appuyés sur leurs réseaux. li170207

4- « Europe ». Quelle représentation dans la presse française ?

IV.2.7.3- Une entité composée de sociétés.

Dans ce contexte d'une entité où évoluent des sociétés, des habitants, c'est un comportement relativement homogène qui est mis en valeur pour l'ensemble de l'« entité composée de plusieurs pays ». Cependant, des divergences peuvent apparaître au niveau des différentes législations en vigueur dans ces sociétés, ou encore au niveau de leur fécondité. Le journal Libération insiste sur une tendance législative inquiétante, alors que les autres journaux portent davantage l'accent sur les problèmes démographiques rencontrés dans les sociétés de cette entité. Il est intéressant de noter que dans certains cas, les divergences de ces comportements sociaux peuvent amener à un enrichissement mutuel.

des lois et des pratiques très disparates en Europe , entre l'Irlande où l'IVG est interdite, et le fi100207
minalisation du sida devient la règle, y compris en Europe . li120407
sud (...) vont devenir essentiels à la survie de cette Europe sans enfants su260307
Les discussions, en France, en Europe et dans le monde s'articulent autour d'un cons ou100307

Cette entité se composant de multiples sociétés, celles-ci sont parfois sujettes à divers trafics. Ce sont surtout les journaux régionaux qui mentionnent ce type de problèmes, exception faite toutefois de Libération. L'Europe est alors soit la zone dans laquelle ces trafics se déroulent, ou bien celle qui alimente financièrement les trafiquants, qu'il s'agisse des Farcs dans le cadre du trafic de drogue, de braconnage d'ivoire, ou de vol d'œuvres d'art.

Deux sociétés sont plus particulièrement enclines à ces problèmes selon Libération : le nord de l'Europe pour le trafic d'art, et l'Europe de l'Est pour tous les autres types de trafics. Sud Ouest renforce cette idée selon laquelle, des sociétés européennes, celles se situant en Europe de l'Est seraient beaucoup plus sujettes à tous types de trafics. Ouest France confirme ce tableau d'une Europe de l'Est comme zone dans laquelle de nombreux trafics ont lieu. Ce journal insiste par ailleurs sur les difficultés liées à la drogue dans le sud de l'Europe.

vingtaine de caïds du Sud-Est et de trafiquants d' Europe de l'est. li120207
difficultés de la lutte contre ce fléau en Europe , un sénateur s'est indigné que la Belgique et la li150307
En cause : des ventes d'armes d' Europe de l'est vers l'Angola. su290307
reviendrai jamais dans ce pays [l'Albanie] dont l' Europe ignore la brutalité. Il aurait été plus honnête de su090207
héroïne est en (...) progression constante Europe de l'est. ou010307
L'Espagne est la porte d'entrée, en Europe de la cocaïne sud-américaine et du haschisch ou160207

4- « Europe ». Quelle représentation dans la presse française ?

IV.2.8- Une entité sportive.

Le troisième sens potentiel dans lequel le mot « Europe » apparaît en termes de fréquence est celui d'une zone où l'on pratique le sport, d'une entité qui organise ainsi de fréquentes compétitions entre les divers pays la composant. Sans qu'il y ait de réelle surprise, le principal pays participant à ces compétitions est la France (le corpus y ayant été collecté). Les pays qui sont ensuite les plus abordés et les domaines dans lesquels il est question de l'Europe, considérée comme un bloc, varient assez peu d'un journal à l'autre et peuvent donner une indication intéressante pour savoir quels pays font réellement partie de cette entité. Dans une première partie, nous verrons comment la France est représentée par la presse lorsque celle-ci actualise le sens potentiel pour « Europe » d'entité sportive. Dans une seconde partie, nous verrons également quels sont les différents pays cités par la presse française comme composant cette entité sportive. Enfin, dans une dernière partie, nous verrons comment l'entité, elle-même, est présentée de façon générale dans le domaine sportif.

IV.2.8.1- La France dans l'Europe sportive.

Quand la France est présente au niveau européen dans le domaine sportif, les différents journaux en donnent une image relativement homogène avec cinq grands sports importants qui sont le football, le rugby, l'athlétisme, les sports aquatiques ou nautiques et le patinage artistique (cf. annexe 9). Lorsqu'il est question de football pour Le Figaro, la France est un pays qui a de bons résultats au niveau de son équipe nationale en Europe, différents clubs de ce pays aimeraient participer ou participent à la coupe d'Europe, cette coupe revêtant un caractère très valorisant. Toutefois, à demi mot, est présenté le fait que les clubs de football français jouant la coupe européenne sont plutôt moins bons que leurs homologues européens : il n'y a pas de honte à perdre contre une autre équipe européenne, et c'est un sésame qui n'est jamais obtenu. Libération donne une image beaucoup moins favorable à la France, ne reprenant pas les résultats de l'équipe nationale et s'attachant davantage aux déconvenues des clubs de football français (Lyon, notamment, meilleur club français lors de la collecte du corpus).

Sud Ouest est le seul journal qui fait apparaître le football en deuxième position,

4- « Europe ». Quelle représentation dans la presse française ?

derrière le rugby. Cela s'explique par la forte implantation culturelle du rugby dans la région de diffusion du quotidien. La coupe d'Europe est alors comme dans Le Figaro une compétition très valorisée, et les difficultés des clubs français dans ce domaine sont peu stigmatisées (par rapport aux échecs en rugby des clubs français dans ce même journal). L'équipe féminine nationale y est dépeinte comme obtenant de bons résultats, c'est le seul journal qui en parle. Dans le journal Ouest France, bien que les résultats des clubs français dans la compétition européenne ne soient pas réellement très flatteurs pour le football national, l'accent est essentiellement porté sur les bons résultats de petits clubs plutôt que sur le fait qu'aucun club français ne jouera parmi les huit meilleures équipes de la compétition. La plupart des occurrences dans ce journal (50 sur 53) donnent en conséquence une image positive du football en France.

ambassadeurs d'une France championne d' Europe et donnée comme favorite de la compétition.	fi050207
perdre face au double champion d' Europe n'aurait rien d'infamant.	fi070307
tasme lyonnais d'une progression vers le faite de l' Europe des clubs	li180407
En coupe d' Europe , rappelons-le, un 0-0 à domicile n'a rien d'un m	su150207
Bruno Bini, champion d' Europe avec les moins de 19 ans féminie en 2003	su240207
Bretons auront la volonté de bonifier ce résultat, l' Europe leur tendant encore les bras.	ou140407

Le rugby est le deuxième sport dont il est question en termes de fréquence dans le corpus. Trois grands contextes sont alors présents dans à travers la presse française, quel que soit le journal concerné. La coupe d'Europe est la plus présente à deux niveaux, du fait de son possible boycott dont la menace est brandie par les clubs anglais et français (qui sont ceux qui la gagnent généralement) ou alors des résultats d'équipes françaises dans cette coupe (et surtout de l'élimination du club de Biarritz). Enfin, il est aussi question de la coupe d'Europe, mais cette fois des équipes nationales, où la France est présentée comme une équipe de qualité par un procédé métonymique.

et anglais menacent toujours de boycotter la prochaine coupe d' Europe .	fi100207
Toulousain, il a remporté deux titres de champions d' Europe (2003 et 2005).	li060307
Dominer l' Europe une année de coupe du monde.	si260207
Nothampton, dimanche dernier, en quart de finale de la coupe d' Europe .	ou070407

La présence de l'athlétisme est moins empreinte de passions, les occurrences dans lesquelles la France est présente en association avec « Europe » sont alors plus factuelles et donnent des informations sur des résultats pour les championnats d'athlétisme d'Europe qui se sont déroulés lors de la collecte du corpus. Les résultats montrent dans les quatre journaux

4- « Europe ». *Quelle représentation dans la presse française ?*

concernés une bonne santé de l'équipe de France et ses diverses branches :

illeure performance de ces deux jours, laissant loin derrière Wignanitz, 3e aux Europe en 2005 su090407

Les sports aquatiques concernent essentiellement la natation (les occurrences sont alors quasiment exclusivement dédiées aux bons résultats de Laure Manaudou lors des derniers championnats d'Europe), mais aussi en fonction des préférences régionales le surf, la voile ou le canoë. Ce sont toujours des domaines où les sportifs français obtiennent de bons résultats qui sont cités.

deuxième femme sous la minute : 59s87 (record d' Europe). su280307

Enfin, les sports d'hiver concernent dans une large majorité le succès du patineur artistique Brian Joubert lors des championnats d'Europe s'étant déroulés au moment de la collecte du corpus. Ce succès est d'autant plus mis en relief qu'il intervient après une longue période d'échecs français dans cette compétition.

Le champion d' Europe et n°1 mondial, invaincu cette saison ou220307

IV.2.8.2- *Quels pays dans les compétitions sportives européennes ?*

L'annexe 10 donne un aperçu global des différents pays participant à une compétition européenne. Par ordre de fréquence, le premier pays ainsi considéré comme européen est le Royaume-Uni, avec un ensemble dans le corpus de 65 occurrences. Le football et le rugby sont les deux sports dans lesquels ce pays s'illustre le plus souvent. Dans les deux cas, les éloges sont unanimes : on parle du « meilleur championnat » de ligue 1 du monde pour le football, et d'équipes trustant également les premières positions en ce qui concerne le rugby. Dans ce dernier sport, le problème posé par l'éventuel boycott de la future coupe d'Europe par les clubs anglais et français se pose aussi avec insistance. L'image du sport britannique conserve cependant une forte valeur positive. Le Figaro modère néanmoins cette vision élogieuse en notant que ce pays connaît également un grave problème de violence récurrent dans ses stades.

Ses province terrorisent l' Europe entière du rugby su100207

a deux équipes anglaises en finale de la coupe d' Europe , c'est peut être que le rugby national n'a pas to su230407

la preuve que le football anglais est le meilleur en Europe s'est rengorgé Alex Ferguson qui a vécu mardi fi120407

ysel, en mai 1985, lors de la finale de la Coupe d' Europe entre Turin et Liverpool, et celui de Sheffield, qu fi290207

L'Italie est le deuxième pays européen cité au niveau sportif en termes de fréquence.

4- « Europe ». Quelle représentation dans la presse française ?

Ce pays est alors une nation majoritairement consacrée au football dans la presse française. L'image qui en est donnée y est pour le moins contrastée : problèmes de violence dans ses stades, clubs qui ne sont jamais aussi bons que lorsqu'ils comptent des sportifs français évoluant en leur sein, trucages des résultats... Cette image négative se retrouve dans la majorité des journaux du corpus, si ce n'est Ouest France qui en fait un portrait plus élogieux, mais il s'agit alors uniquement des bons résultats d'un club dans lequel un joueur breton évolue. Le cyclisme confirme l'Italie dans cette mauvaise fortune, puisque le sujet abordé dans chacune des quatre occurrences concerne les condamnations pour dopage d'Ivan Basso dans le cadre du Tour de France. Le rugby est le seul sport qui en donne une image un peu plus favorable, ce qui peut être étonnant car dans ce domaine, l'équipe nationale obtient pourtant de faibles résultats, se faisant régulièrement battre par l'équipe française.

Italienne était bien plus agressive que (...) ailleurs en Europe ?	fi050207
nt préféré à l'Italie. Jamais attribution d'un championnat d' Europe des nations n'aura fait l'effet d'une	su190407
Ivan Basso, qui a fait sa rentrée en Europe (...) Contador a été cité dans le dossier	ou160307
Yoann Gourcuff sur le toit de l' Europe avec Milan	ou020507

L'Allemagne et l'Espagne semblent connaître une fortune équivalente dans le domaine sportif. Ces deux pays sont essentiellement présents via leurs championnats de football, ils sont alors présentés comme possédant des équipes historiquement très fortes, mais qui ne trustent plus les premières marches des podiums européens du fait de l'hégémonie britannique. D'autre part, l'Allemagne est très présente dans le journal Sud Ouest du fait d'une équipe de tennis de table bénéficiant d'une organisation exceptionnelle. Quand l'Espagne est citée dans le domaine du cyclisme, c'est la propension des sportifs espagnols au dopage qui est alors décrite.

Un grand d' Europe quittera la compétition (...) Le Bayern	ou110407
pour sauver la face par rapport à l'Angleterre, reine de l' Europe	fi260407

L'Irlande vient ensuite avec des occurrences qui concernent exclusivement le monde du rugby. Son équipe nationale est présente dans tous les journaux du corpus comme ayant de très bons résultats au niveau européen, il en va de même pour ses clubs. Un de ses sportifs est néanmoins présenté comme violent à la suite d'une altercation avec un reporter. Ce pays bénéficie cependant d'une image largement favorable au niveau sportif européen.

hamière Stringer-O'Gara du Munster, champion d' Europe , associés à la flamboyante ligne de trois-quarts	fi100207
--	----------

La Lituanie ne peut pas prétendre à cette faveur. Si elle semble obtenir de bons

4- « Europe ». Quelle représentation dans la presse française ?

résultats en basket historiquement en Europe, l'accent est majoritairement porté dans les occurrences la faisant apparaître dans la presse française sur un problème de xénophobie, voire de racisme des supporters lituaniens qui ont déroulés une banderole à caractère raciste lors d'un match France/Lituanie, ce genre d'incident s'est déjà produit par le passé lors d'autres matchs de basket. Ce pays est donc montré de façon très défavorable dans le corpus IntUne 2007.

Championne d' Europe en 1937(..) la Lituanie a toujours été au sommet fi190307
avec une formule lapidaire : bienvenue en Europe . Une allusion raciste fi260307

La Pologne est présentée dans les différents journaux du corpus sous un jour essentiellement favorable : l'organisation de la coupe d'Europe de football en 2012 lui est confiée au détriment de l'Italie, ce qui inscrit pleinement ce pays en Europe. Le Figaro insiste notamment sur la « chance » offerte par le reste des États européens d'organiser cet événement. D'autre part, les très bons résultats d'un pongiste polonais évoluant dans une équipe française sont notés.

qui n'avait jamais organisé de championnat d' Europe de football, et qui était empêtrée dans ses rancœurs fi200407

La Serbie apparaît dans le même type de contexte que la Pologne : il s'agit alors d'un pays qui organise une compétition à l'échelon européen, cependant, la portée symbolique de cette organisation est moindre que pour la Pologne puisqu'il s'agit d'athlétisme et non de football. Le Portugal est montré comme un pays où le principal sport est le football, ses sportifs obtenant de bons résultats dans ce domaine. La Grèce est citée essentiellement au niveau de la coupe d'Europe de football, son équipe nationale ayant des difficultés à se qualifier pour cette coupe dont elle a pourtant gagné la précédente édition. Les Pays Bas, la Suisse, la République Tchèque, le Danemark et la Suède sont aussi présents, mais à travers des sports qui ne sont pas nécessairement marquants pour le public français, et dans des informations factuelles.

Enfin, l'Arménie, la Russie et Israël semblent bien faire partie de l'Europe puisque ces pays s'illustrent dans des compétitions européennes, qu'il s'agisse de basket, ou d'haltérophilie. Néanmoins, les sports alors concernés font partie des moins populaires au niveau européen, ce qui en fait des participants moins emblématiques de l'Europe du sport.

Tel Aviv que défilera en finale de la coupe d' Europe des clubs champions, appellation fi140207
Revenue au champion d' Europe Junior, l'Arménien Tigran su200407

4- « Europe ». Quelle représentation dans la presse française ?

Aux Russes de Samar, championnes d' Europe en 2005 et finalistes en 2006

su020407

IV.2.8.3- L'Europe sportive en règle générale.

Parfois, il n'est plus question de différents pays participant à une compétition européenne, mais de phénomènes sportifs qui touchent l'ensemble de l'Europe. Dans ce contexte, les deux thèmes les plus fréquemment abordés sont les problèmes rencontrés pour l'organisation de la coupe d'Europe des clubs de rugby, menacée par un boycott franco-anglais, et d'autre part la proportion croissante des violences dans les stades européens.

Le journal Libération introduit une dernière idée dans ce contexte : l'Europe du football serait devenue davantage une histoire de finance que de sport, ce qui créerait d'une part de forts déséquilibres budgétaires, et d'autre part, ferait perdre « son âme » à ce sport.

Partout en Europe , on constate ces derniers mois un regain de violence fi050207

L' Europe du football désormais administrée par Platini a perdu quelq li260307

Les clubs (..) qui trustent les titres en Europe dépensent sans compter et creuses des déficits abyssaux, à li170207

IV.3- Conclusion.

L'analyse d'un corpus de presse française tel que le corpus IntUne montre ainsi que l'Europe est un concept complexe, flou et fluctuant, qu'en fonction des structures et des contextes dans lesquels il apparaît, les significations qu'il peut recouvrir varient fortement. Parler d'Europe sans définir précisément de quelle entité il est question pose un réel problème d'interprétation. Face à une réalité aussi polysémique, l'approche par prototype semble la plus justifiée, et permet d'appréhender la multiplicité de sens recouverts dans le corpus.

Le prototype de Williams (2009) liste les sens suivants pour ce concept :

- *« Une masse terrestre.*
- *Situé dans l'hémisphère Nord.*
- *Entouré d'îles plus ou moins grandes.*
- *Une péninsule d'Eurafrasie.*
- *Une péninsule d'Eurasie.*
- *Un continent.*
- *Situé entre l'Asie et l'océan Atlantique.*
- *Délimité à l'ouest par l'océan Atlantique.*

4- « Europe ». Quelle représentation dans la presse française ?

- Délimité au Sud par la méditerranée.
- Délimité au Sud par la chaîne du Caucase.
- Délimité à l'Est par la mer Caspienne et l'Oural.
- Composé de plusieurs pays.
- Dont les pays sont regroupés dans des blocs géographiques.
- Dont les pays sont regroupés dans des blocs politiques.
- Partiellement composé d'une unité politique : l'Union Européenne.
- L'Union Européenne est divisée en sous-ensembles.
- L'Europe politique est une idéalisation.
- Composé de pays continentaux par rapport à la Grande Bretagne.»

(Williams 2009)

Les sens grisés sont clairement activés dans la presse française, tandis que les sens soulignés posent davantage de problèmes. L'Europe est bel et bien une masse terrestre à de fréquentes occurrences dans le corpus IntUne, son climat est parfois mentionné, les moyens de transports permettant de s'y déplacer également.

Cette entité se compose aussi de plusieurs pays dont, nous l'avons vu, les comportements peuvent diverger dans plusieurs domaines. La liste des pays qui composent le concept « Europe » est complexe à établir. Selon que l'on se situe dans un contexte économique, énergétique, sportif, culturel ou politique, la Turquie, l'Israël, la Russie, l'Arménie en feront partie ou non. Le fait que ces pays soient regroupés dans des blocs géographiques et politiques est indéniable et très présent dans le corpus IntUne, avec une progression de ces sens entre 2007 et 2009. A une Europe occidentale ou de l'Ouest, viennent s'opposer une Europe centrale ou de l'Est et une Europe orientale. L'image donnée de la première montre de nombreuses difficultés dans le corpus 2007, au profit de la seconde, mais ces difficultés s'inversent en 2009. D'autre part, la dimension politique qui donne à l'Europe de l'Est une réelle autonomie vis à vis de la Russie en 2007 tend à basculer également en 2009, à la faveur d'une influence sensiblement plus présente. L'Europe orientale est assez peu définie, et souvent amalgamée à l'ensemble Europe de l'Est, elle partage alors les mêmes déconvenues économiques en 2009, et connaît de nombreux trafics sur son sol en 2007 aussi bien qu'en 2009.

Le sens qui montre l'Europe comme « partiellement composée d'une unité politique : l'Union Européenne. » est aussi très présent dans les deux corpus. C'est le premier sens qui apparaît dans le corpus 2007. La division de l'Union Européenne s'active dans le corpus entre différentes entités selon le nombre de pays qui les compose, ou selon la

4- « Europe ». *Quelle représentation dans la presse française ?*

monnaie qu'ils utilisent, notamment. L'idéalisation de l'Europe politique apparaît avec le plus de vigueur en 2007, lors de la campagne présidentielle française, elle n'a pas disparue en 2009, mais semble néanmoins beaucoup moins prégnante dans le débat public. Dans tous les cas, cette Europe est fortement souhaitée, mais peu réelle, et le corpus 2009 fait apparaître que cette entité idéale est davantage l'apanage de la gauche, parti n'étant pas au pouvoir à l'époque de la collecte du corpus.

Le sens d'une opposition entre « Europe continentale » et Royaume-Uni est présent dans le corpus IntUne en 2007 ainsi qu'en 2009. Il montre notamment qu'à un bloc continental composé de la France et de l'Allemagne pour l'essentiel, s'oppose la Grande Bretagne, principalement au niveau de l'économie, et de la méthode envisagée pour la sortie de crise, mais aussi dans le domaine diplomatique ou social. Le Royaume-Uni se rapproche alors généralement des États-Unis.

Les trois sens soulignés sont présents dans le corpus mais posent des difficultés, en effet, comme nous l'avons dit plus haut, l'Europe est une entité dotée d'une géométrie extrêmement variable selon les contextes dans lesquels elle est citée, il semble dès lors risqué de définir ses frontières de façon fixe. Si la plupart des pays généralement considérés comme européens conviennent à ces frontières, il est possible que des doutes soient parfois émis quant à leur valeur absolue. Ainsi, parler de l'Europe sans décrire plus précisément quel concept on sous entend, ne pas inscrire ce mot dans un contexte d'utilisation, porte en soi le risque d'un décalage certain dans l'interprétation qui en sera faite entre le locuteur et l'interlocuteur. En cela, le prototype permet d'être davantage conscient des multiples réalités que peut recouvrir ce concept dans l'imaginaire d'une communauté, et de mieux comprendre ce qui peut parfois être sous tendu par son usage.

L'Europe, nous l'avons vu, est une entité composée de pays, ces pays varient, mais certains peuvent être jugés prototypiques de l'Europe, voire typiques de telle ou telle zone européenne. La représentation dans la presse française de ces pays permet en conséquence de mieux comprendre comment les citoyens français peuvent se représenter les autres citoyens qui forment avec eux l'ensemble Union Européenne. Dans le chapitre suivant, il sera procédé à l'analyse puis à l'élaboration de la présence de deux pays dans le corpus IntUne : l'Italie et la Pologne.

5- Deux cas particuliers : l'Italie et la Pologne.

V.0- Introduction.

Le projet IntUne a pour titre complet « *IntUne, Integrated and United ? A quest for citizenship in an ever closer Europe* ». La représentation de l'Europe dans la presse française a été abordée dans le chapitre précédent : on a vu à travers la constitution d'un prototype sémantique pour ce mot que la réalité qu'il recouvrait était pour le moins multiple. On peut en effet parler de plusieurs « Europes » bien différentes les unes des autres, dotées de traits propres clairement identifiés dans la presse française.

Si la question de l'intégration et de l'unité passe par des ensembles supranationaux, ces ensembles se composent avant tout au rang inférieur d'États, de nations, de pays. Ces trois mots touchent à des concepts proches, mais distincts, qu'il convient de préciser davantage. Leur définition dans un des dictionnaires les plus consultés en France est la suivante :

« État, n. m. 1- Entité politique constituée d'un territoire délimité par des frontières, d'une population et d'un pouvoir institutionnalisé. 2- Ensemble des pouvoirs publics ». (Larousse 1992:401)

« Nation n. f. 1- Grande communauté humaine, le plus souvent installée sur un

5- Deux cas particuliers : l'Italie et la Pologne.

même territoire et qui possède une unité historique, linguistique, culturelle, économique plus ou moins forte. 2- Communauté politique distincte des individus qui la composent, et titulaire de la souveraineté. » (Larousse 1992:671)

« Pays n. m. 1- Territoire d'une nation ; nation, État. Ensemble des habitants d'une nation. 2- Région, contrée envisagée du point de vue physique, climatique, économique, etc. Pays chauds (...) 3- Lieu, région d'origine. Mal du pays (...) 4- Village, localité. Un petit pays de deux cents habitants. » (Larousse 1992:736)

La définition du mot pays pose toute la proximité de ces trois concepts dans les définitions du Larousse, puisqu'il s'explique notamment par les mots « nation », « état ». Le mot « nation » porte également dans ses définitions une forte importance sur des concepts proches de l'identité à travers les mots « unité historique, linguistique, culturelle ». Parmi ces unités, le mot « État » semble être le plus neutre, celui qui se rapproche le moins du concept d'identité, et dont la définition porte moins sur les hommes qui constituent cette unité. En conséquence, il serait intéressant d'étudier quels sens les mots « pays » et « nation » recouvrent dans le corpus IntUne.

Des différentes « Europes » relevées dans le chapitre précédent, deux ensembles semblent clairement se distinguer : une Europe de l'Ouest et une Europe de l'Est (cf. partie IV.2.2), ces deux « Europes » posant clairement certains problèmes de coexistence et semblant pouvoir mettre en péril une unification totale de l'entité européenne. L'analyse de deux pays issus de ces deux ensembles permettrait alors de voir en quoi l'Europe est dotée de composants qui seraient plus ou moins intégrés et unis, et en quoi l'image fournie par la presse française de deux de ses membres serait homogène. Les deux pays choisis dans ce chapitre sont l'Italie et la Pologne. Ce choix obéit à différentes motivations.

L'Italie est le pays où a été signé le traité de Rome en 1957, acte de naissance de l'Union Européenne actuelle. D'autre part, il est généralement bien établi comme faisant partie de l'Europe de l'Ouest, et sa proximité culturelle ou géographique avec la France fait qu'il est largement présent dans la presse française. Ce pays est donc un bon candidat pour

5- Deux cas particuliers : l'Italie et la Pologne.

représenter l'Europe de l'Ouest dans la presse française.

La Pologne a intégré l'Union Européenne en 2004, lors d'un élargissement de cet ensemble à 8 nouveaux pays membres : l'Estonie, la Lettonie, la Lituanie, la République Tchèque, la Pologne, la Slovaquie, la Hongrie et la Slovénie. Cet élargissement, outre le fait qu'il ait été le plus important en Union Européenne (les élargissements précédents n'avaient jamais intégré plus de trois pays à la fois) intégrait des pays issus de l'ancien bloc communiste, et dont l'histoire s'était construite en opposition à l'Union Européenne pendant presque 50 ans. La Pologne représente le plus grand et le plus peuplé des États ayant participé à cet élargissement, c'est pourquoi elle est un bon représentant pour le deuxième sous ensemble européen. De plus, nous avons vu dans le chapitre précédent que dans la presse française, c'est le pays prototypique -au sens roschien du terme- appartenant à la catégorie « Europe de l'Est ».

Afin d'étudier l'intégration de différents sous ensembles dans l'Europe, ce chapitre va dans un premier temps concerner l'élaboration d'un prototype sémantique pour les mots « pays » et « nations ». Ce prototype sera construit de la même façon que dans le chapitre précédent. Un premier prototype sera proposé à partir de la confrontation de plusieurs sources lexicographiques traditionnelles, puis il sera confronté au corpus comme cela a été fait pour le mot Europe. Dans une seconde partie, nous verrons quelle compagnie les mots « Italie » et « Pologne » ont dans le corpus, et un prototype en sera produit à l'aide de ceux de « pays » et « nation », ainsi que par l'étude des réseaux collocationnels de ces deux pays. Enfin, dans une troisième partie, nous verrons comment les prototypes de ces deux pays sont plus particulièrement activés dans les corpus IntUne 2007 et 2009.

V.1- Qu'est-ce qu'un pays ou une nation dans la presse française ?

Les mots « pays » et « nation » sont ceux qui, dans le dictionnaire Larousse, donnent l'image la plus proche du concept d'identité des différentes entités qui composent l'Europe. Ces deux mots sont présents avec diverses fortunes au sein du corpus IntUne. « Nation » y est cité en 807 occurrences en 2007 et 339 occurrences en 2009. Le mot « pays » est présent à 15 448 reprises en 2007 et 13 319 reprises en 2009.

5- Deux cas particuliers : l'Italie et la Pologne.

De ces deux mots, le plus présent dans le corpus est de loin « pays ». L'analyse du corpus afin d'établir un prototype pour ce mot sera en conséquence menée dans un premier temps, puis dans un deuxième temps, il sera procédé à l'élaboration d'un prototype pour le mot « nation » dans la presse française.

V.1.1- Élaboration du prototype de « pays » dans la presse française.

Le mot « pays » apparaît davantage dans le corpus 2007 que dans le corpus 2009 si l'on prend en compte sa fréquence par rapport au nombre total de mots de ces corpus. Il est intéressant de voir si cette tendance à la baisse de l'usage du mot a été suivie d'une tendance à un usage de sens différent entre les deux corpus IntUne. Pour cela, nous allons dans une première partie procéder à l'élaboration d'un premier prototype sémantique qui se basera sur les ressources lexicographiques traditionnellement disponibles. Puis, nous verrons quels sont les réseaux collocationnels qui peuvent s'organiser autour de ce mot dans le corpus IntUne en 2007 et 2009. Nous pourrions alors voir s'il y a une modification dans ces réseaux entre les deux années de collecte du corpus. Enfin, une troisième partie nous permettra d'établir un prototype sémantique pour le mot pays.

V.1.1.1- Qu'est-ce qu'un « pays » selon les dictionnaires ?

Différentes ressources lexicographiques donnent du mot « pays » des définitions faisant apparaître les sens suivants :

« 1- Territoire d'une nation ; nation, État. Ensemble des habitants d'une nation.
2- Région, contrée envisagée du point de vue physique, climatique, économique,
etc.

3- Lieu, région d'origine.
4- Village, localité. » (Larousse 1992)

« 1- Territoire habité par une collectivité et constituant une réalité géographique dénommée ; nation.

2- Les gens, les habitants du pays.

3a- Le pays de quelqu'un, patrie à laquelle on appartient par sa naissance. b- Le pays de quelque chose, terre d'élection, milieu particulièrement favorable à, riche en.

4- Région géographique plus ou moins nettement limitée, considérée surtout dans son aspect physique.

5- Petite ville, village. » (Petit Robert 2008)

5- Deux cas particuliers : l'Italie et la Pologne.

« A- Division territoriale habitée par une collectivité, et constituant une entité géographique et humaine. Dans une acception large, nation, État. Dans une acception restreinte, partie plus ou moins étendue d'une nation : province, région, canton.

B- Nation, région à laquelle on appartient par la naissance. Dans une acception large, nation d'origine, patrie. Dans une acception restreinte, province natale, village natal.

C- Par métonymie : ensemble des habitants, des citoyens d'une nation.

D- Au figuré : domaine fabuleux, irréel, généralement peuplé d'êtres et de choses imaginaires, rêvés. » (Trésor de la Langue Française Informatisé, TLF⁷²)

Le dictionnaire Larousse nous fournit une définition du mot « pays » se composant d'essentiellement cinq sens : le pays peut être le territoire d'une nation, ou la totalité des habitants de cette nation, il peut s'agir d'une région considérée du point de vue de son histoire, son climat, son économie, etc. Il peut s'agir d'un lieu de naissance, d'une unité administrative : un village ou une petite ville. Le deuxième dictionnaire, le Petit Robert, nous donne une autre définition. Il y a encore cinq sens principaux. « Pays » est un territoire habité par une communauté, et par métonymie, c'est aussi cette communauté. Il peut s'agir d'un lieu de naissance, d'un lieu particulièrement favorable à quelque chose, d'une région géographique, d'une petite ville ou d'un village. Le TLF⁷² reprend ces différents sens en les organisant différemment, d'autre part, il ajoute l'usage fréquent au figuré du mot pays dans le sens de « domaine fabuleux », par exemple, un « pays des merveilles », « pays du sommeil » ou « pays de cocagne ».

Dans ces trois définitions, on retrouve le sens de territoire, avec des frontières, d'une entité géographique, d'un lieu où se trouve une communauté, d'une unité administrative, d'un village et d'un lieu de naissance. On peut également noter que le Larousse qui est un dictionnaire de nature prescriptive utilise le mot « nation » très fortement marqué au point de vue sémantique et politique, tandis que le Petit Robert, plus récent préfère un « communauté » plus neutre. Le TLF⁷² utilise les mots « collectivité » et « nation », formant une sorte de synthèse de ces deux définitions.

A partir du moment où un prototype sémantique liste un ensemble de sens potentiels qui doivent être simples et aisément traduisibles dans d'autres langues, l'utilisation de dictionnaires étrangers peut être une aide dans la construction de ce prototype. Le dictionnaire Oxford Advanced Learner's Dictionary (OALD) donne du mot « country » les

72 <http://atilf.atilf.fr/tlf.htm>

5- Deux cas particuliers : l'Italie et la Pologne.

sens suivants, sensiblement différents de ceux trouvés dans les dictionnaires français :

« 1- An area of land that has or used to have its own government and laws. 2- (often following an adjective) An area of land, especially with particular physical features, suitable for a particular purpose, or connected with a particular person or people. 3- (sing.) The people of a country, the nation as a whole. 4- (sing.) Any area outside towns and cities, with fields, woods, farms, etc. » (OALD⁷³)

Cette définition est relativement plus courte que les définitions françaises, moins de sens y sont distingués, le pays est alors une étendue de terre qui a une existence politique, ainsi que différentes caractéristiques physiques ou culturelles. Le mot peut aussi représenter les personnes de ce pays, dans ce cas, l'OALD utilise le mot « nation ». Enfin, le quatrième sens fourni par cet ouvrage n'est pas du tout présent dans les dictionnaires français. Cela peut s'expliquer par le fait qu'en France, il y a un autre mot pour désigner ce concept d'étendues entre les villes qui est “campagne”, en conséquence, le mot “pays” n'est pas utilisé dans ce cas. Les sens un et trois peuvent représenter les premier et quatrième sens des dictionnaires français, tandis que le troisième sens peut correspondre au second dans le Larousse (région considérée du point de vue de son climat, de son histoire, son économie, etc.), mais il en diffère sensiblement du fait de l'emploi du mot « particular ». Il peut aussi inclure le troisième sens du Petit Robert qui traite de l'origine.

Le “MEDAL”, Macmillan English Dictionary for Advanced Learners, basé sur corpus présente également une définition sensiblement différente du mot “pays” par rapport aux dictionnaires français :

*“1- An area of land that has its own government and official borders.
2- Areas away from cities and large towns.
3- An area that is known for a particular product, activity, person...” (MEDAL)*

Le MEDAL nous donne 3 sens : une étendue de terre qui a son propre gouvernement et des frontières officielles, une étendue entre les villes, et une étendue qui est connue pour une activité, un produit ou une personne particulière. Cette définition est nettement plus courte que celles du Larousse ou du Petit Robert. Cela peut être un avantage (plus une définition est concise, plus elle est claire) ou un inconvénient (moins une définition est longue et moins elle est précise et signifiante). Dans ce dictionnaire, le premier sens inclut le premier et le quatrième sens des dictionnaires français (territoire d'une nation et unité administrative), excluant le sens métonymique de pays comme la communauté qui y

73 http://www.oup.com/oald-bin/web_getald7/index1a.pl

5- Deux cas particuliers : l'Italie et la Pologne.

vit.

Le troisième sens est plus difficile à analyser car plus complexe. Il peut correspondre au second sens du OALD. Une des questions ici repose sur l'absence du sens « communauté » dans le MEDAL, par rapport aux dictionnaires français. Une explication pourrait être que le Larousse et le Petit Robert sont des dictionnaires basés sur des méthodes intuitives et prescriptrices, tandis que le MEDAL est basé sur corpus, or Hanks notait (Hanks 1994) que l'être humain a tendance à enregistrer l'inhabituel plutôt que l'habituel.

A l'aide de ces informations, un premier prototype lexicographique peut être élaboré pour le mot « pays » :

Un pays est :

- Une étendue de terre physique.
- Une étendue de terre avec un gouvernement et des frontières officielles, un lieu administratif.
- Une étendue où se trouvent des habitants.
- Un lieu de naissance, d'origine.
- Une étendue qui peut se considérer du point de vue de son histoire, son économie...
- Une petite ville, un village.

V.1.1.2- « Pays » dans le corpus IntUne.

La répartition de ce mot dans le corpus IntUne 2007 s'effectue de la façon suivante :

Pays dans IntUne 2007					
Source	occurrences	Total de mots	%	Textes avec « pays »	Total des textes
Les Échos	4314	5015207	0.086	63	63
Le Figaro	4098	4861988	0.084	77	77
Libération	2715	3602857	0.075	77	77
Sud-Ouest	1640	3627131	0.045	76	76
Ouest France	2351	3226744	0.073	77	77
TF1	185	251551	0.074	31	32
France 3	145	183741	0.079	31	34
Total	15448	20769219	0.074	432	436
Pays dans IntUne 2009					
Les Échos	4282	4473616	0.096	63	63

5- Deux cas particuliers : l'Italie et la Pologne.

Le Figaro	3985	4858636	0.082	77	77
Libération	2279	2973339	0.077	75	75
Sud-Ouest	1521	3786883	0.040	73	73
Ouest France	1252	2146102	0.058	77	77
Total	13319	18238576	0.073	365	365

(Figure 17, répartition du mot « pays » dans le corpus IntUne)

Le journal de presse française qui mentionne le plus, proportionnellement ce mot est Les Échos, tandis que celui dans lequel il est le moins présent est Sud Ouest, cet écart se creusant entre 2007 et 2009. Il est présent dans chacun des journaux et chaque jour, c'est un mot important du corpus. Pour élaborer le prototype du mot « pays », l'analyse du corpus va maintenant se baser sur deux étapes essentielles. L'étude des collocations de ce mot d'une part, permettant de mettre en relief les différentes thématiques associées à ce mot. Et l'étude des structures les plus saillantes dans lesquelles il apparaît d'autre part. Ces structures ont été identifiées dans un échantillon aléatoire de 100 occurrences en notant les mots présents à plusieurs reprises en position de N-1 et de N+1 par rapport à pays. Puis, chacune d'entre elles a été reprise selon sa signification en relation avec les ensembles thématiques distingués à l'aide des collocations.

L'étude des collocations donne des informations sur « la compagnie » que garde un mot, dans Xaira, selon la méthode déjà utilisée, les collocations du mot « pays » sont les suivantes dans le corpus IntUne :

2007			2009		
Collocations	Fréquence	z-score	Collocations	Fréquence	z-score
émergents	368	244,9	émergents	340	246,1
basque	249	148,4	basque	248	155,6
galles	127	148,1	loire	196	155,5
loire	219	146,6	développés	154	153,4
industrialisés	81	130,9	galles	153	149,5
dans	3393	113,4	industrialisés	58	112,6
européens	335	99,4	dans	2764	100,2
développés	99	97,1	européens	318	97,8
pauvres	155	83,4	riches	157	90,9
notre	638	82,1	pauvres	152	90,7
nordiques	38	71,1	baltes	43	83,4

5- Deux cas particuliers : l'Italie et la Pologne.

baltes	38	70	occidentaux	86	67,8
scandinaves	44	69,2	les	3421	58,3
du	2915	68	du	2329	55,8
riches	111	61,4	arabes	72	54,9
arabes	90	60,9	nordiques	23	54,5
lorient	254	57,9	autres	452	51
anglo-saxons	49	56,2	anast	11	49,9
occidentaux	73	55,6	notre	376	49,3
civilisé	25	54,7	natal	27	48,9
exportateurs	38	53	africains	54	48,7
où	523	51,9	anglo-saxons	35	43,6
les	3427	50,6	scandinaves	19	42,3
africains	65	49,1	exportateurs	25	41,9
autres	475	48,4	auge	12	41,3
natal	29	47	où	407	39,2
ocde	47	46,7	membres	143	39,2
voisins	86	45,8	golfe	47	38
origine	168	44,7	dogon	11	36
bigouden	15	43,7	europe	227	35,3

(Figure 18, trente premières collocations du mot « pays » dans le corpus IntUne)

Une grande majorité des collocations reste identique entre les deux années de collecte du corpus, ce qui confère une certaine stabilité sémantique au mot pays. La dimension économique du concept est indubitablement présente à travers les collocs « émergents », « industrialisés », « développés », « pauvres », « riches », etc. colorés en bleu turquoise dans la figure 18. En cela, le mot se rapproche de « État » tel qu'il est défini dans les dictionnaires.

D'autre part, la notion d'étendue géographique est marquée par des mots comme « dans », « où », colorés en rose vif et dans une certaine mesure, par les collocs « nordiques », « occidentaux », colorés en rose saumon. Précisons cependant que ce dernier type de collocat, bien que se construisant sur une appellation géographique, peut se baser sur des critères largement culturels ou économiques. Les collocations qui rendent compte d'une étendue géographique montrent un mot fréquent qui est « du ». Cette préposition apparaît dans la colligation « du pays », qui est utilisée pour situer un événement d'une façon géographique. Les journaux parlent essentiellement de faits de guerres dans ces occurrences,

5- Deux cas particuliers : l'Italie et la Pologne.

ainsi, « du pays » est utilisé massivement pour l'Afghanistan, l'Irak, le Soudan. Quand « du pays » est utilisé pour désigner une étendue géographique de façon plus précise, il apparaît dans la vaste majorité des cas que le pays en question a des problèmes de terrorisme ou est en guerre (cf. annexe 11). Le fait que les journaux préfèrent généralement traiter de problèmes peut être une explication à ce phénomène. Néanmoins, Les Échos, qui accorde moins d'importance à ce type de faits est aussi concerné par ce comportement, puisque quand il est question de situer quelque chose à l'aide de la colligation, les informations sont le plus souvent mauvaises. Par ailleurs, en 2009, ce type d'usage apparaît toujours avec la même importance.

Au Darfour, région de l'ouest du pays , en proie à la guerre civile	fi170307
Dans le nord du pays , 105 personnes sont tuées	li050207
èvement de 2 français et de 3 afghans dans le sud ouest du pays .	su060407
fanatiques hindous avaient détruit une mosquée, au nord du pays . Des émeutes inter-religieuses s'en ét	ec160409
400 soldats supplémentaires, principalement dans le sud du pays , a annoncé hier le ministre de la défen	ec270407

Une entité politique est également présente, notamment avec le mot « européen » qui concerne les pays qui sont membres de l'Union Européenne. Ce type de collocats est coloré en bleu foncé dans le tableau. Une façon intéressante de dépeindre un espace politique dans la presse française est alors l'usage de la formule « notre pays ». Il y a 607 occurrences de cette locution dans le corpus, et le journal qui l'utilise le plus est Le Figaro avec 208 occurrences, puis Libération avec 123 occurrences, tandis que Les Échos n'apparaît qu'en troisième position. Cela peut s'expliquer par le fait que ces deux premiers médias ont une forte orientation politique et que la formule est particulièrement appréciée par les politiciens. Cette colligation leur est en fait très utile parce qu'elle accentue le fait que les pays leur appartiennent ainsi qu'à la communauté des votants. Elle est, assez logiquement, beaucoup moins présente dans le corpus 2009, ce qui s'explique par le fait que ce deuxième corpus n'a pas été collecté au moment d'une échéance électorale jugée importante pour la France. Le pays, lorsque cette formule est utilisée, a différents problèmes.

Notre pays est en train de « décrocher sous anesthésie »	fi240307
Notre pays est dans une situation difficile et effectivement menacé	fi370307
ausse du chômage intervient alors que notre pays fait	fi030209
Se pose à l'ensemble de notre pays . Faut-il pour autant les ignorer ? Je m'y refuse. Sans	li050209

A partir du moment où des problèmes sont exposés, des solutions doivent être suggérées, cela peut arriver mais avec une importance moindre en termes de fréquence :

5- Deux cas particuliers : l'Italie et la Pologne.

aurer pour pouvoir lancer les réformes utiles à notre pays .

fi130307

Toujours dans cette thématique politique, « Pays » peut aussi être utilisé pour désigner une entité qui appartient à un groupe avec une réglementation politique commune. Sans surprise, le groupe le plus présent est l'Union Européenne. En 2007, tous les pays en faisant partie peuvent y avoir une bonne croissance économique à l'exception de la France, et la Commission européenne peut aussi connaître certaines difficultés dans ses relations avec les habitants des pays membres de l'Union. En 2009, le contexte d'élection législative européenne apparaît dans ce cadre, d'autre part, la crise économique fait ressortir de lourds problèmes d'entente entre les pays la composant.

croissance dans les principaux pays de l'Union en 2007 à l'exception de la France	ec190207
xelles et les opinions publiques des principaux pays de l'Union	fi240307
nne de présenter la liste dans l'ensemble des pays de l'Union européenne. Sous ce même drapeau,	li2110209
FMI décoche (...) une flèche en direction des pays de l'Union européenne dont le manque de coordinati	ec230409

Le collocat « africain » peut également recouvrir une zone où différents pays ont des rapports politiques très forts, dans ce cas, ce sont surtout des endroits où les dirigeants semblent montrer une tendance assez nette à la corruption dans la presse française, et qui refusent un arrêt de la Cour Pénale Internationale dans le cadre du conflit au Darfour. Mais il peut aussi s'agir d'une étendue économique dont les différents pays ont un comportement assez semblable. C'est alors essentiellement la pauvreté de la zone qui est mise en avant, et les conséquences de la crise internationale sur l'aide qui lui est prodiguée par le monde occidental.

crise financière où les pays africains redoutent de voir les pays occid	fi100209
Khadafi, qui préside l'Union africaine, veut convaincre les pays ayant adhéré à la CPI de s'en retirer	li050309

La dimension culturelle est présente à travers les collocats colorés en jaune qui montrent une entité complexe. Des pays peuvent alors être regroupés selon un point commun, par exemple langagier (« anglo-saxon », « arabes ») ou économique et religieux (les pays du « golfe » sont majoritairement exportateurs de pétrole et ont une population musulmane). Un autre type de collocat rend également compte de cette thématique culturelle, cette fois, les pays sont compris en tant qu'entité infra-nationales, les collocats concernés sont colorés en gris dans le table. Le mot « basque » est, par exemple, une partie de la locution « pays Basque » qui désigne une nation qui n'est pas reconnue et n'a pas d'existence officielle. Le mot « Loire » est utilisé essentiellement dans l'expression « pays de

5- Deux cas particuliers : l'Italie et la Pologne.

Loire » qui désigne un endroit où l'on trouve de magnifiques châteaux de style Renaissance ou de bons vins.

L'unité administrative est présente via différents collocats colorés en vert. « Lorient » désigne par exemple une entité administrative qui est particulière à l'organisation française et qui concerne une étendue dépendant d'une ville « le pays de Lorient » regroupant plusieurs communes. Le collocat « Galles » présente également une unité administrative reconnue, mais fonctionnant cette fois au Royaume-Uni. Le lieu d'origine est présent via le pronom « notre » et le collocat « natal » ou « origine ». Enfin, cette idée d'origine est reprise dans les collocats « autres » ou « voisins », colorés en jaune pâle, qui montrent que ce dont on n'est pas originaire s'exprime par une altérité plus ou moins proche : les lieux d'où l'on ne vient pas sont d' « autres pays ».

Ces collocations font émerger différentes thématiques, différents sens pour le mot pays. D'autre part, nous avons vu que ceux-ci pouvaient se relier les uns aux autres. En conséquence, après ce panorama mené à partir de l'étude des collocations, complété par l'analyse des structures les plus saillantes dans lesquelles le mot pays apparaît, le prototype lexicographique préalablement construit peut évoluer de la façon suivante :

Un pays est :

- Une étendue économique.
- Une étendue géographique.
- Un espace politique.
- Une unité administrative plus ou moins grande.
- Un lieu d'origine.
- Une entité complexe constituant une réalité culturelle.

Maintenant que nous avons réalisé un prototype et que nous en avons brièvement étudié l'actualisation dans la presse française en 2007 et 2009, nous allons pouvoir voir comment il peut s'articuler à celui de « nation ».

V.1.2- Élaboration du prototype de « nation » dans la presse française.

Le mot « nation » est beaucoup moins utilisé que « pays » dans le corpus IntUne. 807 occurrences sont présentes en 2007 et 339 sont citées en 2009. Ce mot est beaucoup plus présent en 2007, notamment du fait de la campagne électorale et de son fort poids sémantique politique.

V.1.2.1- Qu'est-ce qu'une « nation » dans différents dictionnaires ?

Les ressources lexicographiques utilisées plus haut donnent les définitions suivantes du mot nation :

« Nation n. f. 1- Groupe d'hommes auxquels on suppose une origine commune. 2- Groupe humain, généralement assez vaste, qui se caractérise par la conscience de son unité (historique, culturelle) et la volonté de vivre en commun. 3- Groupe humain constituant une communauté politique, établie sur un territoire défini ou un ensemble de territoires définis, et personnifiée par une autorité souveraine. 4- Ensemble des individus qui composent ce groupe. » (Petit Robert 2005)

« Nation n. f. 1- Grande communauté humaine, le plus souvent installée sur un même territoire et qui possède une unité historique, linguistique, culturelle, économique plus ou moins forte. 2- Communauté politique distincte des individus qui la composent, et titulaire de la souveraineté. » (Larousse 1992:671)

« Nation, subst. fém. A- Groupe d'hommes dont les membres sont unis par une origine réelle ou supposée commune et qui sont organisés primitivement sur un territoire.

B- 1. Groupe humain, généralement assez vaste, dont les membres sont liés par des affinités tenant à un ensemble d'éléments communs ethniques, sociaux (langue, religion, etc.) et subjectifs (traditions historiques, culturelles, etc.) dont la cohérence repose sur une aspiration à former ou à maintenir une communauté.

2. Groupe humain stable, établi sur un territoire défini constituant une unité économique caractérisée par une auto-conscience ethnique (marquée par l'idée de la communauté d'origine et de destinée historique), une langue et une culture communes, formant une communauté politique personnifiée par une autorité souveraine et correspondant à un stade évolué du mode et des rapports de production.

3. Par métonymie : territoire occupé par ce groupe humain.

C-1. Ensemble des personnes qui composent ce groupe (cf. B2)

D- Désigne une collectivité par opposition à un individu, une classe, un groupe. Collectivité nationale. » (TLFI)

5- Deux cas particuliers : l'Italie et la Pologne.

« **Nation: noun**, 1- A country considered as a group of people, with the same language, culture and history, who live in a particular area under one government. 2-All the people in a country. » (OALD)

De nombreux sens potentiels communs émergent de ces quatre définitions. Un premier prototype lexicographique pour le mot nation pourrait être le suivant.

Une nation est :

- un pays considéré comme un groupe humain.
- un groupe humain d'une origine commune supposée.
- un groupe humain généralement assez vaste.
- un groupe humain qui se caractérise par la conscience de son unité (historique, culturelle).
- un groupe humain qui se caractérise par une langue commune.
- un groupe humain qui se caractérise par une religion commune.
- un groupe humain qui a la volonté de vivre en commun.
- un groupe humain constituant une communauté politique.
- un groupe humain établi sur un territoire défini.
- un groupe humain établi sur un ensemble de territoires définis.
- un groupe humain personnifié par une autorité souveraine.
- une communauté politique distincte des individus qui la composent.
- l'ensemble des individus qui composent ce groupe.

Maintenant que nous avons un premier prototype lexicographique pour le mot « nation », nous allons pouvoir le confronter à l'actualisation de ce mot dans le corpus IntUne. Une analyse détaillée des collocations et des contextes dans lesquels elles apparaissent permettra de voir comment ce concept est présent dans la presse française en 2007 et en 2009.

V.1.2.2- « Nation » dans le corpus IntUne.

Le mot se répartit de la façon suivante dans le corpus IntUne 2007 :

Nation dans IntUne 2007					
Source	Occurrences	Total de mots	%	Textes avec nation	Total de textes
Le Figaro	290	4861988	0.006	74	77
Libération	196	3602857	0.005	58	77
Les Échos	130	5015207	0.003	48	63
Sud-Ouest	71	3683377	0.002	38	77
Ouest France	97	3226744	0.003	49	77
TF1	11	251551	0.004	8	32

5- Deux cas particuliers : l'Italie et la Pologne.

France 3	12	183741	0.007	9	34
Total	807	20769219	0.004	284	436
Nation dans IntUne 2009					
Le Figaro	133	4473616	0.003	56	63
Libération	88	4858636	0.003	52	77
Les Échos	70	2973339	0.002	40	75
Sud-Ouest	26	3786883	0.001	21	73
Ouest France	22	2146102	0.001	19	77
Total	339	18238576	0.002	188	365

(Figure 19, répartition du mot « nation » dans le corpus IntUne)

Les deux journaux quotidiens politiques sont ceux qui font le plus large usage de ce mot. On peut dès lors se demander dans quelle mesure le mot nation ne désigne-t-il pas avant toute chose une réalité politique. Une recherche des collocations de ce mot dans le corpus IntUne avec Xaira donne les trente premiers collocats suivants :

2007			2009		
Collocations	Fréquence	z-score	Collocations	Fréquence	z-score
pupilles	3	50,7	live	22	161,6
la	645	49,2	ticketmaster	5	131,3
exaltation	4	45,5	la	232	25,8
symboles	9	42,4	fondamentaux	4	24,1
intégrante	4	35,5	entière	5	22,6
ethnie	3	30,6	terres	4	17,7
conceptions	4	29,2	indépendante	3	16,1
agonie	3	25,5	concerts	3	12,6
fondent	3	23,9	reconnaissance	3	12,3
république	21	21,9	père	6	11
musulmane	4	21	une	58	9,6
discours	17	20,9	seule	6	9,5
unie	3	19,8	football	4	9,4
patrie	4	19,2	sauver	3	9,3
islam	6	18,1	américaine	5	8,8
nationalisme	3	17,7	toute	8	8,6
menacée	4	16,8	de	172	8,4
symbole	6	15,3	petite	5	8,3
entière	5	15,2	économique	7	8,1
une	139	15	menace	3	7,8

5- Deux cas particuliers : l'Italie et la Pologne.

monopole	4	14,9	classement	3	7,7
race	3	13,3	supérieur	3	6,8
fermée	3	13,1	grande	6	6,5
identité	9	12,8	ordre	4	6,4
vives	3	12,8	veulent	3	6,4
grande	18	12,7	sous	7	5,8
turque	3	12,4	reste	6	5,6
unité	6	12,3	sociale	3	5,5
idée	12	12	république	3	4,9
ouverte	6	10,9	notre	6	4,5

(Figure 20, collocations du mot « nation » dans le corpus IntUne)

Ces différents collocats font intervenir différentes thématiques concernant ce concept. Une première thématique forte qui est présente dans les deux corpus IntUne est la dimension symbolique du mot « nation ». Cette dimension se retrouve dans les collocats colorés en jaune, l'entité semble alors très propice à la métaphore, elle est presque humaine, se comportant parfois en mère et étant menacée d'une mort rapide. Elle est aussi associée à des idées positives et fortes : elle intègre, on l'exalte. Ce côté hautement symbolique de la nation se retrouve en 2009. Elle y est alors personnifiée de façon plus paternelle, notamment avec le mot « père » qui peut représenter soit Gandhi ou encore Abraham Lincoln pour la nation américaine.

Irak : l'agonie d'une nation , confirmant notamment

li190207

L'arrière petit-fils du père de la nation Indienne, Tushar Gandhi

li050309

Un autre contexte qui apparaît dans ces collocations est celui de la religion, la nation est quelque chose qui se construit par ou contre cette religion. Les collocats s'y rapportant sont colorés en rose dans le tableau. Dans le corpus, c'est surtout la religion musulmane qui est présente, et elle est plutôt vue comme menaçant la nation. Il est à noter qu'aucune occurrence ne rapproche la religion chrétienne et le mot nation dans un contexte problématique dans le corpus IntUne 2007. Ce contexte religieux semble moins saillant en 2009.

dénoncé une agression contre les lieux saints de la nation musulmane.

fi260207

Le contexte politique est présent de diverses façons, les collocats qui s'y rapportent sont colorés en bleu foncé. Il peut s'agir de la présence importante du concept de nation dans les discours d'hommes politiques avec des mots comme « discours », « monopole », mais aussi du fait que la nation est une réalité politique avec une existence administrative

5- Deux cas particuliers : l'Italie et la Pologne.

reconnue avec le mot « république ». Dans ce dernier cas, la nation est alors associée à la République comme représentant toutes deux des principes essentiels et fondamentaux de la France. En 2007, cette étendue politique est très souvent représentée par la personne de son président. En 2009, le contexte politique apparaît également via des collocs comme « indépendante », « ordre », « république ». La nation est alors avant tout quelque chose qui bénéficie d'une reconnaissance par d'autres ensembles politiques, et qui en tant que tel, peut agir (honorer quelqu'un en lui donnant l'ordre de la nation, par exemple). Ce contexte d'une demande de reconnaissance par d'autres ensembles (par exemple l'État dont dépend telle ou telle nation) est également présent en 2007, mais ne se manifeste pas de façon aussi importante dans les collocations.

président de la République, c'est l'homme de la nation , ce n'est pas l'homme d'un clan, d'un parti	li170407
membre des comptes économiques de la nation dresse un bilan financier terrifiant de la décentr	ou090207
grands principes politiques – notamment laïcité, nation et république – gardent un profond enracinemen	li030507
regrettant qu'il ne qualifie pas l'Andalousie de nation dans son préambule, comme la Catalogne.	ec160207
La cinquième nation du monde, la France, prenne toute sa place en Eur	ou050507
donner par étape à la Corse un statut de nation indépendante en Europe. Officiellement formé	fi020209
soldats le soutient et la reconnaissance de la nation . A l'issue de cette visite, je voudrais	fi220409

La nation est aussi un concept très fortement associé à travers ses collocations à l'idée d'unité (en jaune pâle dans le tableau). L'unité rejoint dans les deux corpus IntUne la thématique politique puisque dans la vaste majorité des cas, quand cette dimension est évoquée, c'est par des hommes politiques. Les différents articles définis et indéfinis « la », « une », renforcent cette idée puisqu'ils sont actualisés au singulier, de même que des mots comme « unité », « unie » ou « entière » qui montrent que cette entité est fédératrice et porteuse d'une forte unité. Cette volonté d'unité est néanmoins assez régulièrement évoquée comme non réelle et simplement souhaitée. On la retrouve en 2009, notamment dans un contexte où cette nation est unie dans le discours du président la représentant.

réaliser le rêve d'une nation unie, chacun doit pouvoir participer au destin	li120307
Ami, entends-tu ce soir l'hommage que t'adresse la nation tout entière que tu as si noblement servi.	fi210409

Cette idée d'une unité de la nation rejoint le fait que c'est une communauté humaine qui a des origines culturelles et historiques supposées communes, une « ethnie », une « patrie », les collocs représentant la dimension des origines culturelles communes sont colorés en violet dans le tableau. Dans ces cas, les deux substantifs sont associés à travers des énumérations au mot « nation ». Ces différents concepts sont ainsi présentés comme

5- Deux cas particuliers : l'Italie et la Pologne.

proches ou équivalents. En 2007, le contexte culturel montre essentiellement la nation française et note sa grande postérité, notamment à travers le musée du Louvre dont une antenne pourrait ouvrir dans un pays du golfe persique. Le contexte historique est présent à la même période pour expliquer l'importance du concept de « nation » dans la campagne présidentielle ou encore pour expliquer d'anciens conflits qui ont pu avoir lieu.

posée, des personnes décédées à une ethnie, une nation , une race ou une religion déterminée. su230407
qui vient rappeler que le Louvre est au service de la nation tout entière », la lettre poursuit : « ce projet fi050207
concouraient, chacun à sa façon, à maintenir la nation en ordre de marche. li060307

Le corpus 2009 fait apparaître des thématiques qui n'étaient pas présentes dans les collocations de 2007. Il s'agit essentiellement de contextes sportifs, économiques et territoriaux. Le contexte sportif apparaît déjà dans le corpus 2007, mais sans que cela soit mis en relief par les collocations. Il donne plutôt une image positive des nations concernées, on parle alors d'une grande nation dans un sport quelconque. D'autre part, le Tournoi des Six Nations, manifestation sportive de rugby est présent dans de nombreuses occurrences. Enfin, lorsque « nation » apparaît dans un contexte économique, l'image qui en est donné est essentiellement défavorable : il s'agit alors de la France, et d'une France qui est en difficulté. En 2009, le collocat « football » fait intervenir une nation s'apparentant à une entité sportive, essentiellement française et dont les occurrences sont plutôt favorables :

avait décroché 5 médailles à Bordeaux en 2006 (2e nation derrière la Grande-Bretagne), quels sont les ou290307
La France, grande nation du football, avec ses 2,5 millions de licenciés li060309

L'apparition du contexte économique en 2009 se manifeste via différents collocats colorés en bleu turquoise dans le tableau. Le collocat « économique » apparaît uniquement dans le cadre de prévisions qui sont faites par un organisme, la « Commission économique de la nation ». Cet organe est le plus souvent jugé obsolète et comme n'étant pas adapté à la gravité de la crise économique qui sévit alors. D'autre part, le collocat « sociale » fait également intervenir la crise économique en ce qu'elle met en péril l'organisation sociale de la nation. Cette dimension sociale de la nation apparaît donc comme soumise à sa dimension économique. Enfin, le collocat « Live » se rapporte au nom d'une société musicale, « live nation », particulièrement active lors de la collecte de ce corpus.

Dans le cadre de la commission économique de la nation , elles sont obsolètes. ec110309
ives et les dangers. Pour la cohésion sociale de la nation , il est peut être encore temps de changer ou140209

Le contexte territorial apparaît également dans les collocations en 2009, montrant

5- Deux cas particuliers : l'Italie et la Pologne.

que lorsqu'on parle d'une nation, il peut aussi s'agir d'une étendue de terre pourvue de frontières. Les occurrences du corpus relatent essentiellement des faits d'agressions dans ce domaine. Ce sont le plus souvent les mots du président iranien qui sont repris et qui notent les problématiques territoriales entre pays occidentaux et pays arabes.

militaire pour priver de terres une nation entière sous le prétexte de la souffrance juive » ec210409
Israël, qui a privé « de terres une nation entière sous le prétexte de la souffrance juive su210409

V.1.3- Conclusion.

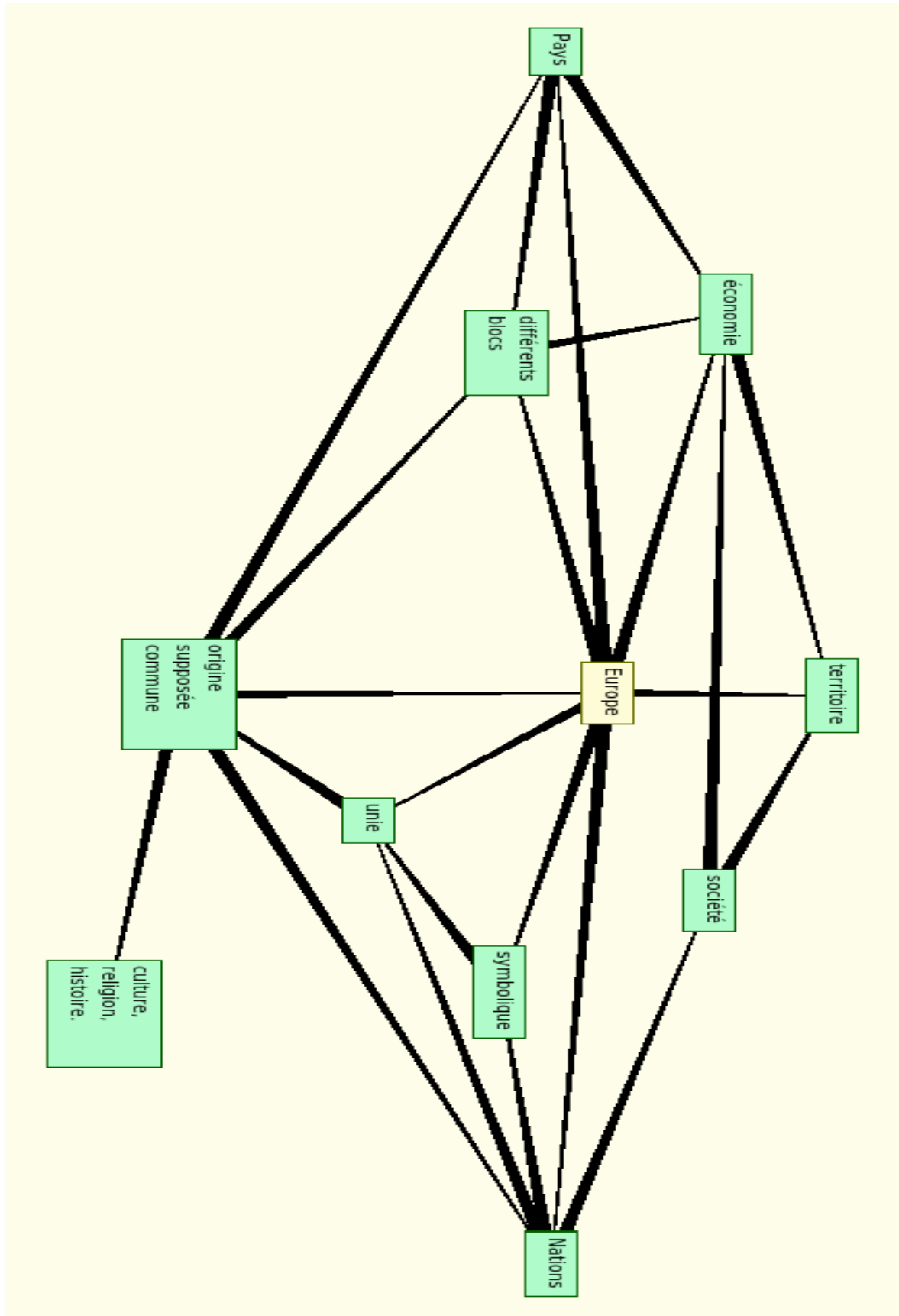
L'élaboration des prototypes sémantiques de pays et nation met en lumière un certain nombre de points par rapport à leur activation dans la presse française.

Tout d'abord, les différences entre ces deux notions sont mises en relief lors de l'analyse de leur présence dans le corpus. La plus importante d'entre elles réside dans le fait que si les pays peuvent être regroupés selon différents critères dans des ensembles à un échelon supérieur, les nations restent dans l'individualité et ne sont associées à d'autres nations qu'à partir du moment où une certaine forme de concurrence (territoriale, religieuse, sportive) se met en place. Le pays semble également être un concept plus économique que la nation, celle-ci se prêtant davantage à toute forme de symbolisation, notamment à travers une métaphore parentale.

Cependant, quand bien même ces deux entités montrent des différences qui justifient l'emploi de deux mots distincts, elles sont très proches et investissent des contextes qui se retrouvent de l'une à l'autre : contexte politique, économique, géographique, social, etc. Cette proximité peut être représentée à l'aide des réseaux collocationnels des deux mots dans la figure 21.

Dans cette figure, on voit notamment que les nations et les pays ont des relations complexes et étroites avec des ensembles créés à des échelons supérieurs : ils sont à la fois influencés par, et constituants de ces unités. La deuxième partie de ce chapitre va maintenant viser à montrer quelle représentation la presse française donne de l'Italie et de la Pologne en 2007 et 2009. Ces représentations présentent un angle de vue intéressant sur la façon dont l'image de l'Europe se construit peu à peu, pays par pays, dans les médias.

5- Deux cas particuliers : l'Italie et la Pologne.



(figure 21, relation entre les concepts d'Europe, nation et pays.)

V.2- La représentation de l'Italie dans la presse française en 2007 et 2009.

L'Italie est un pays de l'Union Européenne qui a des frontières communes avec la France. Berceau de la civilisation latine, c'est aussi un pays qui entretient de forts liens culturels avec la France. C'est enfin un des partenaires privilégiés du projet IntUne puisque ce projet est dirigé par l'université de Sienna. Dès lors, en étudier la représentation dans la presse française est intéressant et pourra permettre par la suite de réaliser des études comparatives au sein du groupe IntUne.

V.2.1- Vue globale et élaboration d'un prototype.

“Italie”, dans le corpus IntUne 2007 a une image plus complexe que celle accordée au concept « pays ». Si elle partage avec celui-ci les traits d'étendue géographique, d'un espace économique, ou d'une étendue culturelle, politique et historique. Elle peut aussi être un pays sportif ou un lieu avec une société dans laquelle des hommes vivent, ce qui la rapproche davantage du mot « nation ».

La répartition de ce mot dans le corpus IntUne lors des deux années de collecte est la suivante :

<i>Italie dans IntUne 2007</i>					
Source	Occurrences	Total de mots	%	Textes avec nation	Total de textes
Le Figaro	616	4861988	0.013	76	77
Libération	294	3602857	0.008	72	77
Les Échos	663	5015207	0.013	63	63
Sud-Ouest	270	3683377	0.007	71	77
Ouest France	221	3226744	0.007	67	77
TF1	15	251551	0.006	13	32
France 3	12	183741	0.007	9	34
Total	2091	20769219	0.010	371	436
<i>Italie dans IntUne 2009</i>					
Le Figaro	537	4473616	0.011	77	63

5- Deux cas particuliers : l'Italie et la Pologne.

Libération	232	4858636	0.008	68	77
Les Échos	410	2973339	0.009	63	75
Sud-Ouest	258	3786883	0.007	54	73
Ouest France	106	2146102	0.005	56	77
Total	1543	18238576	0.008	328	365

(Figure 22, répartition du mot « Italie » dans le corpus IntUne)

Ce mot est plus présent dans le corpus que nation, mais il est beaucoup moins représenté que le terme générique « pays ». Les deux journaux le mentionnant le plus sont les deux quotidiens nationaux avec une ligne éditoriale politique, la majorité des occurrences est dans Le Figaro en 2009.

Pour avoir une idée plus précise du sens que peut prendre ce mot, l'étude de sa « compagnie », des collocations qu'il a dans le corpus est intéressante. Xaira, avec la méthode précédemment exposée fournit les collocations suivantes du mot Italie en 2007 et en 2009.

2007			2009		
Collocat	Fréquence	z-score	Collocat	Fréquence	z-score
espagne	131	120,1	espagne	117	125,3
jozsef	12	74,3	grèce	28	67,5
scanavaca	3	70,4	allemagne	72	64,4
en	1035	68,3	en	680	51,1
allemagne	91	66,1	l'	795	48,6
emoi	3	54,5	portugal	17	43,2
l'	1019	52,7	pays-bas	19	40,2
20-juil	3	49,8	aquila	10	37,4
anterselva	3	46,1	france	120	31,4
prodi	25	44,9	royaume-uni	20	31,1
donati	3	40,6	jozsef	3	29,7
irlande	24	39,4	berlusconi	12	29
fasciste	7	39,3	belgique	17	28,4
unipublic	3	36,7	turin	7	27,8
belgique	25	35,4	grande-bretagne	19	27,6
pays-bas	21	33,2	euthanasie	4	27,6
portugal	15	28,6	castorama	3	26,5
grèce	12	28,5	malte	6	25,6
berbizier	4	27,8	séisme	8	24,1

5- Deux cas particuliers : l'Italie et la Pologne.

royaume-uni	19	27,5	débarquer	4	22
romano	13	26,3	elwana	4	21,2
luxembourg	12	26	benelux	3	20,7
truqués	3	25,3	abruzzes	4	20,3
brésil	18	25	,	965	18,7
pologne	17	24,7	silvio	6	18,4
benelux	5	24,1	32e	4	18,3
grande-bretagne	21	23,8	canada	9	18,2
extradé	3	22,5	et	344	17,8
cittanova	5	22,1	suède	8	17,4
01-janv	8	21,9	irlande	10	17,3

(Figure 23, collocations du mot « Italie » dans le corpus IntUne).

Ces collocations font apparaître plusieurs contextes. Les noms propres « Jozef Scanavaca », « Donati » et « Berbizier » sont des noms de footballeurs italiens ou celui d'un rugbyman français, sélectionneur de l'équipe nationale d'Italie. Le mot « unipublic » qui est une organisation gérant diverses compétitions cyclistes italiennes et le collocat « 32e » (32^e journée du championnat de football) appartiennent également à ce contexte sportif, ils sont colorés en rouge dans le tableau. Le contexte sportif semble donc bien présent.

Le contexte politique est aussi présent via des mots comme « prodi », « romano » ou « fasciste », l'image de l'Italie est alors quelque peu défavorable puisqu'on lui reproche une certaine instabilité politique -Romano Prodi, alors premier ministre a été dévoyé par sa propre majorité- voire des problèmes persistants de fascisme. Les collocats intervenant dans ce contexte sont colorés en bleu foncé dans le tableau. En 2009, l'entité montre toujours cette dimension politique, notamment avec les collocats « Silvio » et « Berlusconi » qui désignent le dirigeant du gouvernement italien. L'image qui en transparait via la lecture des collocations semble néanmoins beaucoup plus neutre alors, puisqu'elle n'est pas reliée au fascisme.

C'est une étendue géographique qui apparaît à travers les collocats colorés en rose saumon dans le tableau. Notons « en » qui est une préposition indiquant un lieu et « anterselva » qui est une ville située au nord de cette étendue. Cette dimension se renforce considérablement en 2009 avec de nouveaux noms de ville comme « Turin », la persistance de la préposition « en », mais surtout, l'apparition de la ville de l' « Aquila » et de la région dans laquelle elle se situe, les « Abruzzes ». Ces derniers collocats sont associés à un

5- Deux cas particuliers : l'Italie et la Pologne.

phénomène physique, un « séisme » qui a eu lieu dans cette région lors de la période de collecte du corpus. La réalité géographique et physique de l'Italie est donc très présente dans le corpus 2009. Un collocat « débarquer » pourrait appartenir à la fois à ce contexte géographique, mais aussi à un contexte plus politique. Quand ce mot est associé à « Italie » dans le corpus, il s'agit en fait d'un problème d'immigration. Le fait que l'Italie soit une étendue géographique associée à un pouvoir politique fait qu'elle est dotée de frontières, et c'est l'existence de ces frontières qui suppose la possibilité de problèmes migratoires.

L'Italie peut aussi être mise sur le même plan que de nombreux autres pays faisant pour la plupart partie de l'Union Européenne, ce contexte est marqué par la couleur jaune dans le tableau (quand des pays ne font pas partie de l'Union Européenne, ils sont cette fois colorés en jaune pâle). C'est le cas des collocats « Espagne », « Allemagne », « Irlande », « Benelux », « Pays-Bas », « Belgique », « Luxembourg » et « Grèce ». Ces différents pays auxquels l'Italie est associée en 2007 sont tous situés en Europe occidentale, que ce soit au nord ou au sud. Les termes « extradé » et « Brésil » sont dus à un contexte particulier : celui de l'arrestation d'un ancien terroriste italien au Brésil. En 2009, elle reste associée à de nombreux pays, mais ceux-ci ne font plus partie exclusivement de l'Union Européenne. En effet, l'Italie est citée notamment à proximité du « Canada », et cela ne concerne pas un contexte particulier, ces deux pays peuvent se rapprocher du fait de décisions financières, de caractéristiques culturelles ou de décisions sanitaires communes. L'association à l'« Espagne » persiste et montre une relation forte avec ce pays puisqu'il s'agit du premier collocat dans les deux corpus. Les autres pays de l'Union Européenne auxquels l'Italie est associée via ses collocats sont « Grèce », « Allemagne », « Portugal », « Grande-Bretagne », « France », « Finlande » et « Royaume-Uni ».

Les collocats marqués par la couleur bleu turquoise dans le tableau montrent que l'Italie est aussi une entité qui peut être associée à un contexte économique. Le mot « émoi » apparaît à plusieurs reprises dans le journal les Échos pour parler d'une offre publique d'achat : « *Emoi en Italie après l'offre d'ATT sur Telecom Italia* » (ec030407), tandis qu'en 2009, ce sont les résultats de « castorama » qui sont fournis.

Enfin, les collocats colorés en vert, « euthanasie » et « eluana » interviennent dans une actualité particulière. Une jeune femme atteinte d'une maladie incurable s'est faite euthanasier selon sa volonté, ce qui a déclenché les foudres d'un grand nombre d'italiens se

5- Deux cas particuliers : l'Italie et la Pologne.

réclamant de la foi catholique, l'Italie est alors fortement associée dans la presse française à l'idée d'un pays très catholique.

Grâce à cette première lecture des collocations du mot « Italie », nous pouvons voir que ce pays est présent selon différents sens dans la presse française. D'autre part, le fait que l'Italie soit un pays peut également nous renseigner sur les sens principaux qu'elle peut prendre dans la presse française. En conséquence, le prototype sémantique suivant semble pouvoir être proposé pour ce concept :

L'Italie est :

- Une étendue géographique.
- Une étendue culturelle.
- Un espace historique.
- Une étendue économique.
- Un espace politique.
- Un lieu où des personnes vivent.
- Un pays sportif.

Une analyse plus poussée des lignes de concordance dans lesquelles le mot « Italie » apparaît va alors permettre de voir de façon détaillée dans quels sens, et selon quelle actualisation de ces sens, ce pays est présenté dans le corpus IntUne. D'autre part, lors de l'étude de chaque sens potentiel, nous pourrons voir si l'Italie a une représentation stable dans la presse française de 2009 par rapport à 2007, si les choses se sont accentuées ou si au contraire, elles se sont affaiblies.

V.2.2- Actualisation des sens potentiels du prototype dans la presse.

V.2.2.1- L'Italie comme étendue géographique.

V.2.2.1.1- En 2007.

Tout d'abord, l'Italie peut être considérée comme un territoire physique, une étendue géographique. Cela peut être un territoire clairement séparé en deux parties : le nord et le sud. Le nord est dans ce cas décrit comme la partie la plus développée économiquement, cette région est désignée par des mots comme « poumon économique », tandis que le sud est décrit dans les termes d'un retard économique. Ces deux parties sont considérées comme ayant un grand nombre de « disparités »

5- Deux cas particuliers : l'Italie et la Pologne.

Le développement du sud de l'Italie , qui accuse un retard toujours plus prononcé sur
inculpés sont originaires du même village du sud de l'Italie . fi290307
tandis que dans le nord, poumon économique de l'Italie . ec230307
les mêmes disparités qu'entre le sud et le nord de l'Italie . li120307

L'Italie peut aussi être vue comme une étendue purement physique, dotée d'un climat particulier et d'une géographie. Dans la presse française de 2007, l'emphase est portée sur le climat chaud et sec de ce pays, décrivant avec force de superlatifs « hiver le plus chaud », « l'été le plus doux », une sécheresse qui cause une certaine anxiété aux italiens. Ces faits climatiques concernent dans la grande majorité des occurrences le nord de l'Italie. Une autre particularité en est décrite en rapportant un volcanisme très actif. Cette particularité apparaît dans le sud du pays.

L'Italie a déclaré l'état d'urgence (...) par crainte d'une sécher fi050507
L'Italie a connu cette année son hiver le plus chaud depuis fi050507
nombreux endroits, le long fleuve du nord de l'Italie est presque à sec. li280407
La petite île éolienne du sud de l'Italie a été mise en état d'alerte après un regain d'activité de ou010307
son volcan.
700 volcans (...), de l'Italie à l'Azerbaïdjan. li170307

V.2.2.1.2- En 2009.

Dans le corpus IntUne 2009, l'Italie est toujours présentée comme étant une étendue géographique. La dimension physique du pays reste présente via le séisme de l'Aquila qui montre que, quand il ne rencontre pas des problèmes de volcanisme en 2007, ce pays subit de violents tremblements de terre. Cette représentation d'une Italie aux prises avec des phénomènes physiques problématiques est donc plutôt cohérente entre 2007 et 2009. D'autre part, les distinctions entre une Italie du Sud et une Italie du Nord ont disparu en 2009. L'Italie, comme étendue géographique organisée, est alors présente lorsqu'il est question de son découpage administratif. Elle est comparée à la France, et le fait que le découpage y soit semblable avec le fonctionnement français vient cautionner ce dernier.

La colligation « d'Italie » fait apparaître clairement cette idée d'une étendue géographique en étant fréquemment précédée de verbes de mouvement comme : « venir », « amener », « ramener », etc. Quand le mot « Italie » est introduit par le déterminant défini élidé « l' », ce contexte géographique se manifeste à travers la réalisation de différentes infrastructures de transports (qu'il s'agisse de transports de camions, de trains, ou de matières

5- Deux cas particuliers : l'Italie et la Pologne.

premières) entre l'Italie et des pays frontaliers (France, Algérie). Il peut également être question dans ce cas d'une Italie qui est mal située au niveau environnemental et qui est alors associée à des pays généralement associés à l'idée de pauvreté ou de mauvaise gouvernance. L'introduction de « Italie » par un déterminant défini semble ainsi lui être assez peu favorable dans ce sens d'une étendue géographique.

séisme qui a touché les Abruzzes, la région centrale d'	Italie	donnant sur l'Adriatique.	ec070409
séisme qui a frappé la région des Abruzzes en	Italie	le 6 avril et causé la mort de	su230409
	L'	Italie a pris du retard dans la réalisation	fi240209
environnement, liste noire comprenant la Libye, Panama, (...) et l'	Italie		fi260309
territoires équivalents aux départements, j'en connais en	Italie	, en Espagne et en Allemagne.	ou070309
on ait perdu la trace du camion qui devait les amener d'	Italie	.	su050309

L'Italie dans la presse française est donc une étendue géographique. Ce sens potentiel est bien représenté dans les deux corpus IntUne, néanmoins, certaines persistances et certaines évolutions peuvent être notées. D'une part, ce pays est toujours représenté comme étant sous la menace d'une nature instable : volcans et séismes s'y produisent. D'autre part, il semble toujours poser problème du fait de son environnement jugé trop austral : association à certains pays connotés défavorablement en France (la Libye ou Panama, notamment), problèmes qui se déroulent précisément toujours dans le sud du pays (pauvreté structurelle, emprise de la mafia sur la gestion des déchets, etc.). Nous allons maintenant voir comment cette entité est présente dans la presse française lorsque c'est le sens potentiel d'étendue culturelle qui est activé.

V.2.2.2- L'Italie comme une étendue culturelle.

V.2.2.2.1- En 2007.

La presse française donne une image de l'Italie comme étant une étendue culturelle, et plus précisément, une étendue artistique. Cette dimension apparaît plus particulièrement avec la colligation « en Italie » et quand le pays est introduit par un déterminant défini. La colligation « en Italie » est utilisée dans ce contexte pour désigner une étendue générale où les artistes vivent ou dans laquelle ils viennent améliorer leur talent. C'est le cas dans ces occurrences les plus typiques :

est considéré comme un des artistes capitaux de la période baroque en	Italie	.	ec200407
---	--------	---	----------

5- Deux cas particuliers : l'Italie et la Pologne.

artistes qui préparaient le concours de Rome et voyageaient en Italie . fi080307

Quand « Italie » est introduit par un déterminant défini dans ce contexte artistique, la France y est très souvent associée. Ces deux pays sont alors considérés par la presse française comme deux lieux artistiques majeurs, d'une importance égale : quand la France possède le fameux « Penseur de Rodin », l'Italie dispose de certaines peintures du Titien, et ces deux endroits influencent les arts de la meilleure façon, c'est notamment le cas avec la musique.

France a envoyé le célèbre Penseur de Rodin, l' Italie a prêté un tableau du Titien tf1230307
musique d'une simplicité raffinée, influencée par la France et l' Italie . ec120407

V.2.2.2- En 2009.

La presse française présente toujours le mot « Italie » comme ayant le sens d'une étendue culturelle en 2009. Deux « Italie » sont présentées selon la préposition ou le déterminant qui les introduit. « En Italie » nous donne en effet à voir un espace où la musique a une grande importance et qui a fortement influencé sa pratique dans tout l'occident, la même influence se fait sentir dans le domaine de la peinture. C'est également un haut lieu de la haute couture et de la mode. « En Italie » fait aussi apparaître un espace où les romanciers sont inspirés, voire un espace qui inspire les romanciers dans leur écrits.

« L'Italie » semble cette fois davantage insister sur les domaines de la gastronomie, où ce pays est présenté comme ayant une tradition riche, et celui du cinéma où la richesse de tout un pays est également rapportée. « L'Italie » donne enfin à voir une zone où le patrimoine religieux est particulièrement développé et intéressant. Ces deux façons d'introduire le mot « Italie » semblent donc avoir chacune leur spécialisation dans la presse française de 2009 : l'article défini donne à voir une « Italie » intemporelle, avec ses grands traits caractéristiques : la gastronomie, la religion, le cinéma de l'époque de Cinecitta, tandis que la préposition indiquant un lieu montre une Italie plus actuelle : étendue où la haute couture s'épanouit, ainsi que lieu de prédilection des romanciers modernes.

Amalfitaine, Jamie Olivier savoure l' Italie et la restitue à merveille. Les recettes su130309
importance du patrimoine religieux, (...) ne connaissait pas l' Italie du sud. fi250309
Défilés. Tout le week-end, les shows se sont succédé en Italie . Alors que des griffes comme Ji. fi020309
plume (...) passe l'été dans une villa de famille en Italie . li130209

5- Deux cas particuliers : l'Italie et la Pologne.

En règle générale, lorsqu'il est question de culture, la presse française tend à présenter l'Italie comme un pays intemporel qui a eu un passé important et qui continue actuellement à influencer culturellement le monde entier. Cette influence se fait plus particulièrement en association avec la France dans le corpus 2007.

V.2.2.3- L'Italie comme un espace historique.

V.2.2.3.1- En 2007

L'Italie peut aussi être vue comme un lieu historique. Quand ce contexte est présent, l'Italie semble être la victime d'une histoire violente et plutôt défavorable dans la presse française. Un type particulier de collocation dans ce contexte est quand le mot Italie apparaît juste devant un adjectif. Sur 15 occurrences, 8 concernent en effet une dimension historique, et 6 des adjectifs relevés en antéposition à Italie sont « fasciste ». Ce collocat est le plus important dans le journal Libération. L'Italie, quand elle est caractérisée par un adjectif, peut ainsi être fasciste, mais aussi « ravagée par le terrorisme des brigades rouges ».

L' Italie fasciste fut bel et bien antisémite. li220207

L' Italie fasciste ne disposa jamais de savants. li220207

Quand la colligation « à l'Italie » est trouvée dans un contexte historique dans le corpus, cela traite essentiellement des traités et de certaines modifications des frontières de ce pays au cours du vingtième siècle.

En 1964, Trieste a été rattachée à l' Italie . Il faudra attendre le traité fi150207

Les traités de paix de 1919 pour que cette région revienne à l' Italie ec220307

Ainsi, l'histoire italienne semble-t-elle être quelque peu agitée du point de vue de la presse française en 2007.

V.2.2.3.2- En 2009.

L'Italie est aussi présente dans le corpus IntUne 2009 sous la forme d'un espace historique. Dans ce cas, différentes époques de l'histoire de ce pays peuvent être rapportées. Il peut s'agir de l'antiquité et les débuts du catholicisme via une martyr. Il peut aussi être question de la renaissance italienne à travers une famille très importante de cette époque, les Borgia, même si davantage que l'apport patrimonial de cette famille, ce qui y est relevé en est la décadence. L'histoire italienne se concentre surtout dans la presse française sur la

5- Deux cas particuliers : l'Italie et la Pologne.

période des années 1930 et 1940 avec le pouvoir mussolinien, la seconde guerre mondiale, et la naissance d'un mouvement à la fois politique ou artistique : le futurisme. L'histoire européenne communiste est commémorée par ce pays. Enfin, c'est un espace qui a connu dans les années 1980 et 1990 de nombreux problèmes de terrorisme.

Lituanie, la République Tchèque et l' Italie ont adopté une journée d'hommage aux victimes du communis fi030209
Borgia. Cette famille a régné sur l' Italie pendant la Renaissance (...). Accusée de fratricides, d'incestes ou030309
terrorisme fut avant tout politique en Italie . li210209

L'image historique de ce pays est ainsi plutôt défavorable dans la presse française. Ceci se retrouve à la fois en 2007 : pays fasciste, antisémite, etc. et en 2009 où alliance avec l'Allemagne d'Hitler, familles régnantes corrompues et décadentes, association avec des pays ayant eu un lourd passé communiste, terrorisme, etc. brossent un tableau plutôt sombre. La représentation de l'Italie comme étendue économique dans les médias français va maintenant être précisée.

V.2.2.4- « Italie » comme étendue économique.

V.2.2.4.1- En 2007.

Dans le corpus, l'Italie est souvent une étendue économique. Dans ce cas, la presse française donne principalement des informations objectives à propos de la situation économique de l'Italie. Quand « Italie » est utilisé comme un complément verbal, on trouve le plus souvent la colligation « En Italie » et les exemples montrent des compagnies qui font des affaires dans le pays. Cette image est plutôt positive pour l'Italie à l'exception des occurrences qui traitent de la compagnie aérienne Alitalia.

Fiat a décidé d'anticiper le lancement de la nouvelle 500 en Italie . fi160307
Axa en est convaincu, il a trouvé le « meilleur partenaire » possible en Italie . ec260307

Mais dans la presse française, l'Italie peut aussi être considérée d'une façon quasi métaphorique comme une entreprise qui connaît un certain développement, une expansion, ou qui peut exporter certains produits. Dans ces occurrences où l'Italie est considérée comme une compagnie, la majorité montre une situation bonne et dynamique pour le pays :

de plus en plus palpable, le réveil de l' Italie ec110407
l' Italie a accru la valeur de ses exportations hors Union ec190207

L'Italie, en tant qu'espace économique, peut être décrite à l'aide de propositions

5- Deux cas particuliers : l'Italie et la Pologne.

relatives, en utilisant des verbes d'action ou des verbes d'état. Dans ce cas, l'Italie est également considérée comme une compagnie. Cette image est une fois de plus positive dans les majorité des occurrences du corpus.

Une opération d'envergure en Italie qui devrait lui permettre d'être demain une	fi200207
l' Italie a enregistré une croissance de 2% l'an dernier, rattrapant une partie	ec280207
l' Italie est le deuxième organisateur de salons en Europe	fi250407

Une colligation plus intéressante donnant une image économique de l'Italie est « et l'Italie ». Avec cette colligation, l'Italie est associée à d'autres pays et il peut être intéressant dès lors de voir quels pays impliquent quels contextes. Ainsi, quand l'Italie est associée à la France dans une dimension économique, cette association est dans la majorité du corpus plutôt mauvaise. Ces deux pays ne sont pas compétitifs en comparaison d'autres états membres de l'Union Européenne comme l'Allemagne.

La France et l' Italie ont perdu des parts de marché au profit de l'Allemagne	ec160407
C'est le risque, la France et l' Italie notent déjà que l'Allemagne	li180407

Quand l'Italie est associée à l'Allemagne, la situation est à l'exact opposé : l'Italie est dans ce cas une entité compétitive et, dans certaines occurrences, plus compétitive que la France. Enfin, quand l'Italie est associée à l'Espagne, les deux pays sont décrits comme étant des espaces dynamiques, avec une bonne croissance économique.

La Grande Bretagne, l'Allemagne et l' Italie constituent ses principales bases de clientèle.	fi200307
France se classe désormais devant (..) l'Allemagne et l' Italie en termes de pression fiscale.	fi190207
La Grande Bretagne, l'Espagne et l' Italie sont beaucoup plus attractifs que nous	li040407
Dont le regard se porte sur l'Espagne et l' Italie , marchés en forte croissance.	ec160207

Ainsi, la presse française donne-t-elle une image positive de l'Italie en tant qu'entité économique, si ce n'est quand celle-ci est associée à la France.

V.2.2.4.2- En 2009.

L'Italie est également présentée dans le corpus IntUne 2009 comme une zone économique. Les contextes sont alors beaucoup plus tranchés qu'en 2007 où le pays était présenté de façon plutôt neutre et factuelle dans l'ensemble. Ici, les occurrences montrent dans une vaste majorité un pays qui rencontre de nombreuses difficultés. Quand il est question du pays, des comptes de la nation et de son économie structurelle, les constats sont assez peu favorables : entre déficits gigantesques, hausse du chômage et très mauvaises

5- Deux cas particuliers : l'Italie et la Pologne.

notations pour emprunter sur les marchés internationaux. Cette représentation structurelle défavorable se trouve le plus souvent associée à une colligation formée de « Italie » introduite par un déterminant défini.

La Grèce, l'Espagne, l' Italie et le Portugal sont déjà considérés comme fragiles	fi230209
L' Italie continue à perdre des emplois	fi110309
La dette allemande est montée (...) pire pour (...) l' Italie	li040309

Lorsque les PME et grandes entreprises d'Italie sont évoquées, la description est sensiblement meilleure, elles peuvent avoir de bons résultats ou exporter leurs produits phares (de la haute couture, notamment). Il est intéressant de noter que dans ce contexte, leurs résultats ne sont jamais aussi bons que lorsque ces entreprises ont des partenariats avec la France. Les deux domaines de production dans lequel ce pays semble se distinguer sont ainsi la production de vin et la haute couture, en partenariat avec la France. A l'inverse, dans les domaines de l'automobile ou du textile moyenne gamme, les difficultés s'amoncellent sur les comptes des entreprises.

deuxième secteur manufacturier en Italie , il faut le protéger.	fi240209
L' Italie possède la viticulture la plus compétitive en 20	su030309
Fiat n'exclut pas de possibles fermetures d'usines en Italie .	li280309

Ce pays défend enfin, au niveau économique, un plan de relance de son industrie automobile devant la Commission européenne. Dans ce cas, il est un soutien de la France qui défend le même type de plan et cette démarche apparaît comme légitime et efficace.

La France, accusée par plusieurs voisins à l'exception notable de l' Italie d'entrave à la libre concurrence

fi130209

Ainsi, l'image économique de l'Italie dans le corpus IntUne 2009 est sensiblement plus tranchée et défavorable que celle qui en était donnée en 2007. Cela peut s'expliquer par un contexte de crise internationale, alors qu'en 2007, les constats étaient plutôt euphoriques, à l'exception des occurrences où l'Italie était associée à la France. D'autre part, il est étonnant de constater le revirement de situation lorsque dans ce contexte économique, l'Italie est associée à la France : en 2007, il s'agissait de mauvais résultats quand en 2009, ils deviennent meilleurs. Cela peut montrer que la presse, à la faveur de la crise économique, tend à montrer qu'il s'agit là de deux pays dont les économies sont à la fois proches et saines

5- Deux cas particuliers : l'Italie et la Pologne.

V.2.2.5- L'Italie en tant qu'espace politique.

V.2.2.5.1- En 2007

L'Italie peut aussi être une entité politique, avec des institutions définies. En 2007, ces institutions peuvent être financières comme la « banque d'Italie ». Si cette banque peut alors avoir des actions bénéfiques pour l'économie italienne dans le corpus, elle peut surtout avoir connu des problèmes lors de l'exercice d'un gouverneur passé.

Ancien gouverneur de la banque d' Italie a usé de tout son pouvoir de nuisance ec280307

En tant qu'espace politique, l'Italie est montrée comme ayant une politique intérieure, avec un gouvernement et des hommes politiques. Dans la presse française, ce contexte montre essentiellement une Italie qui doit trouver une solution au problème de l'immigration et, dans la majorité des cas, l'Italie dans une crise politique. Le gouvernement rencontre en effet des problèmes de stabilité du fait d'une très large coalition au pouvoir. Une explication à la représentation de ce contexte peut être que les deux journaux politiques du corpus sont de droite et de gauche et non pas du centre, en conséquence, ils montrent une Italie qui rencontre des difficultés du fait de la position centriste de son gouvernement.

Chaque jour, l' Italie doit faire face à l'arrivée d'immigrés clandestins fi250407

L' Italie plonge ainsi dans la crise politique la plus grave ec220207

Mais l'Italie a également une politique étrangère. Le contexte concerne plus particulièrement deux pays étrangers qui sont l'Afghanistan et le Brésil. L'explication de la sur-représentation de ces deux pays se trouve dans l'actualité : l'Italie souhaite l'extradition d'une personne au Brésil et un otage ressortissant italien a été libéré en Afghanistan. Les occurrences concernant ce dernier contexte sont dans la majorité négatives, notamment parce que l'Italie a accepté des négociations avec les terroristes. A l'opposé, les occurrences qui concernent le Brésil sont dans l'ensemble positives. Dans ce contexte, elle est décrite comme un état de loi, et la personne accusée (et appréhendée à l'aide des autorités françaises), Cesare Battisti, comme un terroriste.

L' Italie a affaibli le gouvernement Karzaï en le contraignant à fi220307

extradition est en vigueur depuis 1993 entre l' Italie et le Brésil su190307

Ainsi, l'image politique de l'Italie est-elle nuancée : c'est un pays qui a connu des problèmes au niveau de ses institutions financières par le passé, mais ces problèmes sont

5- Deux cas particuliers : l'Italie et la Pologne.

résolus. Il a certains problèmes intérieurs, notamment du fait de l'instabilité de son gouvernement, et ce pays a négocié avec des terroristes, mais il peut néanmoins juger des personnes dans le respect de la loi.

V.2.2.5.2- En 2009.

La presse française montre toujours ce pays comme un espace politique en 2009. C'est une étendue qui dispose ainsi de frontières et d'un gouvernement avec un chef : Silvio Berlusconi, son nom est cité à maintes reprises et il figure dans les vingt premiers colloques de « Italie ». La dimension nationale de l'Italie, à travers la plus forte identification via la personnification de l'autorité souveraine est ainsi plus présente en 2009. Cette étendue politique peut agir dans différents domaines. Elle mène tout d'abord une politique intérieure, et dans ce contexte, c'est essentiellement le problème de l'immigration qui est traité : l'Italie, ainsi que d'autres pays situés au sud de l'Europe sont largement confrontés à des « vagues » de migrants, et différentes mesures sont mises en place pour tenter d'endiguer ces flux de population. Toujours au niveau intérieur, la forte immigration conduit également à une résurgence du racisme dans ce pays, ce qui est noté dans la communauté internationale, notamment du fait de lois discriminatoires à l'égard de la communauté tzigane.

afflux continu de clandestins sur leurs côtes (...) l' Italie , la Grèce fi270209
L' Italie , seul pays occidental dans ce cas, est en train de li130209
se précipiter vers l'abîme des lois raciales.

La politique intérieure de l'Italie est aussi étudiée par différents personnages, et ce, essentiellement à travers le personnage de Silvio Berlusconi, sa popularité et les raisons de son succès sont analysées et expliquées à travers l'absence d'opposition stable. En fonction de la ligne éditoriale des journaux français se penchant sur cette question, les descriptions qui en sont données sont neutres ou péjoratives.

Il ne veut pas gouverner l' Italie . Il veut en devenir le patron fi230209
auteur de l' Italie Sur le fil du rasoir. A l'intérieur du PDL, on retrouve des (...) post-fascistes li270309

Le gouvernement, lorsqu'il est personnifié par son chef, Silvio Berlusconi, mène aussi une politique économique. Cette politique économique prend deux visages. Si elle touche au domaine énergétique et à un partenariat avec la France sur le nucléaire, elle est décrite comme étant très positive, étant en rupture avec le « fanatisme » des écologistes. A l'opposé, si elle est considérée par les chefs d'entreprise italiens, alors ceux-ci regrettent une

5- Deux cas particuliers : l'Italie et la Pologne.

politique trop inactive pour remédier à la crise économique internationale.

a fustigé le fanatisme écologique qui a privé l'Italie d'énergie nucléaire pendant près d'un quart de siècle. L'Italie fait un pas significatif vers une nouvelle stratégie énergétique. La grogne des patrons monte en Italie contre le chef du gouvernement, Silvio Berlusconi

fi250209
fi240209
ec170309

Cette étendue politique s'insère pleinement d'autre part dans un espace politique plus large qui est l'Union Européenne. Elle fait alors doublement partie de cette entité : elle en partage la monnaie, ce qui lui permet d'éviter de douloureuses dévaluations, contrairement au Royaume-Uni, par exemple. De plus, elle partage avec cette entité des visions communes au niveau des relations internationales. Le rapport de l'Italie à l'Union Européenne est cependant loin d'être idyllique, surtout lorsqu'il s'agit de « Bruxelles », la Commission européenne reprochant à ce pays divers manquements au niveau de la salubrité des eaux publiques de certaines de ses villes ou s'opposant au protectionnisme de la politique économique qui y est menée.

l'euro. Sans celui-ci, l'Italie, l'Espagne et la France auraient déjà subi 3 dévaluations. L'Union européenne a mis en demeure l'Italie, hier, de collecter et traiter les eaux. L'Italie a demandé aux entreprises (...), s'attirant les foudres de Bruxelles.

li260209
fi200209
fi280209

Au niveau politique, cette entité entretient également des rapports avec le reste du monde, elle a ainsi une politique de relations internationales plus ou moins facile, critiquée par la première dame française (d'origine italienne), et ayant parfois un certain décalage entre ses déclarations d'intentions et ses actes effectifs. C'est aussi un pays qui est montré par la presse française comme ayant une politique de défense. L'Italie est alors dans la vaste majorité des occurrences un pays qui participe pleinement à l'effort de guerre portant sur l'Afghanistan demandé par le président Obama. Cette vision sera plutôt positive ou négative selon la ligne éditoriale du journal qui la mentionne. C'est enfin un pays qui accueille un dirigeant à la renommée négative dans le monde occidental en la personne du colonel Kadhafi.

Kadhafi devrait effectuer une visite officielle en Italie. Première dame de France a accusé l'Italie de ne pas avoir versé les (...) promis par le pays à ce sujet. Pressions américaines, la France, l'Italie et les Pays-Bas (...) envoi d'une force de plusieurs centaines

fi040309
su130209
fi280309

En 2009, l'image politique de l'Italie est ainsi très fortement marquée par la personne de Silvio Berlusconi, on analyse son succès électoral, et ses décisions sont admirées ou critiquées dans le domaine de l'économie. D'autre part, les problèmes intérieurs

5- Deux cas particuliers : l'Italie et la Pologne.

qui étaient décriés en 2007 n'ont plus cours, ce pays est plutôt comparé à la France dans le domaine de sa politique internationale (surtout par Le Figaro) en ce qui concerne par exemple ses difficultés avec la Commission européenne, ou bien encore en recevant des personnalités décriées.

V.2.2.6- L'Italie comme un lieu où des personnes vivent.

V.2.2.6.1- En 2007.

Dans le corpus IntUne 2007, l'Italie est parfois représentée avec une dimension sociale. Dans ce contexte, les personnes qui y vivent peuvent avoir certaines particularités sociétales, ou cette société peut rencontrer des problèmes avec la loi et la légalité. Le comportement sociétal de l'Italie décrit, dans la majorité des occurrences, un pays qui est archaïque et très catholique, parfois même d'une façon caricaturale. Quand un adjectif est utilisé pour caractériser « Italie » dans ce contexte, cela peut être « catholique » ou d'une façon plus marquée, plus accentuée : « processionnaire » et « superstitieux ». Le pays est vu comme se positionnant contre l'avortement et comme utilisant d'« antiques » systèmes pour les bébés abandonnés.

En Italie , il est encore mal vu d'habiter ensemble sans être marié	fi200307
En Italie , la consultation d'un médecin est un préalable (...), refuser un IVG	fi100207
évoque la superstitieuse Italie processionnaire	su240307

Quand des enquêtes judiciaires ou bien la légalité au sein du pays sont concernées dans la presse française, l'Italie semble alors être dépeinte de façon particulièrement marquée comme un pays où l'illégalité est la règle. Beaucoup des occurrences traitent de toutes sortes de trafics et de crimes, et présentent le pays comme ayant un sens des plus relatifs de la légalité.

l' Italie est celui qui connaît le mieux la pratique des enlèvements criminels	ec280307
Trafic (...) à destination de l' Italie	ou160207
l' Italie est incontestablement le leader en Europe pour le trafic de contrefa	ou220307
l' Italie Est déjà connue à l'étranger pour avoir un sens relatif de la légalité	ec300307

V.2.2.6.2- En 2009.

L'Italie est toujours en 2009 une étendue sur laquelle des personnes vivent, où une société s'organise. Au niveau des caractéristiques communes fortes de cette société, la presse

5- Deux cas particuliers : l'Italie et la Pologne.

française accentue alors toujours particulièrement l'importance de la religion catholique dans ce pays d'un point de vue moral. Les représentants de cette religion prennent ainsi position dans divers sujets du domaine public dans ce pays, notamment dans le contexte de l'euthanasie d'une personne atteinte d'un mal incurable selon sa volonté, ou celui de l'usage de plus en plus important des nouveaux modes de communication virtuels.

Évêque de Modène, dans le nord de l'Italie a invité les jeunes de son diocèse à renoncer à la prière se sont organisées (...) un peu partout en Italie à l'initiative des mouvements catholiques. su040309
fi100209

La société active est aussi plus particulièrement considérée, dans ce cas, il peut être question du fait que cette population est une de celles qui veillent à rester les plus élégantes en Europe, mais encore du fait que les italiens font grève régulièrement. En termes de démographie, cette société est considérée comme étant vieillissante et particulièrement moins dynamique que la France.

L'Allemagne reste faible, ainsi que l'Italie et l'Espagne. Une démographie dynamique reste un facteur d'attractivité. fi020309
L'Italie, première à soigner sa mise en valeur. li060309

La société italienne montre dans la vaste majorité des occurrences une difficulté majeure en ce qui concerne la légalité. La presse française relate ainsi un système de justice à l'efficacité relative, ainsi que de nombreuses affaires de corruption ou de problèmes liés à la mafia. Elle montre en outre que cette société est contrainte de s'organiser de façon privée pour lutter contre les très nombreuses agressions qui s'y produisent.

Le contraste entre l'Italie touristique et les problèmes dénoncés par le rapport Ecomafia. (...) Le système à démonter est énorme, li190309
le fonctionnement de la justice, l'Italie est à la 70^e place mondiale, derrière l'Angola et le Gabon. li070209
Le sud de l'Italie. La mesure (...) faits de corruption et des appels d'offres truqués pour lutter contre (...) agressions sexuelles, l'Italie a institué hier le principe de rondes citoyennes (...) pour assurer ec170209
fi210209

L'image véhiculée de l'Italie à travers sa société dans la presse française en 2007 comme en 2009 reprend ainsi une bonne partie des clichés habituels concernant ce pays. Ses habitants suivent la mode de près, ont un sens de la légalité relatif, l'influence de la mafia y est très importante et les principes hérités du christianisme y sont très prégnants. On peut tout de même noter que les problèmes de légalité sont plus accentués dans le corpus IntUne 2009, notamment à travers les évocations de la mafia, ou encore du classement mondial des systèmes de justice. Ce sens potentiel d'un lieu où des personnes vivent est celui qui connaît la plus grande stabilité entre sa représentation en 2007 et celle en 2009.

5- Deux cas particuliers : l'Italie et la Pologne.

V.2.2.7- L'Italie comme un pays sportif.

V.2.2.7.1- En 2007.

Cette dimension de l'Italie dans la presse française peut être intéressante du fait d'une rivalité entre ces deux pays, particulièrement en 2007, après la finale de la coupe du monde de football que l'Italie a gagné contre la France. Dans la presse française, quand l'Italie est montrée comme un pays sportif, cela concerne essentiellement deux sports : le football et le rugby.

Dans un contexte de football, dans la vaste majorité des journaux du corpus, l'Italie a une image des plus douteuses. Nous pouvons ainsi trouver des mots comme « douteux », « truqués », « exclusion » et « scandale ». Nous pouvons aussi noter que l'équipe championne du monde n'est pas une si grande équipe que cela, surtout lorsqu'elle joue un match contre la France. L'autre sport italien, le rugby, a une représentation sensiblement différente dans la presse française. Le plus souvent, l'Italie y est décrite comme connaissant une grande progression et nous pouvons alors trouver des environnements comme « victoire à l'extérieur » ou « membre à part entière du Tournoi ».

Transferts douteux pour alimenter des caisses noires et des matchs truqués comme en Italie		ec090207
L'exclusion de l' Italie	a jeté un froid dans le calcio, renvoyant le football transalpin à ses scandales	li190407
l' Italie	a commis une grosse bêtise : elle a concédé le nul à Naples	li240307
L' Italie	s'est épuisée contre les bleus	li050207
Qui bénéficie depuis l'an passé à l' Italie	devenue membre à part entière du Tournoi.	fi140407

Cela peut s'expliquer par le fait qu'en France, le sport le plus important est le football et la récente coupe du monde en 2007 peut être un souvenir cuisant. Le rugby, en comparaison, ne bénéficie pas d'une telle importance et, en outre, dans la vaste majorité des matchs de rugby entre ces deux pays, c'est la France qui gagne.

Il n'y a qu'un journal du corpus qui donne une image différente. Dans ce journal, l'Italie est une bonne équipe de football, mais au rugby, c'est une mauvaise équipe nationale. Le journal qui fournit cette image de l'Italie est Sud Ouest. C'est un régional du sud de la France, et cela peut être une explication à ce point de vue différent puisque dans cette région, les lecteurs ne sont pas intéressés par le football, mais ils sont passionnés de rugby.

5- Deux cas particuliers : l'Italie et la Pologne.

L' Italie qui, Totti mis à part, va se reposer sur ses champions du monde su280307

Une Italie qui ne l'a jamais emporté dans le Tournoi et a enregistré trois cuisantes raclées su130407

Ainsi, l'image de l'Italie dans la presse française dès lors qu'il s'agit de sport semble quelque peu manquer d'objectivité. Le récent championnat du monde de football a laissé un souvenir difficile, et toutes les difficultés que pourrait rencontrer le football italien sont alors accentuées dans les journaux français, à l'exception du seul qui considère le rugby comme étant plus important que le football : Sud Ouest.

V.2.2.7.2- En 2009.

La majorité des occurrences du corpus 2009 qui mentionnent le mot « Italie » le font dans un contexte sportif. Trois sports sont alors essentiellement concernés, il peut s'agir du rugby, du football ou du tennis. Le rugby est le sport qui est le plus présent lorsqu'il est question d'Italie. Cela peut paraître relativement étonnant dans la mesure où ce pays s'illustre davantage au football. Néanmoins, la plupart des occurrences proviennent du journal Sud Ouest dont la diffusion favorise une importance plus large accordée au rugby. Dans ce contexte sportif, l'Italie est un protagoniste de deuxième importance dans le Tournoi des six nations : ce pays doit affronter la France, mais il est noté que l'équipe française n'a pas de gloire à gagner contre cette équipe, et qu'une victoire serait trop peu significative contre l'Italie pour lui permettre d'oublier un échec face à l'Angleterre.

Même une victoire en Italie ne pourra effacer l'amertume de la déroute de Twickenham. su160309

L' Italie est la dernière du Tournoi avec deux défaites assez lourdes dues li280209

Concernant le football, l'Italie est présente non pas de façon métonymique à travers son équipe nationale mais plutôt à travers sa fédération qui est une concurrente sérieuse de la France dans le cadre de l'organisation d'une manifestation sportive en 2016. Cependant son échec pour l'organisation de la même compétition en 2012 est fréquemment rappelé. Elle est aussi présente dans ce sport via ses clubs de football dont les bons résultats sont particulièrement soulignés dès lors qu'ils emploient des sportifs français.

Football. L' Italie envisage de se porter candidate à l'organisation de l'Euro 2016 après sa défaite surprise fi040309

Gourcuff (...) a été critiqué quand il est parti en Italie mais je pense que cela lui a beaucoup apporté. su060309

Enfin, les sportifs de ce pays, plus particulièrement lorsqu'il s'agit de cyclisme, semblent rencontrer de nombreux problèmes avec les instances anti-dopages. La presse

5- Deux cas particuliers : l'Italie et la Pologne.

française insiste d'autre part sur le fait que les dépistages semblent plus efficaces lorsqu'ils sont effectués sur le tour de France que lorsqu'ils sont faits en Italie, laissant à penser implicitement à un certain laxisme des autorités italiennes dans ce domaine.

Jamais été contrôlé positif
(notamment sur le Tour d' Italie où il avait été terminé 2è) mais a été pris facilement sur le Tour de France su040209
en Italie permettraient une identification par ADN et une comparaison par rapport li120209
aux poches de sang saisies.

L'image de ce pays semble être moins tranchée que ça n'était le cas en 2007 dans la presse française de 2009. Néanmoins, les bons résultats des clubs italiens ou de l'équipe nationale italienne dans le domaine du football sont très peu cités, et le portrait qui est brossé de ce pays reste globalement défavorable : sportifs se dopant, équipe nationale de rugby de peu de qualité, etc.

V.2.2.8- Conclusion.

L'Italie en 2009 présente une image relativement stable avec celle proposée en 2007. Nonobstant, certaines différences se font ressentir, notamment dans le domaine économique où ce pays n'est plus vu de façon aussi favorable, et où sa comparaison avec la France se fait au profit de cette dernière. Dans un contexte de politique intérieure, le chef du gouvernement semble également marquer davantage la représentation qui est donnée de ce pays par la presse française que ça n'était le cas en 2007. Enfin, dans le domaine de l'illégalité, les problèmes de l'Italie semblent s'être accrus à travers une recrudescence de la présence mafieuse dans les journaux français et un système judiciaire particulièrement mal situé au niveau mondial.

La presse française donne une image de l'Italie qui est donc relativement stéréotypée dans les corpus IntUne 2007 et 2009, ces stéréotypes étant d'autant plus prégnant dans des domaines comme la société, la culture ou bien le sport. Néanmoins, l'Italie est fréquemment associée à la France et il semble exister au sein des médias une proximité indéniable entre ces deux pays qui font partie de l'Union Européenne. Bien que subissant des représentations fortement empreintes de clichés, ces deux pays semblent donc pouvoir dans certains contextes être unis et intégrés au sein de cet ensemble. Nous allons maintenant voir dans quelle mesure la Pologne est présentée dans la presse française selon le même mode opératoire.

V.3- La représentation de la Pologne dans la presse française en 2007 et en 2009.

La Pologne est un vaste pays européen qui partage des frontières avec l'Allemagne à l'ouest et la Russie à l'est. Ce pays est souvent représenté dans le corpus IntUne comme étant prototypique de la zone « Europe de l'Est » : quand cette zone est suivie de différents pays la composant, la Pologne est le pays le plus fréquemment cité. D'autre part, outre sa position stratégique (entre Allemagne et Russie) en Europe, ce pays a connu une histoire des plus mouvementées. Ce fut notamment le cas tout au long du vingtième siècle. La plus récente étape de cette histoire en a été l'adhésion à l'Union Européenne en 2004, lors du plus vaste élargissement que cette union ait jamais connu. Il est ainsi intéressant de voir comment ce pays est représenté dans la presse française de 2007 et de 2009 : quelles évolutions ont éventuellement pu apparaître ? Présente-t-il des similitudes avec un pays présent depuis 1957 dans l'Union Européenne ? Est-il montré comme en faisant partie intégrante ? L'élaboration d'un prototype sémantique pour ce pays en 2007 et 2009 permettra de mieux appréhender ces différentes questions.

Afin de répondre à ces différentes questions, nous allons donc dans un premier temps élaborer un prototype sémantique en procédant selon le même mode opératoire que pour l'Italie. Une vue globale de la présence du mot « Pologne » dans le corpus sera étudiée, notamment à travers l'étude de ses collocations. Une fois cette première vue proposée, nous pourrons ainsi élaborer un prototype sémantique pour ce pays. Dans une seconde partie, nous verrons comment chaque sens potentiel de ce prototype est activé dans la presse française.

V.3.1- Vue globale et élaboration d'un prototype.

La Pologne est moins présente que l'Italie dans le corpus IntUne 2007. Cela pourrait assez simplement s'expliquer par la moins grande proximité culturelle et géographique entre ce pays et la France. Il apparaît à 741 reprises en 2007 et à 462 reprises en 2009. Le mot « Pologne » est ainsi sensiblement moins présent dans le corpus IntUne 2009. Cela peut notamment s'expliquer par le fait que l'actualité du bouclier antimissile américain, très forte en 2007, s'est quelque peu atténuée en 2009. Ce mot est réparti de la façon suivante dans le corpus.

5- Deux cas particuliers : l'Italie et la Pologne.

<i>Pologne dans IntUne 2007</i>					
Source	Occurrences	Total de mots	%	Textes avec Pologne	Total de textes
Le Figaro	231	4861988	0.005	67	77
Libération	85	3602857	0.002	38	77
Les Échos	231	5015207	0.005	57	63
Sud-Ouest	83	3683377	0.002	39	77
Ouest France	100	3226744	0.003	41	77
TF1	4	251551	0.002	4	32
France 3	7	183741	0.004	3	34
Total	741	20769219	0.004	249	436
<i>Pologne dans IntUne 2009</i>					
Le Figaro	162	4473616	0.003	54	63
Libération	56	4858636	0.002	38	77
Les Échos	156	2973339	0.003	57	75
Sud-Ouest	55	3786883	0.001	30	73
Ouest France	33	2146102	0.002	24	77
Total	462	18238576	0.003	203	365

(Figure 24, répartition de « Pologne » dans le corpus IntUne)

Les deux journaux mentionnant le plus ce mot sont Le Figaro et Les Échos. Cette répartition est relativement stable avec celle observée dans le corpus 2009. Notons néanmoins la plus faible représentation de ce pays parmi les régionaux de notre corpus en 2009. Nous pourrions nous demander si cette évolution de la répartition du mot entre 2007 et 2009 conduit à une évolution du sens qui lui est donné. Une analyse de ses collocations permet alors d'appréhender ses sens dans le corpus. Elles peuvent se présenter dans le tableau suivant.

2007			2009		
Collocat	Fréquence	z-score	Collocat	Fréquence	z-score
tchéquie	6	94	pruszkow	12	207,5
intercepteurs	7	85,9	cracovie	4	78,6
gorak	7	76,5	hongrie	15	70,3
hongrie	15	76	tchécoslovaquie	3	62,8

5- Deux cas particuliers : l'Italie et la Pologne.

tchèque	20	69,1	tchèque	19	58,1
maltraiter	3	68,3	ukraine	9	42,2
cracovie	5	61,3	antimissile	6	40,4
solidarnosc	3	48,3	délocaliser	4	39,8
slovaquie	6	46,6	en	268	38,8
bulgarie	8	44,8	slovaquie	4	30,3
jaruzelski	3	44,7	roumanie	7	30,1
en	377	42	allemagne	18	29,3
ukraine	11	41,6	lituanie	5	28,4
république	37	41	république	18	27,5
antimissile	6	38,3	eurosport	4	25,8
geremek	3	32,7	bulgarie	3	24,5
assistés	3	30,8	incendie	6	23,2
roumanie	8	30,3	hitler	3	21,8
avortement	6	28,3	la	234	20,1
la	395	27,9	orientale	3	20,1
antimissiles	3	27	belgique	6	18,4
baltes	3	25,3	tragique	3	17,3
malte	3	24,9	espagne	9	17,2
italie	17	24,7	royaume-uni	6	17
missiles	5	24,5	italie	8	16
installer	10	23,3	pologne	4	14,8
n°	6	22,6	et	128	13,7
kaczynski	3	20,1	russie	7	13,6
varsovie	4	20	bouclier	4	12,4
lituanie	4	19,2	autriche	3	11,2
staline	3	18,3	grande-bretagne	4	10,5

(Figure 25, collocations du mot « Pologne » dans le corpus IntUne)

Parmi ces trente premiers collocats, beaucoup représentent des pays européens. Ces collocats peuvent appartenir à l'Union Européenne. Ils ont alors été colorés en jaune (Europe de l'Est, c'est-à-dire les pays entrés dans l'Europe politique après le milieu des années 1990) et en vert (Europe de l'Ouest, c'est-à-dire les pays présents depuis plusieurs décennies dans l'Europe politique). La République Tchèque concentre à elle seule trois collocats en 2007 et semble donc avoir un fort lien dans la presse française avec la Pologne. Nous pouvons noter qu'en 2007, la Pologne est davantage associée à des pays situés en Europe de l'Est qu'en 2009, où, dans le bas du tableau, des pays d'Europe de l'Ouest apparaissent en plus grand nombre. Ce groupe de collocats contribue donc à donner de la Pologne le sens d'étendue

5- Deux cas particuliers : l'Italie et la Pologne.

politique.

Ce sens politique peut également s'associer à un sens plus géographique lorsque la Pologne s'associe à d'autres pays, ces collocsats sont colorés en marron dans le tableau. La presse française tend alors à montrer que ce pays est avant tout « oriental », et qu'il entretient de forts liens avec d'autres pays de l'Est n'appartenant pas à l'Union Européenne comme l'Ukraine ou la Russie. En 2007, l'Ukraine est présente dans les collocsats en ce que la Pologne semble en être une sorte d'aînée, aidant celle-ci à se développer économiquement, et en appuyant la candidature à l'Union Européenne. Ce collocat n'apparaît plus en 2007, il est alors remplacé par les collocsats « Russie » et « orientale », qui ne sont plus associés à l'idée d'Union Européenne.

La Pologne peut également parfois être présentée via ses collocsats comme étant simplement une étendue géographique, cette dimension a été marquée en rose saumon dans le tableau. Il s'agit alors notamment de la préposition « en » mais aussi de diverses villes importantes dans ce pays.

Un contexte qui semble avoir une forte importance dans ces collocsats est celui de la défense (coloré en rose vif dans le tableau), cela rejoint alors le sens d'une étendue politique pour ce pays. En effet, en tant qu'étendue politique, la Pologne dispose de frontières qu'elle doit défendre. Les collocations font alors apparaître des mots comme « missiles », « bouclier ». Il s'agit du souhait des États-Unis d'installer ce type d'armes sur le territoire polonais, ce qui irrite fortement la Russie. Cette affaire se déroule surtout en 2007 mais nous pouvons voir qu'elle trouve quelques échos dans le corpus 2009. Ce contexte explique également la forte association de la République Tchèque à la Pologne, puisque ce pays est aussi concerné par l'installation militaire.

L'histoire de la Pologne est évoquée dans ces collocations à travers les collocsats colorés en violet. « solidarnosc » correspond à un syndicat créé dans les années 1980 et dont le leader, Lech Walesa permettra à la Pologne de se libérer de l'emprise de la Russie soviétique. Les collocsats colorés en violet « Jaruzelski », « Geremek » et « Staline » montrent également une forte influence de la Russie dans le passé de la Pologne. Dans le corpus 2009, ce sont davantage les exactions allemandes et le fait que ce pays ait été envahi au cours de la seconde guerre mondiale qui sont notées à travers la présence du collocat

5- Deux cas particuliers : l'Italie et la Pologne.

« Hitler ». Le collocat « Tchécoslovaquie » peut également être considéré comme appartenant à ce contexte dans la mesure où ce pays n'existe plus depuis l'effondrement du bloc soviétique.

La Pologne est un pays sportif, ce contexte est présent à travers les collocats colorés en rouge. « Gorak » est le nom d'un pongiste polonais jouant dans un club français, « eurosport » diffuse quant à elle des compétitions cyclistes organisées dans la ville de « Pruskow ». Différents faits de société semblent aussi apparaître dans ce pays, notamment avec le collocat « avortement », il est alors question de la législation particulièrement contraignante qui y est en place et qui est connotée défavorablement dans le corpus 2007. La Pologne est aussi présentée comme une entité économique à travers le collocat « délocaliser » en 2009. Enfin, un fait divers est relaté par 4 des journaux du corpus IntUne en 2009, il s'agit d'un « tragique incendie » qui s'y est produit.

L'analyse des collocations du mot « Pologne » dans le corpus IntUne permet de mieux appréhender les différents sens potentiels que cette entité peut prendre. Il est alors possible d'en proposer le prototype sémantique suivant :

La Pologne est :

- Une étendue géographique.
- Une étendue culturelle.
- Un espace historique.
- Une étendue économique.
- Un espace politique avec des frontières à défendre.
- Un pays sportif.

Maintenant que nous avons un prototype sémantique pour ce pays, nous allons pouvoir en analyser l'actualisation dans la presse française en 2007 et en 2009. Pour cela, nous procéderons de la même façon que lors de l'étude sur l'Italie.

V.3.2- Analyse des lignes de concordances.

V.3.2.1- La Pologne comme une étendue géographique.

La Pologne est présente dans le sens d'étendue géographique en 2007, elle y est notamment un indice spatial qui permet de situer géographiquement la Lituanie, sa proximité

5- Deux cas particuliers : l'Italie et la Pologne.

est posée de fait avec ce pays. Elle est toujours présente dans le corpus 2009 en tant qu'étendue de terre. Il s'agit alors d'une masse physique dont on décrit la faune et la flore comme y étant encore préservée, quand bien même le développement économique de ce pays pose des problèmes de plus en plus importants au niveau environnemental.

Zone frontalière de la Pologne , au nord de Brest-Litvosk. A cheval sur la frontière, se trouve l'un des derniers vestiges européens de la forêt primaire que connurent les Celtes et les Goths.	li070209
ce petit pays (..) est frontalier de la Pologne et de la Biélorussie au sud, de la Lettonie au nord.	su240307

V.3.2.2- « Pologne » comme une étendue culturelle.

La Pologne peut également être une étendue culturelle. Deux contextes émergent principalement de ces concordances : d'une part, c'est un pays qui inspire différents romanciers de talent, voire un pays qui voit naître ces romanciers de talent sur sa terre. D'autre part, un contexte plus musical peut intervenir, dans ce cas, la Pologne est un grand pays culturel, qui a vu naître de nombreux compositeurs ou interprètes de grande qualité et qui a une grande tradition de musique classique. En cela, c'est un pays dont les différents concours musicaux sont renommés dans le monde entier. En 2009, son image est moins marquée par la prégnance de son excellence musicale, elle s'illustre dans divers domaines qui vont de l'architecture à la littérature en passant par la photographie.

L'écrivain partage désormais son temps entre la Pologne et les Etats-Unis	ou030307
Lauréat de concours en Russie, Estonie, Pologne , violon solo, clavecin.	su280207
talents, coup de cœur qu'elle découvrait lors de ses nombreux voyages en Pologne .	su250309

V.3.2.3- La Pologne comme un espace historique.

En 2007, l'un des principaux sens que la presse française attribue à la Pologne est celui d'un espace historique. Cet espace historique semble se construire essentiellement au cours du vingtième siècle. Dans ce contexte, le corpus donne à voir un pays qui a été très marqué par son passé communiste, avec de nombreuses occurrences qui se rapportent à cette partie de l'histoire polonaise.

chars soviétiques écrasent l'insurrection comme ils ont raison des émeutes en Pologne	fi230307
---	----------

D'autre part, c'est un espace dans lequel de nombreuses persécutions ont eu lieu sur les populations juives au cours du vingtième siècle, mais également bien avant l'invasion nazie :

5- Deux cas particuliers : l'Italie et la Pologne.

Vers 1880 pour fuir les persécutions dont les juifs étaient victimes en Pologne li170407

Lettres sur l'anéantissement des juifs en Pologne su240307

Enfin, la Pologne est également présente au niveau historique comme protagoniste malheureux de la seconde guerre mondiale : ayant été envahie par l'Allemagne nazie qui lui a fait subir de nombreuses atrocités, puis ayant ensuite été partagée par le vainqueur Russe de façon autoritaire. De cette dernière page de son histoire, la Pologne semble conserver une méfiance certaine vis à vis de l'Allemagne voire parfois une hostilité sourde.

Staline ampute la Pologne d'une partie de son territoire et en expulse les habitants. ou030307

En Pologne, dans l'ancien Gouvernement général, il découvre les atrocités commises par les allemands fi080207

La Pologne existe aussi en tant qu'espace historique dans le corpus IntUne 2009. Les principales parties de l'histoire de ce pays qui sont évoquées concernent alors son passé communiste, et notamment les déportations de populations allemandes qui ont eu lieu après la seconde guerre mondiale. Les persécutions des populations juives qui se sont produites lors de l'occupation nazie sont aussi fortement présentes lorsque la Pologne est considérée comme un espace historique. Cette occupation nazie est alors traitée sous le jour d'une invasion particulièrement dure pour la population juive et non plus pour la Pologne dans son ensemble.

URSS attraper dans sa nasse les pays de l'Est : Pologne, Tchécoslovaquie, Hongrie, Roumanie, Bulgari fi040209

Extermination des juifs, des Roms, des Sintis, des Yéniches du gouvernement général en Pologne su050209

L'actualisation du sens potentiel « espace historique » est donc plutôt stable entre 2007 et 2009 dans la presse française. On peut néanmoins noter qu'en 2009, l'accent est davantage porté sur la responsabilité conjointe des polonais dans les atrocités commises lors de ce passé violent, notamment à travers la mention du gouvernement général de ce pays. Tandis qu'en 2007, le peuple polonais était présenté comme une victime des régimes nazis et soviétiques.

V.3.2.4- La Pologne comme une étendue économique.

Le deuxième sens en termes de fréquence dans lequel le mot « Pologne » apparaît dans la presse française de 2007 est celui d'une étendue économique au dynamisme particulièrement remarquable. La vaste majorité des occurrences montrent ainsi un pays dynamique, à la croissance soutenue et stable, dont les fondamentaux sont bons : main

5- Deux cas particuliers : l'Italie et la Pologne.

d'œuvre formée, gestion budgétaire saine, croissance basée sur ces fondamentaux. Ce pays est également montré par la France comme lui « prenant » de nombreux emplois à travers les multiples délocalisations d'usines vers la Pologne.

en Pologne et en République tchèque, l'économie de ces pays croît deux ou trois fois plus vite fi160407

Cependant, la Pologne rencontre quelques difficultés, malgré ses réussites économiques, elle ne parvient pas à retenir ses travailleurs qui migrent vers les pays d'Union Européenne plus à l'ouest. Elle connaît également quelques difficultés sur le plan environnemental ou pour s'approvisionner au niveau énergétique.

départ (...) de polonais vers l'Europe de l'Ouest, la Pologne se retrouve en pénurie de travailleurs ou160207
livraisons de pétrole à destination de l'Allemagne et de la Pologne avaient curieusement cessé ou160307

C'est malgré tout un pays qui apparaît sous un jour très favorable dans la presse française en 2007 lorsqu'il est question de son économie, bien qu'encore pauvre, il comble son retard avec une rapidité qui fait une forte impression.

En 2009, ce sens devient le plus important en termes de fréquence. Son activation est alors diamétralement opposée à celle qui avait lieu en 2007. En effet, d'un espace dynamique basé sur une compétitivité et des fondamentaux stables, la Pologne devient un pays dont la monnaie s'effondre, loin de la stabilité affichée par la zone euro. De plus, il subit de plein fouet et plus gravement que la France la crise de l'industrie automobile dont il comptait de nombreuses unités de production. Au niveau structurel, son chômage augmente et son déficit commercial extérieur est en nette croissance par rapport à 2007.

ffondrement du zloty a fait chuter les prix de 30 à 70% en Pologne .Une aubaine pour les consommateurs al fi110309

ise automobile, les régions en subissent les effets(...) en Pologne . ou140209

Déficit du commerce extérieur de la Pologne en(..) hausse par rapport (...) en 2007 fi170209

L'image économique extrêmement favorable de la Pologne de 2007 semble donc s'être transformée en un portrait des plus sombres. Ce basculement de la vision donnée par la presse française de la Pologne est d'autant plus spectaculaire que ce sens d'étendue économique est le plus présent dans les journaux français de 2009.

V.3.2.5- La Pologne comme une étendue politique.

Le sens le plus important en termes de fréquence dans lequel la presse française montre le mot « Pologne » en 2007 est celui d'une entité politique. De nombreux contextes

5- Deux cas particuliers : l'Italie et la Pologne.

interviennent alors dans ce domaine. Tout d'abord, c'est une entité dotée d'un gouvernement qui est très largement décrié par la communauté internationale, et qui fait montre d'une certaine instabilité au niveau de son équipe ministérielle parmi laquelle les remaniements sont fréquents. Cette instabilité, et ce désamour de la communauté internationale pour la classe dirigeante polonaise sont en partie dues à la politique de mémoire prônée par ce gouvernement. La politique « anti-russe » de cet État vise à éloigner des postes publics toute personne ayant par le passé collaboré avec les instances communistes. Le climat de suspicion y devient ainsi la règle, et l'Union Européenne pèse fortement en défaveur de cette politique, soutenant en de nombreuses occurrences un député polonais qui s'y oppose.

La Pologne	fait valser ses ministres	li100207
Pologne	: nouvelle démission de ministre	ou090207
Pologne	. Plusieurs responsables politiques européens ont ovationné hier l'eurodéputé polonais Bronislaw Geremek, (ancien de Solidarnosc) menacé d'être déchu de son mandat.	su270407

D'autre part, le gouvernement polonais semble s'éloigner de la communauté internationale du fait d'une politique de plus en plus intolérante sur le plan des mœurs, que ce soit vis à vis des communautés homosexuelles ou concernant le droit à l'avortement. La presse française décrit également le peuple polonais comme tendant de plus en plus à une intolérance forte dans ces domaines. C'est aussi un pays qui rencontre un décalage croissant entre sa classe politique et son peuple. Les écarts gouvernementaux par rapport à la diplomatie internationale, mais aussi la volonté dirigeante d'une vaste politique de mémoire forcée, semblent en effet être dans un éloignement croissant des volontés de la majorité des polonais.

Ce pays entretient au niveau politique des rapports très complexes avec l'Allemagne. Des minorités allemandes sont présentes en Pologne et des minorités polonaises sont présentes en Allemagne. De plus, les deux états tendent à se rapprocher depuis l'ouverture au monde occidental des années 1990 aidée par l'Allemagne, néanmoins, la politique de mémoire du gouvernement de 2007 conduit à la résurgence d'une certaine méfiance polonaise vis à vis de ce voisin qui, au cours du siècle passé, lui a imposé une invasion terriblement destructrice.

Pologne	. Le chef du gouvernement, Jaroslaw Kaczinski a convoqué hier, en raison de ses déclarations sur l'avortement et les homosexuels, son vice-premier ministre et ministre de	su060307
La majeure partie de la Pologne	se débat dans de multiples difficultés et se laisse séduire par	ou030307

5- Deux cas particuliers : l'Italie et la Pologne.

	les discours populistes	
querelles décrédibilisent la Pologne	sur la scène européenne, (...) montrent que l'image de l'Allemagne auprès des polonais ne cesse de s'améliorer	fi260207
occultant le soutien de l'Allemagne (...)	l'adhésion de la Pologne à l'Union européenne	fo260207
Merkel (...) tenter de convaincre la Pologne	de surmonter ses réticences envers l'Allemagne	su170307

En tant qu'espace politique, la Pologne a des frontières qu'elle doit défendre. Dans ce domaine, elle présente dans la presse française le portrait d'un allié parfait des États-Unis, acceptant les négociations de ce partenaire dans le cadre de l'installation de systèmes antimissiles sur son territoire. Ce faisant, elle irrite fortement la Russie présente à travers la métonymie « Moscou ». Ce contexte est présent dans tous les journaux du corpus en de nombreuses occurrences.

États-Unis ont demandé(...) à la Pologne de servir de bases européennes à son bouclier antimissile. ec220207

Enfin, cet espace politique s'insère dans un espace plus vaste qui est l'Union Européenne depuis 2004. Les relations entre la Pologne et l'Union sont alors pour le moins complexes. D'une part, il semble que l'Union Européenne soit prête à donner beaucoup à cet état selon la presse française, dans différents domaines : fonds d'adhésion, aides à son activité de pêche. En outre, l'Union Européenne inclut cet état dans la zone « Europe du Nord » selon la presse française, zone qui jouit essentiellement d'une vision très favorable. D'autre part, la Pologne ne semble pas disposée à faire preuve de coopération vis à vis de cette institution. Les journaux français montrent le pays comme la bloquant à plusieurs reprises et comme avançant à reculons dans le domaine de la monnaie unique. Les difficultés entre Pologne et Union Européenne semblent dès lors selon la presse française surgir pour leur grande majorité du gouvernement polonais, c'est du fait d'une poignée d'hommes que ces relations sont aussi tendues. La Commission Européenne, alors parfois présente via la métonymie « Bruxelles » n'hésite plus dans ces cas à tancer ce pays.

mission semble vouloir faire plaisir à (...)	Pologne	ou270307
L'Espagne, la Pologne	et l'Italie sont les 3 grands bénéficiaires (...) du fonds européen pour la pêche	fi280307
depuis son entrée dans l'UE en 2004, la Pologne	s'est distinguée en bloquant plusieurs initiatives de l'Union européenne	su170307
ommissaire européen a mis en garde la Pologne	sur son homophobie	li260407
	La Pologne semble trouver un malin plaisir à jouer avec les nerfs de la Commission	fi170307
Bruxelles refuse que la Pologne	construise une autoroute	fi010307
Président de la banque centrale de Pologne	(...) une créature des frères jumeaux, préfère avancer	fi260207

5- Deux cas particuliers : l'Italie et la Pologne.

la date de... 2012

Ainsi, l'actualisation du sens « espace politique » se fait de façon plutôt défavorable pour la Pologne dans la presse française en 2007 : celle-ci reçoit beaucoup de l'Union Européenne mais fait peu d'efforts en revanche. Sa classe dirigeante pose de plus de nombreux problèmes à travers un discours jugé « populiste ».

En 2009, la presse actualise aussi le sens d'étendue politique pour le mot « Pologne ». Il y est néanmoins beaucoup moins présent qu'en 2007. Cela peut notamment s'expliquer par le changement de la classe dirigeante de ce pays, le gouvernement de 2009 ayant davantage l'assentiment de la communauté internationale. Deux contextes apparaissent alors : d'une part, le gouvernement actuel mène une politique économique et financière afin de faire face à la crise économique qui sévit au niveau international, d'autre part, l'archaïsme d'un ancien membre du gouvernement de 2007 est décrit, bien qu'étant en phase d'être dépassé.

Pologne (..) Kaczynski (..) longtemps passé d'un compte en banque, a fini par en ouvrir un fi030309

Pologne . Le gouvernement polonais se dit prêt à intervenir pour défendre le zloty fi180209

La politique de défense militaire reste présente en 2009. Deux contextes sont présents, d'égale importance en termes de fréquence. Il s'agit d'une part du projet d'installation de bouclier antimissile, mais la crise, née en 2007 entre Russie, Pologne et Etats-Unis, semble alors s'être apaisée suite notamment à l'élection du président Barack Obama. Des négociations ont ainsi été rendues possibles entre USA et Russie. Le deuxième contexte concerne l'effort de guerre sollicité par le gouvernement américain pour le conflit en Afghanistan. La Pologne y apparaît alors comme un état qui pratique une politique extérieure d'alliance forte avec les USA, étant parmi les pays qui réalisent l'effort de guerre le plus important.

Washington a en effet adressé un nouveau signal
à Moscou sur le projet de bouclier (...)

éventuellement installés en Pologne et en République tchèque. fi160209

Certains pays européens, la Grande-Bretagne et la Pologne se sont dits prêts à envoyer davantage de troupes fi230309

Au niveau de sa politique diplomatique, la Pologne apparaît dans la presse française comme un pays qui a apaisé ses tensions avec l'Allemagne (tensions alors vives en 2007), notamment via une politique d'apaisement menée par la chancelière Angela Merkel. C'est aussi dans ces journaux un lieu où le président français a rencontré le Dalaï Lama,

5- Deux cas particuliers : l'Italie et la Pologne.

conduisant à une dégradation des rapports entre la France et la Chine.

envenimait (...) relations entre l'Allemagne et la Pologne voisine, s'est enfin éteinte fi050309
rencontrant le dalaï-lama, début décembre en Pologne, le président français n'avait pas l'intention de se ou110209

Les rapports avec l'Union Européenne semblent toujours relativement problématiques, mais cette fois, la Pologne n'est plus mise au ban par les autres États membres de l'Union, au contraire, elle rejette des propositions en étant sur ce point en accord avec d'autres membres, notamment la France. Les problèmes semblent ainsi provenir en 2009 davantage du fonctionnement de l'Union Européenne quasiment personnifiée en la « Commission européenne » que de la Pologne.

Paris a réussi à fédérer dans son sillage 11
autres États membres, dont (...) la Pologne aboutissant ainsi au statu quo réglementaire. A fi170207
l'inverse, la Commission européenne, qui reste
hostile (...) a perdu une manche

Au niveau politique, la Pologne semble donc bénéficier d'une image plus favorable dans la presse française de 2009 par rapport à 2007. Ce n'est plus une zone dans laquelle risquent d'exploser de nombreux dangers, ses rapports avec l'Union Européenne rejoignent ceux de nombreux autres États membres dont la France. Elle reste néanmoins un des premiers alliés américains, contrairement à la position française.

V.3.2.6- La Pologne comme un lieu où des personnes vivent.

La Pologne est aussi présentée dans la presse française sous les traits d'une zone où des personnes vivent. Deux contextes principaux apparaissent dans ce cas, il peut s'agir de la description des mœurs des personnes habitant cette étendue, ou bien de faits judiciaires qui s'y déroulent. Dans un cadre judiciaire, toutes les occurrences du corpus IntUne 2007 concernent un procès se déroulant en France et dont l'un des suspects, d'origine française, est arrêté en Pologne. Au niveau des mœurs sociétares de la Pologne, l'image est sensiblement plus riche et marquée. La presse française fait alors de ce pays le portrait d'une société encore très archaïque, largement intolérante, avec une forte pression de l'Église catholique sur ces habitants. La plupart des occurrences concernent deux contextes. D'une part, la législation anti-avortement qui est très forte dans ce pays, les journaux français insistent alors sur le fait que très peu de pays d'Union Européenne suivent cette législation, et la comparent même à une législation digne d'un régime dictatorial. D'autre part, il peut être question d'une homophobie croissante de la société polonaise, à travers l'instauration de lois

5- Deux cas particuliers : l'Italie et la Pologne.

de plus en plus contraignantes pour cette communauté. Là encore, l'image qui est donnée par la presse française est largement défavorable.

dictature, le Portugal va-t-il rester avec Malte, la Pologne et l'Irlande l'un des 4 pays d'Europe où l'avortement est quasi impossible ?	li100207
Varsovie a pour objectif de dénoncer l'homophobie en Pologne	li270207

Le corpus IntUne 2009 montre aussi la Pologne comme étant une zone où des personnes vivent. Là encore, l'image a évolué sensiblement depuis 2007, les journaux français n'en donnent plus le portrait d'une société archaïque, cette dernière n'y est tout simplement quasiment plus évoquée, si ce n'est à travers un procès de fait divers et le rappel des attentats que l'ancien pape Jean-Paul II y a vécu. L'image en est alors globalement beaucoup plus positive, néanmoins, l'absence de traits sociétares polonais dans les colonnes des quotidiens français empêche de généraliser cette image qui se serait améliorée, on peut penser qu'elle s'est simplement améliorée par omission.

V.3.2.7- Un pays sportif.

La presse française actualise enfin le sens potentiel de « pays sportif » pour la Pologne en 2007. Dans ce cas, ce pays s'illustre dans divers sports relativement exotiques et peu pratiqués en France tels que le tennis de table ou le basket. Il s'illustre aussi dans le domaine du handball où la France dispose d'une équipe nationale ayant remporté plusieurs grandes compétitions. Le contexte le plus repris par la presse française est néanmoins celui qui voit la Pologne se faire confier l'organisation de l'Euro 2012 par l'UEFA, en partenariat avec l'Ukraine et au détriment de l'Italie. Cette décision des instances de football européennes est alors présentée comme une grande chance de développement économique et de rapprochement culturel offerte à la Pologne. Cependant, les journaux français insistent fortement sur le côté étonnant de ce choix, qui privilégie la Pologne et l'Ukraine au détriment de l'Italie, alors que le pays cumule de nombreux handicaps pour l'organisation de cette compétition.

Les stades en Pologne et Ukraine sont ainsi très éloignés les uns des autres	su160407
A la surprise générale, Ukraine et Pologne sont désignées pour organiser le tournoi	li190407
Leur candidature semblait en outre avoir été affaiblie par la menace de la FIFA d'évincer la Pologne de toute compétition officielle après le démantèlement de sa fédération par le gouvernement suite à une affaire de corruption	ou190407

5- Deux cas particuliers : l'Italie et la Pologne.

En 2009, le principal contexte actualisé dans ce sens l'est dans une vaste majorité des occurrences par le régional Sud Ouest. Il s'agit alors de l'organisation d'une compétition cycliste dans la ville de Pruszkow. Le contexte est assez neutre dans ce cas.

à Pruszkow (Pologne). Jean Pitaillier, président de la FFC jusqu'à fin février a été désigné parmi les su090309

V.3.3- Conclusion.

Ainsi, l'image qui est donnée de la Pologne par la presse française en 2007 est relativement mitigée. Le sens d'étendue culturelle semble y bénéficier d'une activation plutôt favorable : musiciens émérites, grand passé en musique classique. Le sens d'espace historique montre une histoire située à mi chemin entre Europe de l'Ouest et Russie, très fortement ancrée dans la mitteleuropa. En tant qu'étendue économique, la Pologne bénéficie d'une image des plus positives, quand bien même certaines difficultés subsistent, un important dynamisme vient les corriger, sa main d'œuvre y est de qualité, les polonais sont « bien formés » et « multilingues », parfaitement insérés dans une Union Européenne au niveau compétitif. C'est considérée comme une étendue politique et sociale que la Pologne voit son image être la plus dépréciée par la presse française : dotée d'un gouvernement non crédible pratiquant des politiques décriées par toute la communauté internationale. Cet état met en péril la paix acquise dans cette zone, d'autre part, sa société est avant tout dépeinte comme intolérante et archaïque. Quand il s'agit d'un pays sportif, l'image est plus nuancée : la Pologne est alors un pays qui gagne des compétitions européennes dans des sports qui ne sont pas les plus pratiqués en France. Mais lorsqu'elle se voit attribuer l'organisation d'une compétition de football majeure, les difficultés surgissent, outre le fait que le choix soit surprenant pour les journaux français, de nombreux obstacles techniques sont évoqués. Ce choix est considéré comme n'allant pas de soi mais plutôt comme étant une grande chance accordée de façon quasi débonnaire à la Pologne.

La Pologne semble en conséquence bénéficier d'une image sensiblement plus positive dans la presse française en 2009. L'amélioration en est la plus sensible dans les différentes activations des ses composantes politiques et sociales. En outre, l'activation sportive de cette entité semble plus neutre quand en 2007 elle tendait à être défavorable. L'image géographique, culturelle et historique en demeure relativement stable. Néanmoins, la forte dégradation du portrait de ce pays au niveau économique semble venir nuancer cette

5- Deux cas particuliers : l'Italie et la Pologne.

représentation française en progrès de la Pologne.

V.4- Conclusion.

L'intégration et l'unification d'ensembles supranationaux passe d'abord par la façon dont les différentes composantes de ses ensembles se voient les uns les autres. L'étude de ces composantes correspond ainsi pleinement aux problématiques du projet IntUne. Nous avons vu dans la première partie de ce chapitre que les pays et nations sont des concepts au sein desquels la question de l'identité est centrale. D'une part, cette identité permet à différents pays de se regrouper entre eux selon une proximité qui sera économique, géographique, culturelle ou encore historique. D'autre part, c'est par l'identité (notion d'une proximité commune entre les citoyens) que différentes nations peuvent se constituer, se différencier les unes des autres. Concernant un ensemble supranational comme l'Union Européenne, l'intégration ne peut fonctionner qu'en reproduisant cette recherche de différents points communs entre ses membres, que ce soit au niveau économique, culturel, ou géographique. A partir de là, il est intéressant de voir comment la presse française représente deux pays prototypes de l'ensemble étudié : l'Italie et la Pologne.

L'étude de deux pays différents de l'Union Européenne qui a été menée dans les seconde et troisième parties de ce chapitre semble alors montrer une certaine différence de traitement et de représentation au sein de la presse française de ces deux membres. Nous avons ainsi vu que l'Italie était représentée d'une façon stéréotypée dans les différents médias. Néanmoins, dans ces clichés et stéréotypes, ce pays reste toujours plus ou moins proche de la France (que ce soit à travers sa culture, son économie ou son histoire). Cette proximité entre les deux nations semble par ailleurs se renforcer en 2007 : l'Italie est économiquement plus proche de la France, et le principal point d'achoppement entre ces deux ensembles, le sport, pose moins de problèmes. Cette association n'est pas parfaite, de nombreuses difficultés subsistent : le sud de l'Italie rencontre des problèmes de légalité croissants, notamment. Cependant, il semble que ces deux entités peuvent s'associer, se comparer, et montrent une véritable proximité culturelle.

Le cas de la Pologne est sensiblement différent. D'une part, ce pays est toujours associé de façon forte et cohérente entre les deux corpus à l'ensemble des pays issus de

5- Deux cas particuliers : l'Italie et la Pologne.

l'ancien bloc soviétique. Cela se voit nettement à travers les collocations. La force de l'activation du sens historique est également présente aussi bien en 2007 qu'en 2009. D'autre part, elle présente un rapport qui est loin d'aller de soi avec l'Union Européenne. Si en 2007, les tensions semblent davantage être le fait d'un gouvernement « ubuesque », en 2009, elles demeurent, et la presse française les présente même comme étant partagées par de nombreux États membres dont la France. La Pologne peut alors devenir dans une certaine mesure un pays dont la France se rapproche. Elle a un fonctionnement sociétair e bien moins archaïque qu'en 2007 et une politique extérieure qui semble s'être apaisée. Des points de différence majeurs existent néanmoins toujours en 2009, et le chemin à parcourir pour ce pays afin d'être pleinement intégré à l'ensemble Union Européenne, ensemble créé par différents pays d'Europe de l'Ouest, semble important.

Cependant, ce qui émerge le plus de cette comparaison diachronique de deux pays à travers une analyse du corpus IntUne est en fait la représentation très stéréotypée qui peut être faite d'un pays dans la presse française et le fait que cette représentation puisse subir différents biais (le résultat d'une coupe du monde de football, par exemple) qui en atténuent l'objectivité. En outre, à travers la description qu'elle fait de la Pologne en 2007 et en 2009, la presse française tend à montrer les problèmes croissants d'une institution, l'Union Européenne. D'une part, ses différents États membres peuvent être problématiques du fait de leurs propres politiques intérieures ou extérieures. D'autre part, ces États membres peuvent avoir de plus en plus de difficultés à « obéir » aux ordres d'une entité quasiment personnifiée, la Commission européenne, d'une façon de plus en plus largement partagée, quelle que soit « l'Europe » (Ouest ou Est) dans laquelle ils se situent.

Ces représentations stéréotypées de pays appartenant à l'Union Européenne se retrouvent ainsi de façon forte dans l'ensemble des médias collectés au sein du corpus IntUne. Nous pourrions maintenant nous intéresser à la présentation qui en est faite au sein de reportages télévisés d'une chaîne privée et publique en France. Cette analyse sera l'objet du chapitre suivant.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

VI.0- Introduction.

Le corpus IntUne se compose de langue écrite à travers les cinq journaux quotidiens du corpus presse, mais également de langue orale à travers les transcriptions de journaux télévisés de deux chaînes, TF1 et France 3. Cette langue orale se rapproche en fait plutôt d'une sorte d'écrit oralisé, se rapprochant du « quasi speech » de Sinclair (1991:16). Le « quasi speech » consiste en des textes préparés écrits, dans le but de simuler un échange oral dans des conditions artificielles, tels que les scripts de films, les textes dramatiques. Les énoncés que l'on peut trouver dans un journal télévisé appartiennent à cette catégorie.

Le nombre de mots du corpus télévisé en 2007 ne représentait pas plus de 420 000 tokens sur un ensemble de plus de vingt millions. Procéder au même type d'étude quantitative dans la recherche de collocats ou de répétitions de structures saillantes ne revêt en conséquence pas la même pertinence que pour le corpus presse. Le projet de recherche IntUne associe ainsi dans son groupe « média » la démarche de recherche quantitative à une démarche d'analyse qualitative. Cette dernière semble mieux adaptée aux données du discours télévisuel.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

C'est dans cette démarche d'analyse qualitative que le chapitre suivant va étudier deux reportages issus du corpus français 2009, concernant un sujet identique avec un retentissement fort au niveau européen : le tremblement de terre qui s'est produit en Italie en avril 2009. Tout comme pour le chapitre précédent, cette analyse fait le postulat qu'en étudiant la façon dont les différents pays européens se représentent les uns vis-à-vis des autres dans les médias, on peut appréhender avec davantage de pertinence leur éventuelle intégration au sein d'un ensemble supranational. L'étude d'un sujet touchant l'Italie et traité le même jour dans les deux journaux télévisés français permet de voir comment ces journaux représentent ce pays, et donc, si ce pays et la France semblent pouvoir être intégrés au sein d'une entité supranationale.

Dans un premier temps, le cadre méthodologique de l'analyse sera brièvement décrit. Puis il sera procédé dans un deuxième temps à l'analyse des deux reportages du 6 avril au soir présentant la catastrophe. Cette analyse procédera de façon linéaire, partie par partie du discours télévisé, et chaque analyse d'une partie du discours télévisé sur France 3 et TF1 sera suivie d'une étude comparative entre la présentation de ces deux chaînes.

VI.1- Cadre méthodologique.

Le discours télévisuel présente une différence flagrante avec la presse, concernant la position dans laquelle il met son récepteur, le téléspectateur. En effet, le récepteur d'un message écrit dans un journal peut choisir quel texte il va lire, quelle partie même du texte il lira, s'attarder plus ou moins sur certains mots, etc. Sa lecture est beaucoup plus active que la réception d'un message télévisuel. La réception d'un message télévisuel ne semble en fait la plupart du temps ne poser de problèmes qu'au niveau technique (Viallon 1996:5).

Le spectateur ne semble pas s'interroger lui-même quant à cet acte de s'informer par le média des journaux télévisés. Les générations antérieures ont du apprendre à lire pour s'informer, mais au niveau de la ressource audiovisuelle, tout effort semble superflu. Le téléviseur est souvent placé pour pouvoir être regardé dans un confort maximum, la présence de télécommande implique un certain pouvoir sur la technique et sur les ressources proposées (le spectateur a l'impression de pouvoir modifier son choix à tout moment). Le

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

message en lui même ne semble pas poser de problèmes, le son est évident et l'image paraît aller de soi. Il pourrait même sembler saugrenu d'envisager qu'il puisse y avoir là la moindre difficulté.

Pourtant, le téléspectateur n'a pas appris à regarder la télévision, on pourrait penser qu'il s'agit d'un comportement inné de l'être humain qui ne nécessite pas d'apprentissage (au même titre que manger ou respirer), ou alors objecter qu'il s'agit d'un faux problème, toute personne étant capable de comprendre sans difficulté le message télévisuel. Cependant, il convient de garder à l'esprit que toute diffusion télévisée nécessite la mise en œuvre de moyens financiers très importants, que chaque diffusion d'un journal télévisé prend place dans un moment stratégique pour la chaîne et qu'elle doit répondre à de nombreux impératifs. En cela, le message télévisuel ne peut être d'une simplicité confondante et différents points peuvent en être analysés plus précisément. Dans ce chapitre, nous nous baserons plus particulièrement sur trois caractéristiques de ce message : le type de communication qu'il implique, les éléments non linguistiques (vidéo, bruitages, musique, etc.) et les éléments linguistiques de ce message.

VI.1.1- La relation de communication télévisuelle.

Tout acte de communication fait intervenir un émetteur et un récepteur d'un message. Dans le cadre du message télévisuel, l'émetteur est à la fois un et multiple.

VI.1.1.1- L'émetteur.

D'une part, le plus souvent, l'émetteur est présent à travers la mise en avant d'un homme seul et unique. C'est le cas le plus fréquent dans les journaux télévisés français. Le présentateur dispose d'une forme d'omnipotence, il peut convoquer sur son plateau des personnalités importantes, maîtriser le déroulement de l'information. Cette fonction d'homme pivot se trouve aussi renforcée par son nom qui s'inscrit en bas de l'écran, ou encore par le cadrage parfois serré sur ce personnage.

Néanmoins, cette omnipotence est nuancée par l'importance que l'émetteur se doit d'accorder au récepteur, par exemple, sa production linguistique se doit de chercher à

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

s'adapter à la compétence supposée du récepteur. Des études comparatives sur la longueur des phrases, la complexité de leur structure ou de leur vocabulaire dans les JT de TF1 et FR2 d'une part et de FR3 en soirée, Arte d'autre part montrent que les journalistes, bien qu'ayant des compétences similaires (les transfuges entre chaînes sont nombreux) ont une production différente et ont donc fait des hypothèses sur leur public. Cela s'inspire notamment des travaux de Kerbrat-Orecchioni (1980 :17) qui montre que dans toute interaction verbale, l'émetteur subit des contraintes, ici, celle des compétences linguistiques et para-linguistiques de son récepteur.

Toujours dans ce domaine de l'adaptation de l'émetteur au récepteur, le présentateur doit notamment adapter son message au savoir culturel supposé du récepteur (Ibid:19). Quand un nouvel événement est abordé, l'émetteur doit décider du fait d'expliquer ou non les toponymes et noms nouveaux qu'il va utiliser, c'est à lui de chercher si oui ou non ces informations font parties d'un contexte référentiel communément connu. Malgré une forme de pouvoir absolu de cet émetteur unique, il semble donc qu'il doive subir quelques contraintes dans la mise en forme de son message.

D'autre part, l'émetteur est en fait souvent constitué d'un groupe de personnes important, tel un iceberg dont le présentateur ne serait que la partie émergée. Les présentateurs de journaux télévisés disposent ainsi le plus souvent d'une équipe importante pour réaliser leur message, cela va des techniciens aux différents journalistes de leur chaîne, en passant par les équipes de diverses agences de presse ou les envoyés spéciaux présents de façon continue dans différents pays jugés stratégiques. Outre les apports de toutes ces personnes, le recours à l'image d'archive peut venir enrichir encore le message proposé. On se retrouve donc face à un conglomérat de discours différents, une succession de filtres et une multiplicité d'émetteurs (Viallon 1996:15).

Le présentateur appartient également à un autre groupe : la chaîne de T.V., avec ses caractéristiques propres. Les statuts des chaînes sont différents, et on ne peut produire le même discours selon que l'on soit sur une chaîne où l'audience est le seul credo, et une autre où les critères retenus sont de l'ordre du service public.

Il convient de garder à l'esprit ces considérations concernant l'émetteur du message

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

d'un journal télévisé lorsque l'on souhaite l'analyser.

VI.1.1.2- Le récepteur.

Un autre acteur important de la communication télévisuelle est le récepteur. Du point de vue de l'émetteur, le récepteur du message télévisuel représente une entité abstraite qu'il se doit d'imaginer, afin d'avoir une communication plus efficace. Il faut donc passer outre les difficultés techniques (s'exprimer comme face à une personne en étant face à une caméra), et inhérentes au fait d'avoir une vaste multiplicité de récepteurs possibles (on ne propose pas le même message à des personnes d'origines, d'âges, etc. totalement différents). Divers outils viennent à son aide dans cette tâche. Actuellement, il existe le système de mesure Mediamat⁷⁴ auquel viennent s'ajouter des sondages téléphoniques ou des questionnaires d'enquête. L'importance de ces mesures est telle que l'expression « dictature de l'audimat » est entrée dans l'usage courant. Il y a aussi une part de construction du message qui appartient au téléspectateur. L'émetteur présuppose la compétence du récepteur : il produit un texte, ouvert ou fermé dans sa compréhension, qui joue sur l'habitude, ou sur la cible, c'est le cas de la fameuse ménagère de moins de cinquante ans (Eco 1985:63).

L'absence inévitable du récepteur lors de la production de ce message est néanmoins flagrante sur les plateaux des JT et différentes initiatives visent à en minorer l'effet. Ainsi, l'utilisation d'un témoin va permettre au téléspectateur de s'y identifier, et le choix de ce témoin sera donc réalisé avec soin en fonction du type de téléspectateur que l'on souhaite toucher en priorité. Dans tous les cas, cet acte de communication n'est pas une donnée naturelle et ne va pas de soi, il pose de nombreuses questions dont les solutions sont apportées par de multiples professionnels afin d'atteindre une efficacité maximale pour la chaîne qui le diffuse (Viallon 1996:21).

VI.1.2- Les éléments non linguistiques du message.

Différents éléments autres que le linguistique interviennent dans la construction du message télévisuel. Il peut s'agir d'éléments sonores (bruits, musiques, sons), mais également

⁷⁴ [Http://www.mediametrie.fr](http://www.mediametrie.fr)

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

et surtout des images vidéos et de leur montage. Nous allons ici préciser davantage dans une première partie quels sont les différents éléments sonores et la fonction que ceux-ci peuvent remplir dans un discours télévisé. Puis, nous préciserons l'usage qui peut être fait de la vidéo, et les rôles joués par les différents choix effectués dans ces usages.

VI.1.2.1- Les éléments sonores.

En règle générale, la présence d'une ambiance riche en bruits divers paraîtra toujours plus normale qu'un silence autour du commentaire d'un reportage. Le décodage des bruits peut se faire de deux manières, soit par leur co-occurrence avec d'autres signes, notamment visuels dans le cadre de messages synchrétiques pour les bruits on (l'image montre un téléphone que le son fait sonner), ou soit par comparaison avec des bruits enregistrés dans la mémoire du récepteur pour les bruits off. Dans une prise de vue extérieure, un vrombissement de moteur sera assimilé à une voiture, même si on ne la voit pas. Ainsi, rares sont les reportages sans bruits d'ambiance, cette nécessité de créer une atmosphère peut même aller jusqu'à conserver les paroles inaudibles d'une personne étrangère interviewée sous la traduction. Le bruit en fait, crée l'espace (bruit de feuilles, on est en pleine nature), le temps ou l'histoire (Ibid :45).

La musique peut aussi être utilisée dans le JT, notamment lors du générique ou de jingles séparant différents temps de ces journaux. Françoise Escal en parlait comme de « *choix restrictifs opérés dans une masse considérable de sons ou de bruits potentiels, sélection inculquée par la communauté musicale/linguistique* » (Escal 1979:11). Elle appartient donc à la culture et non à la nature (Lévi-Strauss 1964:30), ainsi, elle est fortement liée à son temps de production. D'un autre côté, on a l'oreille de sa culture musicale, ce qui nous rend sourd aux autres musiques et explique la difficulté pour l'émetteur télévisuel du choix d'une musique. Une profession « *illustrateur musical* » est dévolue à ce rôle. Différents éléments d'analyse peuvent nous permettre de l'appréhender. La musique a différentes fonctions à la TV, sa première est de soutenir, d'accompagner l'image, de la renforcer (c'est le cas des premiers films muets également). Elle peut aussi tenir un rôle de séparation, pour distinguer une suite de paroles d'une autre suite de paroles (pensons aux *jingles*, très utilisés dans le JT de France 3), elle permet aussi d'attirer l'écoute du téléspectateur sur un moment qu'il reconnaît plus facilement par l'introduction musicale.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

Le silence est tout simplement extrêmement rare à la télévision, on peut y voir, tel Viallon (1996:52) une forme de recherche de communication phatique par le média télévisuel, le but est alors de garder coûte que coûte la communication en marche, d'en éviter toute coupure qu'un silence pourrait créer.

VI.1.2.2- La Vidéo.

Il existe différents types de codes humains présents à la télévision. Il peut s'agir des artifices, tels l'habillage ou le maquillage auxquels un soin important est apporté quand il est question de la personne qui présente un JT. La tendance sera par exemple de préférer les tons bleus qui sont plus nets à la TV aux tons rouges, de même, tel présentateur remplacera ses lunettes par des lentilles pour un contact plus direct avec le téléspectateur. Les codes cinétiques lents, les mimiques et postures adoptées par les personnes présentes à la télévision peuvent également être modifiées pour créer différents effets sur le récepteur. Les codes cinétiques rapides, se composant notamment du regard, de la kinésique, sont également très utilisés par les personnes présentant un JT : ils utilisent des stylos, ont des liasses de papier à la main, etc. Au niveau de la proxémique, les choix s'opèrent essentiellement sur un intérieur et un extérieur qui seront filmés. L'espace intérieur peut être créé de façon différente : à deux ou trois dimensions, comporter différents accessoires, disposés de façon variée. Ainsi, dans les JT, très souvent, le présentateur est au centre d'un dispositif qui peut prendre la forme d'une étoile avec murs d'écrans, ou fenêtres immenses qui montrent une ouverture sur le monde. Rien n'est définitif dans ces dispositifs et ils sont sans cesse renouvelés afin d'atteindre une efficacité toujours plus grande (Ibid 58-59).

L'image est également très construite dans le message télévisuel. Quand on regarde la télévision, on voit le monde à travers un cadre dont la forme a été prédéfinie. Toute image télévisuelle existe dans une norme imposée (généralement, le rapport 3/4). Quand bien même ce rapport est hérité de siècles de représentation picturale, il n'en est pas moins arbitraire (Joly 1994:6).

La prise de vue constitue un premier choix au niveau de l'image télévisuelle : que va-t-on mettre en valeur, que va-t-on nier ? Le même événement filmé par différentes chaînes donnera lieu à différentes images. Le reportage n'est ainsi qu'une réalité parmi une

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

infinité d'autres possibles. Avant d'être une transposition du réel, l'image télévisuelle est une négation par l'absence de multiples facettes de la réalité (Viallon 1996:65). La télévision doit aussi en permanence gérer la quantité d'informations et de données qu'elle délivre au téléspectateur. Il s'agit alors d'une gestion de l'espace et du temps : tel texte ne sera lisible à la TV que s'il est écrit assez gros et s'il reste inscrit suffisamment de temps. D'un point de vue diachronique, cette quantité d'informations offertes par l'écran ne fait qu'augmenter : il semble que le téléspectateur se familiarise de plus en plus à la communication télévisuelle et augmente ainsi ses capacités de visionnement. Les innovations techniques permettent également de multiplier les sources aussi bien auditives que visuelles.

A la manière d'une page de journal, l'écran, considéré comme une surface signifiante, organise le sens par la disposition même des informations qu'il transmet. On s'accorde généralement à penser que la peinture occidentale est essentiellement centrée, l'attention est alors plutôt portée sur ce qui est centré (Arnheim 1973). L'incertitude des limites exactes du téléviseur, la recherche du rendu immédiat de signification et le peu de temps disponible pour lire une image renforcent ce principe à la télévision.

La place du présentateur du JT est aussi assez signifiante. En Allemagne, les deux principales chaînes le placent à droite de l'écran, elles ont fait ce choix sur la base d'une étude qui stipulait que notre cerveau accorde 70 % de son attention à la gauche de l'écran, et 30% à la droite, elles font donc le choix d'une mise en retrait du présentateur au profit de titres, diapositives, images, etc. (Viallon 1996:71). En France, la position privilégiée est l'axe central pour permettre l'affichage d'une vignette, dispositif qui met en valeur le journaliste au détriment de l'information qu'il transmet. Les formes ont toujours eu une signification, et en tant que telles, elles sont aussi utilisées par la télévision. Les lignes droites et les courbes sont à la base de nombreuses figures sous une forme simple ou complexe, que l'une représente l'homme et l'autre la femme sont des lieux communs souvent évoqués. Le cercle clos exprimera l'enfermement ou la protection, le demi-cercle l'ouverture au monde.

Lorsque l'image s'anime, elle se compose de différents plans. Ceux-ci peuvent être fixes et plus ou moins proches du sujet filmé, mais ils peuvent également être mobiles avec des mouvements latéraux ou des zooms. Leur choix peut obéir à différentes motivations dans un JT : un plan rapproché créera une forme d'intimité absente avec un plan général, les

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

mouvements dans un même plan, le changement brusque d'échelle des plans utilisés donnera un effet de dynamisme et attirera le regard du téléspectateur. La durée du plan est un élément essentiel, il peut se situer entre 1/25^e de seconde et plusieurs minutes. Cette durée est surtout à prendre en compte de façon relative, dans un film d'auteur, un plan d'une minute pourra être jugé court, dans un reportage de JT, ce sera exceptionnellement long. Cette durée devient une sorte de signe d'importance de l'émetteur au récepteur au même titre que la typographie grasse pour l'écrit ou l'accentuation pour l'oral (Ibid:75).

A cheval entre l'image fixe et le montage, on trouve les images construites, de plus en plus utilisées dans les JT. Sur l'écran, vient alors se rajouter pour un moment un insert, un élément d'information produit au générateur d'écriture. La tendance croissante de cette pratique peut participer au mouvement général de l'accélération du rythme et de la multiplication des informations.

Le montage peut avoir une fonction narrative, mais aussi une fonction expressive. On trouve parfois aussi des fonctions plus complexes que l'on nomme syntaxiques, sémantiques et rythmiques. La fonction syntaxique correspond alors aux effets de liaison ou de disjonction, d'alternance et de linéarité. Les montages « alterné » ou « parallèle » qui expriment la simultanéité et la comparaison en sont les meilleurs exemples (Aumont et al. 1983:44).

La tendance au raccourcissement des plans crée une accélération du rythme de l'image. La TV américaine a des plans d'une durée moyenne de 3 secondes et demies, et les formats français ont tendance à vouloir copier ce fonctionnement souvent qualifié de plus dynamique. Les téléspectateurs sont sensibles à cette notion du rythme, il suffit d'entendre deux personnes comparer un épisode d'« inspecteur Maigret » à un épisode des « experts » pour s'en rendre compte, de même comparer l'ambiance moyenne d'Arte à celle de TF1 ou de M6 conduira généralement à conclure que la première est plus « ennuyante » que les deux autres, tout simplement parce que les images y sont en général plus longues. La plupart du temps, le téléspectateur préfère le rythme à la contemplation (Aumont 1988:41).

La TV joue énormément sur l'attraction de l'œil pour tout mouvement (l'être humain est carnivore et est en conséquence attiré par le mouvement, qui aidait l'homme primitif à

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

trouver sa nourriture, et aussi à échapper au danger). Cette recherche permanente du mouvement se situe à trois niveaux : un drapeau flottant au vent sera préféré à un drapeau statique par un journaliste, les mouvements de caméra seront ajoutés sans justification réelle, enfin, certains plans de coupe ne s'expliquent par aucune raison technique. Tout cela crée un mouvement artificiel qui a pour conséquence de contribuer à annihiler la capacité du téléspectateur à réfléchir (Viallon 1996:83).

VI.1.3- Les éléments linguistiques.

VI.1.3.1- Les traces de l'émetteur dans l'énonciation.

La situation d'énonciation peut être présente dans le discours télévisuel dans différentes mesures. Diverses personnes peuvent ainsi se manifester. Benveniste considère la troisième personne comme la « non personne », c'est celle qui est absente (Benveniste 1966). C'est alors le discours privilégié de la fiction où l'histoire se raconte toute seule. D'autres genres, comme le documentaire ont aussi un émetteur souvent apparemment absent. La télévision, en l'utilisant, a tendance à vouloir disparaître en tant que sujet de l'acte d'énonciation, elle fait sentir sa présence en tant que canal. Cela présuppose l'existence d'un réel à montrer, d'un réel qui se montre de lui même.

La deuxième personne du pluriel exprimant en français la forme dite de politesse, le « vous » devrait être le mode standard de la communication interpersonnelle médiatisée (Maingueneau 1981). En réalité, on constate que cette forme est surtout employée dans deux cas : comme transition entre les séquences et comme complément du « nous » émetteur, les deux étant souvent simultanés, comme si ces moments n'étaient que des instants privilégiés, phatiques, n'ayant pas d'autre fonction que de créer le contact, ou de vérifier qu'il est bien établi (Viallon 1996:95). Au niveau de l'image, le vous est représenté par exemple avec un micro tendu vers la caméra. Le vous peut être affirmé en tant que tel, l'émission « à vous de juger » met en avant le téléspectateur dès son titre. Dans certains cas plus rares, il peut ne plus concerner le téléspectateur, « j'ai une question à vous poser », dans ce cas, le téléspectateur est représenté par le « je » et non plus par le vous, ce qui est encore plus proche comme relation, or cette émission a connu un succès important.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

Le message télévisuel est le résultat d'une longue suite de compétences et ceux qui ont le droit de parole ne seraient rien sans leurs équipes. Il est donc de tradition que le « nous » ne signifie pas le simple « nous » de majesté, mais exprime la position de l'ensemble de l'équipe productrice. Cependant, les autres valeurs du « nous » ne sont pas négligées, notamment celles qui associent l'émetteur et le récepteur, c'est à dire le journaliste présentateur et le téléspectateur. Il suffit pour s'en persuader de penser au nombre de reportages qui sont introduits par : Que mangeons-nous ? Quelle taille faisons-nous ?, etc. Cette communauté forcée de sentiments, d'actes, est un moyen d'accrocher le destinataire, d'éveiller son intérêt pour quelque chose qui « doit » le concerner.

Le « je » ou la première personne du singulier, s'il est rarement présent de façon explicite, peut être suggéré par la mise en place de différents effets de réel. Ces effets peuvent parfois être introduits par une caméra subjective. Une agression est alors reconstituée par une caméra portée sur l'épaule ou un cadreur qui court, puis tombe. Le téléspectateur court, se retourne, court encore, tombe, revit ce scénario. Dans ce type de démarche où en quelques sortes, la fiction investit l'information. La dramatisation et la scénarisation sont d'autres moyens de renforcer la participation du téléspectateur. La plupart des séquences sont basées sur des stéréotypes narratifs ou actentiels connus du téléspectateur et dont il attend une actualisation. L'amplification est également à la base de nombreuses productions : parmi des personnes affrontant les forces de l'ordre dans la rue, ce seront les plus violents qui seront filmés, les cadrages aidant à la production de l'effet (Ibid:97).

VI.1.3.2- Les marques spatio-temporelles de l'énonciation.

Les séquences du journal télévisé sont construites d'un point de vue spatial sur une opposition que l'on rencontre dans d'autres genres entre le plateau et l'extérieur, le hors plateau. Ce deuxième ne se définissant que par rapport au premier. Le hors plateau du JT est fortement tributaire du plateau : tout est fait pour que le changement soit à la fois fort (il faut que le téléspectateur ait l'impression d'avoir parcouru le monde en trente minutes) et insensible : les chevauchements de son, les fondus enchaînés fréquents, les liens syntaxiques (à la question du présentateur, va succéder la réponse de l'envoyé spécial), les liens logiques (les lancements se transforment de plus en plus en une affirmation du plateau qui est confirmée à l'extérieur), tout concourt à faire que le « ici » et le monde ne fassent qu'un. La

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

succession est l'expression d'une continuité territoriale. En fait, ce qui est mis en scène, c'est le don d'ubiquité du média : la télévision peut être partout et en même temps, les murs de TV que l'on trouve sur de nombreux plateaux figurent bien cela. Le JT espère ainsi satisfaire le désir de dépaysement du téléspectateur. L'espace devient un moyen de communication, le studio exprime une vision du monde, d'une culture.

Les déictiques temporels linguistiques sont facilement repérables : « maintenant », « hier », « demain », sont autant de marques énonciatives. En revanche, si les images sont rarement datées, c'est leur diffusion au sein du continuum temporel de la télévision qui les actualise : qu'elles appartiennent à un passé proche ou lointain, elles sont au présent quand le téléspectateur les regarde. Tout est fait pour faire oublier, ne serait-ce que pour des raisons techniques, qu'elles appartiennent au passé. On a donc affaire à la rencontre de deux temporalités, celle de la réalité et celle de la télévision, cette dernière faisant tout pour allonger le présent.

Au niveau des journaux, l'ensemble de leur production linguistique tend à faire durer le présent 24 h. Les déictiques restreints se multiplient au détriment des déictiques élargis (opposition aujourd'hui/hier plutôt qu'aujourd'hui/autrefois). La notion de direct est également assez extensible pour les journalistes qui n'hésitent pas à annoncer « nous retrouvons tout de suite X sur place », s'ensuit un reportage et le commentaire de l'envoyé spécial en vrai direct conclut. Coincé entre ces deux présents, le reportage par osmose, par contamination, devient présent à son tour. Cette force d'attraction pour le présent détermine une grande partie de la télévision : tous les reportages se doivent d'être d'actualité (Ibid:106). Pour conclure, la télévision peut jouer sur des registres très contradictoires : dire « nous » et montrer « je », ramener à une surface d'écran son lieu d'énonciation qui est le « village global », réduire les différentes temporalités à son temps, présent éternel et cependant toujours renouvelé.

Maintenant que nous avons précisé un cadre théorique d'analyse de données audiovisuelles, nous allons en proposer une étude dans deux reportages télévisés.

VI.2- Analyse comparative des traitements de France 3 et TF1 pour un même sujet.

Les deux chaînes placent toutes deux l'information concernant le tremblement de terre du 6 avril dans la région des Abruzzes en position prioritaire dans leur journal. Dans les deux cas, c'est le premier titre traité, et plusieurs reportages lui sont consacrés. Le JT de France 3 y alloue 10,03 minutes soit environ 31% de sa durée complète. Le JT de TF1 y consacre 13,15 minutes soit environ 38% de sa durée. Proportionnellement, c'est donc TF1 qui accorde le plus d'importance à ce sujet. Les structures des deux débuts de journaux sont sensiblement semblables, on pourrait les résumer comme suit :

TF1	France3
Titres	Titres
Reportage général de présentation de la catastrophe par Michel Scott.	Reportage général de présentation de la catastrophe par Stéphanie Pérez.
Reportage sur les destructions dans les villages alentours par Maurice Olivari.	Reportage sur les personnes sans abris, le côté dramatique, et les difficultés pour les secours à répondre à tous les besoins par Bertrand Boyer.
Reportage sur les sans abris, leur panique, et les nombreux moyens de secours mis à leur disposition par Liseron Boudoul, elle intervient en direct de l'Aquila à la fin du reportage.	Duplexe avec Bertrand Boyer sur les personnes sans abris sans refuges qui paniquent, il clôt sur le fait que le gouvernement italien refuse l'aide de l'Union Européenne.
Reportage didactique de Fabrice Collaro par rapport aux risques de tremblements de terre dans le bassin méditerranéen et notamment en France.	Interview de Patrick Hesters sur le plateau du JT, explication du fonctionnement des tremblements de terre en général.
Interview en duplex avec Maurice Olivari. Se pose la question du bilan, la question des secours également, insiste sur la difficulté de secourir toutes les victimes.	

(Figure 26, structure des deux débuts de journaux sur TF1 et FR3)

Afin d'analyser de façon comparative les deux traitements d'un même sujet, il sera procédé à l'analyse des titres, de la première annonce par la présentatrice du premier reportage, et enfin celle du premier reportage en lui même.

VI.2.1- Les titres.

Les titres sont le premier message transmis lors du journal, ils doivent capter l'attention du téléspectateur en lui annonçant ce qui va suivre dans l'édition. Les deux chaînes ont un procédé de présentation relativement semblable au niveau global, le tremblement de terre est le premier sujet évoqué, le générique du JT continue pendant l'annonce des différents titres tandis que des images illustrent les propos des présentatrices.

VI.2.1.1- France 3.

Dans le journal de France 3, ces titres se décomposent en huit plans.

Le premier plan dure 1,10 secondes, il montre une vue générale du plateau du journal, ainsi que la présentatrice qui s'y déplace en plongée. Celle-ci est habillée d'un jean bleu et d'une blouse blanche : l'ensemble de la tenue dégage une impression de décontraction sans ostentation, il n'y a pas de fantaisie, peu de bijoux (boucles d'oreilles discrètes), le maquillage est très naturel. Ses cheveux sont lâchés, ce qui renforce l'impression d'une présentatrice non ostentatoire. Elle dispose d'un ensemble de feuilles à la main ainsi que d'un stylo. Cette référence explicite à l'écrit confère au message oral qui est de fait diffusé par la présentatrice une part du prestige que l'on accorde généralement à l'écrit, la référence à l'écrit dénote ainsi une dimension plus sérieuse de l'information.

Le plateau se compose de deux parties que l'on voit au début de la séquence des titres, ces deux parties se composent de deux cercles d'une circonférence différente. Les formes dominantes sont courbes, donnant une certaine douceur à l'ensemble, les tons principaux qui s'en dégagent sont le blanc, le bleu et quelques touches d'orangé, qui confirment un aspect non agressif et lumineux.

Le deuxième plan a une durée beaucoup plus longue, de 6,82 secondes. Il débute par un zoom et aboutit à filmer la présentatrice en plan américain (la caméra filme la silhouette à partir de mi-cuisses). Elle dispose toujours des feuilles et du stylo, et marche vers une partie du plateau ce qui confère une impression de dynamisme et peut concourir à capter le regard et l'attention. Elle ne lit pas les feuilles qu'elle a à la main ni ne se sert du stylo, néanmoins, à la fin du plan, on voit une amorce de mouvement pour passer à la feuille

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

suivante. Les deux « cercles » qui définissent l'espace du plateau peuvent prendre leur fonction. Le plus petit semble être un poste pour accueillir la présentatrice lors de l'annonce de ces titres. Le plus grand apparaît derrière la présentatrice lorsqu'elle est montrée en plan américain. Il s'agit d'un bureau entouré d'un cercle, à l'arrière de ce bureau, on voit deux grands écrans sur lesquels s'imprime l'image de fin du générique du journal TV. Ces deux écrans sont eux même « accrochés » sur une sorte de véranda, ce qui permet d'obtenir une double ouverture sur le monde extérieur : l'environnement proche du plateau est suggéré par les lumières changeantes et les formes floues aperçues derrière les vitres de la véranda, tandis que les deux écrans peuvent à tout moment contenir les images des reportages qui vont être diffusés pendant le journal.

Les deux écrans s'animent deux secondes avant la fin du plan, ils se chargent alors d'images de l'événement évoqué par la présentatrice, et l'enchaînement vers le second plan se fait via un zoom sur ces deux écrans. Voici deux images fixes issues de ces deux premiers plans.



Une deuxième partie se déroule ensuite, elle se compose de sept plans où la présentatrice n'est plus visible, mais qui présentent des images filmées à l'Aquila sans que l'on sache de qui elles sont. La durée de cette partie est de 12,4 secondes, le générique du JT se fait toujours entendre. Le premier de ces plans montre un survol de la ville de l'Aquila et dure 1,35 secondes. Le second montre des personnes fouillant les décombres avec un mouvement latéral de la caméra qui peut contribuer à suggérer l'étendue des dégâts et dure 2,57 secondes. Le troisième plan montre une personne sur une civière, extraite des décombres, la caméra suit le mouvement de la civière, ce plan dure 1,47 secondes. Le quatrième plan montre un homme seul qui jette des gravats pour dégager un bâtiment, la

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

caméra suit son mouvement, il dure 1,59 secondes. Le cinquième plan filme un berger allemand évoluant sur les décombres, la caméra suit son mouvement latéral, le plan dure 1,47 secondes. Le sixième plan dure 2,16 secondes, il montre des secouristes vêtus d'uniformes s'affairant autour de gravats en plan fixe, des civils les regardent, ce plan dure 2,16 secondes. Enfin, le dernier plan montre une personne secourue qui est à terre, sur une civière avec des civils qui lui portent secours en plan fixe, il dure 1,76 secondes. Tous les enchaînements entre ces plans se font par des cuts, c'est à dire que le changement d'une image à l'autre est brusque et rapide.

Cette partie se compose ainsi de plans très courts : aucun n'excède les 3 seconde, leur moyenne est d'environ 1,76 secondes. Ils montrent diverses images qui évoquent d'une part la gravité des faits : les ruines sont présentes à chaque instant, leur étendue étant soulignée par la vue aérienne de l'Aquila ou bien par les différents mouvements des caméras. On montre aussi le fait que de nombreux civils soient les secouristes, des personnes en uniforme n'intervenant pas avant le sixième plan. De nombreux blessés sont montrés sur des civières, ainsi qu'un chien secouriste. Les mouvements de caméras, qui peuvent donner du dynamisme et mieux capter le regard du téléspectateur sont le plus souvent liés à un mouvement du sujet filmé (sauf lors du premier plan).

Cette partie montre de façon concise de quoi il est question : séisme, blessés, secours d'urgence, mais peu de secours gouvernementaux, apparemment.

Pendant toute cette seconde partie, un encart apparaît au bas de l'écran avec les mots écrits : « *séisme italien. PRES DE 150 MORTS. 19/20 National. Les titres.* », ce qui permet à tout téléspectateur pouvant lire de rejoindre l'acte de communication en cours de route et de comprendre de quoi il retourne et où il en est dans le JT. Voici deux images de cette partie issues des premier et second plans.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?



La journaliste prononce les énoncés suivants dans un écrit oralisé :

« Bonsoir, dans cette édition : le séisme qui a frappé la région d'Aquila dans les Abruzzes en Italie. Un bilan provisoire fait état de cent cinquante morts, mille cinq cents blessés et cinquante mille sans abris. Pour l'instant, l'Italie dit ne pas avoir besoin de l'aide internationale. Nous verrons l'état de sidération de la région touchée, et entendrons les inquiétudes de la population à chaque réplique. »

Les marques de subjectivité de personne sont le « nous » et les terminaisons en «-ons », le deuxième personnage agissant évoqué est « l'Italie » qui ne requiert pas d'aide. Cette action de l'Italie semble ici mise en contradiction avec d'une part la graduation effectuée dans l'énonciation du bilan, cette graduation est aussi bien effective en terme de chiffres que de syllabes pour chacun de ses trois termes ; d'autre part, avec les mots « sidération » et « inquiétudes » intervenant juste après le discours rapporté de l'Italie qui semblent décrire une situation très grave.

Le mot réplique peut induire une certaine confusion dans l'esprit du téléspectateur, sa polysémie faisant qu'il peut s'agir d'une réplique du tremblement de terre, et donc d'une secousse répétée, ou bien de l'énoncé de personnes interviewées, au vu du contexte, le premier sens est le plus probable, néanmoins, rien ne vient évacuer de façon claire le second sens qui est plus commun dans la langue française. Le montage en parataxe de la première phrase donne une impression d'objectivité, en enlevant tout verbe, on enlève la personne qui active ce verbe, il permet également d'en arriver de façon concise aux faits.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

Ces titres de France 3 produisent ainsi, au niveau visuel, une certaine impression de sobriété et de dynamisme qui opère, l'accent est porté sur le côté sérieux et non ostentatoire de l'information à travers différentes techniques. Le dynamisme est porté par la rapidité de l'enchaînement des différents plans et les mouvements permanents de caméra.

Cependant, la subjectivité apparaît dans le choix des plans de la seconde partie de ces titres : beaucoup de secouristes civils et non équipés, des dégâts immenses. La partie linguistique renforce cette idée d'une objectivité qui diminue peu à peu, à une phrase nominale montée en parataxe, succède une graduation éloquente sur la gravité de l'événement, juxtaposée au discours rapporté de l'Italie de façon métonymique (désignant le gouvernement), et aux adjectifs décrivant l'état de la population.

VI.2.1.2- TFI.

Le premier plan dure 5,36 secondes, il correspond structurellement à la première partie des titres de France 3. La présentatrice parle, la caméra garde un plan fixe sur elle.

La présentatrice est présente en plan rapproché, l'image comporte deux dimensions, un arrière plan composé d'un mur d'écrans symbolisant l'ouverture sur le monde, avec une vue de plusieurs endroits à la fois permettant une forme d'omniscience du média et lui conférant en quelque sorte un don d'ubiquité. Laurence Ferrari est vêtue d'un gilet gris, très sobre, son maquillage est peu présent, et ses boucles d'oreilles sont également sobres. Elle ne quitte pas la caméra du regard. On ne voit pas ses mains, et la présence de notes écrites ou non n'est donc pas envisagée dans ces titres. La date de jour apparaît en bas de l'écran. Les couleurs dominantes du plateau sont le bleu et le rouge : couleurs nationales françaises, ainsi que celles de la chaîne. Le bleu domine, c'est la couleur qui passe le mieux à la télévision, au niveau technique.

Une très forte impression de sobriété se dégage de ce premier plan, le plateau est très simple, en deux dimensions, le sujet principal y est la présentatrice, elle est centrée, et filmée en plan rapproché, cependant, sa mise demeure très sobre.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?



Le deuxième plan est enchaîné avec une arrivée par la droite de l'image. Ce plan dure 2,04 secondes, il montre une vue aérienne de la ville d'Aquila. L'image bouge légèrement. Le troisième plan dure 2,96 secondes, il montre également une vue aérienne de l'Aquila différente, cette fois, la caméra amorce un léger zoom vers les décombres. Le quatrième plan dure 1,96 secondes, il montre des hommes et des secouristes qui fouillent les décombres, à la recherche de survivants. La caméra reste statique. Le cinquième plan montre un homme effondré entouré de ses proches. Ce plan dure 3,64 secondes. C'est le plus long, donc celui sur lequel l'accent est porté parmi les plans composant ce titre. Le sixième plan dure 2,08 secondes, il montre une rue dévastée dans laquelle la caméra amorce un mouvement de zoom qui peut faire croire que l'on y avance. Des voitures détruites côtoient une chaussée pleine de gravats. Le septième plan dure 3,32 secondes, il montre un univers chaotique : la caméra bouge sans ordre, des secouristes en uniforme se fraient péniblement un chemin à travers des monticules de gravats, on ne distingue que les secouristes, les tas de ruines, et une forte lumière en arrière plan. Le huitième plan dure 2,35 secondes, il est une forme de continuation du précédent, les secouristes sont plus loin dans la rue, on aperçoit une voiture au bord d'un gouffre : la chaussée s'est affaissée à un endroit laissant apparaître un gouffre béant vers lequel la voiture penche dangereusement. Une femme est également au bord de ce gouffre, divers projecteurs forts apparaissent à l'arrière plan. Le neuvième plan dure 2,25 secondes, il montre une ambulance arrivant avec célérité vers la caméra et passant à côté, tous feux allumés et sirènes hurlantes. Le plan est fixe. Le dixième plan dure 0,44 secondes, il montre un tas de gravats de plusieurs mètres de haut, et un homme vêtu de noir qui le contemple en portant sa main à la tête. Le onzième plan dure 1,28 secondes, il montre un homme en noir déambulant dans une rue pleine de ruines. Chaque plan comporte un

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

bandeau écrit dans le bas avec la mention du titre dont il est question, du logo de la chaîne et de la date du jour. Voici une image extraite du huitième plan :



Tous les enchaînements de plans dans cette partie se font par des fondus enchaînés rapides. La durée moyenne des plans est de 2,48 secondes, ils proposent une vision de l'ampleur de la catastrophe qui a eu lieu : l'étendue des dégâts est filmée par hélicoptère, de nombreuses victimes sont montrées, des images fortes comme celle d'un scooter tombé dans la chaussée affaissée sont diffusées. Tout concoure ici à accentuer la gravité de ce qui s'est produit en Italie.

Au niveau linguistique, la présentatrice prononce les énoncés suivants :

« Bonsoir à tous et bienvenue dans le journal de TF1. Dans l'actualité ce soir : un violent tremblement de terre a ravagé le centre de l'Italie très tôt ce matin. Mille cinq cents personnes sont blessées et au moins cent cinquante ont trouvé la mort. Il s'agit du pire séisme depuis trente ans dans la région. L'état d'urgence nationale a été décrété par Silvio Berlusconi. On estime que au moins cinquante mille personnes sont sans abri ce soir, alors que la terre continue de trembler. Page spéciale dans ce journal avec nos envoyés spéciaux sur place. »

Les marques de subjectivité sont présentes à travers différentes personnes, le « nous », « nos » et les terminaisons en -ons assimilent la présentatrice avec un groupe. Le « on » est présent, ce qui confère une certaine imprécision à l'information, mais également une forme d'objectivité. Deux noms sont présents : TF1 et Silvio Berlusconi, le premier insiste sur le caractère intrinsèque de la chaîne qui donne l'information et situe le

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

télespectateur. Le second montre un homme qui agit de lui même, ici ses actions sont rapportées à la voie passive. La terre est aussi personnifiée, elle est le sujet du verbe d'action trembler. Le téléspectateur est présent via l'adjectif « tous ».

De nombreux modalisateurs sont présents, l'adverbe « très », l'adjectif qualificatif « violent », le superlatif « pire », ces différents modalisateurs accentuent la gravité des faits qui sont relatés. De même, dans la troisième phrase, on commence par parler de blessés avant d'en arriver à des morts, sorte de graduation macabre pouvant impliquer implicitement un lien de cause à conséquence. Le réseau lexical du séisme est présent, situant rapidement l'information. « tremblement de terre », « séisme », « trembler ». Les mots sont assez simples pour que tout le monde puisse comprendre sans difficulté. Le réseau lexical de la catastrophe est présent à travers les mots « blessés », « mort », « urgence » « ravagé ». La présentatrice reprend également son souffle pour accentuer les mots « du pire séisme », ce qui tend à accentuer la gravité de la situation décrite.

VI.2.1.3- Comparaison.

Ces deux titres proposent une structure relativement semblable. Dans les deux cas, une première partie intervient avec la présentatrice, puis une deuxième partie qui laisse la place à des images filmées sur les lieux de l'événement. Ces images sont parfois communes (la première vue en hélicoptère par exemple), ce qui exclut le fait qu'elles aient été tournées par les journalistes de ces chaînes, leur provenance n'est néanmoins pas explicitée.

Cependant, ces deux titres se distinguent en plusieurs points. Ce sont les premières parties qui diffèrent le plus, à la vue d'ensemble du plateau du JT de France 3 s'oppose la vue en deux dimensions de celui de TF1. L'impression qui peut se dégager de cela est un plus grand dynamisme du côté de France 3, mais aussi une plus forte sobriété pour TF1. Les mises des deux présentatrices sont relativement sobres, celle de Laurence Ferrari l'étant davantage que celle de Audrey Pulvar, moins classique. Les plans choisis pour filmer les deux présentatrices accentuent fortement son importance dans le journal de TF1.

Les plans choisis dans la seconde partie, s'ils semblent provenir du même type de source, montrent également certaines différences, leur durée moyenne est beaucoup plus

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

longue sur TF1 que sur France 3, ce qui peut donner à nouveau un effet de dynamisme pour France 3 contre une certaine sobriété, une profondeur plus importante du traitement du sujet pour TF1. D'autre part, les choix des plans ne suggèrent pas tout à fait le même événement : France 3 accentue sa présentation sur le fait que les civils se débattent dans les ruines pour sauver leurs proches, tandis que TF1 dépeint une situation dramatique, catastrophique, avec des secouristes qui aident les populations.

Au niveau linguistique, cette impression d'une tragédie plus grave dépeinte sur TF1 se confirme, les mots sont plus forts, de nombreux modalisateurs viennent les accentuer, ainsi que l'intonation de la présentatrice, le vocable est relativement aisé à comprendre de tous. Sur France 3, la subjectivité semble peu à peu prendre davantage de place dans les propos de la présentatrice, avec une importance plus accrue à la réaction du gouvernement italien à cette catastrophe.

VI.2.2- Le premier reportage.

VI.2.2.1- FR3.

VI.2.2.1.1- L'annonce par la présentatrice.

Celle-ci elle est filmée en plan rapproché. On ne voit pas si elle est assise ou non. L'image a deux dimensions, en arrière plan, on voit la sorte de verrière et un des deux grands écrans sur lequel l'image d'un gouffre dans une rue est présente : au bord de ce gouffre se trouvent une personne et une voiture. Entre la présentatrice et l'écran, on aperçoit deux micros sur un bureau avec des feuilles et un stylo. La présentatrice a une prononciation neutre, elle accentue néanmoins « cinquante mille personnes », son regard ne quitte pas la caméra. Le bandeau « 19/20 national » est toujours présent, il le reste tout au long du JT. Le nom de la présentatrice apparaît en grand format au dessus, puis disparaît au bout de quelques secondes.

Ce premier plan est introduit par un fondu enchaîné. Il dure 16,08 secondes et se termine également par un fondu enchaîné. Un deuxième plan clôt cette annonce par la présentatrice, il dure 5,18 secondes. Il s'agit d'une carte animée de l'Italie. Les couleurs

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

utilisées sont le blanc pour l'Italie, le bleu clair pour les pays alentours, et un bleu plus soutenu pour la mer. L'Aquila est située par une bannière orange, le code couleur du plateau et de l'ensemble du JT est ainsi rappelé. Des ondes concentriques se développant à partir de l'Aquila peuvent rappeler le sujet sismique du reportage.

Au niveau linguistique, la présentatrice prononce les énoncés suivants :

« Le bilan est encore provisoire, mais on parle déjà d'une centaine de morts dans le centre de l'Italie touché par un séisme de magnitude six virgule trois sur l'échelle de Richter. Une vingtaine de villes et de villages ont été partiellement ou totalement détruits. Cinquante mille personnes, au moins, sont considérées comme sans abris. Stéphanie Pérez. »

Le ton est toujours relativement neutre. Une seule personne est présente, « on », celle-ci permet une certaine objectivité du ton, au détriment d'une précision de l'information. Les modalisateurs « partiellement » et « totalement » accolés rendent cette même impression d'imprécision puisqu'ils évoquent des idées contraires. Les mots « séisme », « magnitude », « Richter », renvoient à une conception scientifique de l'événement et requièrent une compétence minimale dans ce domaine de la part du téléspectateur, ces informations permettent de situer précisément l'importance du séisme qui s'est produit. Les autres chiffres, loin d'utiliser la précision d'une décimale sont exprimés par « centaine », « vingtaine », ce qui est beaucoup plus vague.

Le chiffre cinquante mille est nuancé par la locution « au moins » qui tend à en accroître l'importance, on sous entend que la réalité est probablement plus importante. Le nom de la journaliste faisant le reportage est annoncé sans aucune transition. La seule indication spatio temporelle concerne le lieu où s'est déroulé l'événement, à savoir le « centre de l'Italie ».

VI.2.2.1.2- Le reportage.

Une description de ce reportage est présente en annexe 12. Il se compose de 32 plans, ceux-ci ayant une durée moyenne de 3,87 secondes pour une durée totale de 2,3 minutes. Cette durée moyenne des plans est plus élevée que lors des titres, cela s'explique par la volonté d'aller davantage en profondeur dans un reportage et le fait de procéder à certains interviews de témoins qui donnent lieu à des plans longs.

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

12 plans montrent des images de secouristes autour des gravats, cherchant à extraire d'éventuelles victimes, 10 plans montrent des victimes seules, 7 plans montrent des gravats sans être humain et 3 plans montrent des civils officiant en tant que secouristes à main nue ou avec des pelles. Ces choix opérés, quant aux images à montrer, ont pour effet de donner l'image d'une catastrophe pour laquelle les secours interviennent, mais qui fait surtout de nombreux dégâts matériels et humains. Les plans supérieurs à la durée moyenne montrent des victimes sauvées par les secouristes, des interviews de secouristes, l'étendue des dégâts en filmant avec des plans panoramiques, des victimes seules ou des victimes secourues par des civils. Le plan introductif montre un homme en sous-vêtements qui est extrait des décombres où il était pris au piège, ce plan dure 10,45 secondes, c'est le plus long du reportage, sa charge émotionnelle est très forte, l'homme éclate en pleurs. Le dernier plan montre l'image d'une victime qui prend pour seul soutien une représentation du Christ. Cette dernière image contribue à donner de l'Italie la vision stéréotypée d'un pays très fortement catholique.

Au niveau linguistique, (le commentaire apparaît en annexe 12 en italiques dans le tableau), différents réseaux sémantiques apparaissent. Tout d'abord, certains mots assez précis situent bien le phénomène du tremblement de terre : « épïcéntré », « séisme », « répliques ». Ils appartiennent de façon restreinte à ce domaine, et du fait de cette relative scientificité de leur usage, peuvent ne pas être compris de tous. D'autre part, un réseau sémantique de la catastrophe est mis en place à travers les noms « catastrophe », « dégâts » et « violence », les adjectifs « terrorisés », « survivants », « prostrés », « hagards », « ravagée », « dangereux », les verbes « acharnent » et « redoutent ». La locution « au moins » devant le nombre important « dix mille » renforce cette idée d'une terrible catastrophe qui s'est produite. La métaphore « tas de gravats qu'est devenue leur ville » vient confirmer cet effet de dramatisation forte de la situation dépeinte. Le domaine du religieux est aussi présent avec l'adjectif « miraculés », la locution « fin du monde » ou encore la paraphrase « toutes les protections possibles » qui sous-entend la protection divine puisqu'à ce moment du commentaire, un homme sous un crucifix apparaît. Deux explications peuvent se trouver pour cette évocation du religieux : d'une part, on peut considérer que la situation est tellement hors norme que seule une aide surnaturelle pourrait y remédier, d'autre part, l'Italie serait avant tout un pays de dévots très catholiques qui remettent leur salut entre les

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

mains de Dieu. L'expression « a reconnu ce matin le chef du gouvernement, Silvio Berlusconi » est particulièrement intéressante, en effet, le verbe « reconnaître » est généralement suivi d'une expression peu valorisante : on reconnaît un méfait, une faute, etc. Si ce que le gouvernement reconnaît ici est relativement neutre, on peut donc penser qu'il s'était fourvoyé dans un premier temps, reconnaissant par la suite son erreur. Cette induction est implicite, mais fortement présente par l'usage des mots « a reconnu ».

Trois idées essentielles apparaissent alors à la suite de l'analyse de ce reportage : on veut avant tout y montrer une réalité dont la gravité est importante et un pays qui invoque une aide divine pour sortir de cette impasse. D'autre part, implicitement, le gouvernement est mis en faute à travers le commentaire.

VI.2.2.2- TFI.

VI.2.2.2.1- Le lancement du sujet.

Ce reportage est synthétisé dans le tableau de l'annexe 13. L'introduction par la présentatrice se fait également en deux plans, un premier plan long de 20 secondes cadre celle-ci en plan rapproché avec un zoom au début du plan, puis un second d'une durée de 6,72 secondes donne à voir une image animée afin de mieux situer les faits au niveau géographique. Lors du premier plan, l'attention est très clairement portée sur la présentatrice. Le plan est introduit par un zoom, et elle est filmée en plan rapproché, sa position étant centrale. Sa diction est relativement fluide, l'accentuation qu'elle porte sur certains mots est ainsi plus notable, et permet une importante emphase. Ici, ce procédé est utilisé pour les mots « pleine nuit », « secousse », « supérieure » et « tout », insistant sur le côté surprenant (au milieu de la nuit) de l'événement et sur son importance (ressenti dans « presque tout le pays, même jusqu'à Rome »). Chacun de ces mots est de plus appuyé par un mouvement de tête et un haussement de sourcils. La présentatrice affecte également différentes poses, cela peut avoir pour effet de donner davantage de dynamisme à un plan fixe de 20 secondes.

Le deuxième plan, qui met en place une image animée a une vocation de précision du cadre de l'événement rapporté. Une carte géographique présente l'Italie, un encart dans le

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

coin gauche situe l'Italie en Europe, et l'essentiel de la carte situe Rome, l'Aquila et la région des Abruzzes. Le séisme est figuré par des ondes concentriques rappelant les ondes sismiques. Le code couleur utilise du bleu pour la mer, du beige pour la terre, du noir ou blanc pour les toponymes et du rouge pour le lieu de la catastrophe.

Au niveau linguistique, le commentaire fait intervenir le vocable précis du tremblement de terre : « séisme », « magnitude », « six degrés », mais des mots plus simples de ce réseau sémantique sont également présents : « secousse », « tremblement de terre ». De cette façon, l'ensemble des téléspectateurs peut comprendre de quoi il est question. L'échelle sur laquelle la magnitude est calculée n'y est cependant pas précisée. Les mots « violent », « tout le pays », « alourdir » rendent également compte de la réalité catastrophique qui va être donnée à voir dans le reportage, pour aboutir à la fin de ce lancement à la graduation « au moins cent cinquante morts et mille cinq cents blessés ». Le nom du reporter est introduit par la phrase nominale « Récit de Michel Scott », donnant un peu plus de subjectivité que le simple nom donné sur France 3, quand bien même la phrase nominale demeure très neutre. Aucune personne n'est présente dans ce lancement, le cadre spatio-temporel y est néanmoins très précisément situé : « centre de l'Italie », « trois heures et demie du matin ».

Ce lancement, plaçant la présentatrice au cœur de la présentation de l'information donne donc de multiples informations sur l'événement qui s'est produit. Il est plutôt neutre, insistant néanmoins sur l'importance et la gravité de ce qui vient de se produire.

VI.2.2.2- Le reportage.

Ce reportage est réalisé par Michel Scott, un reporter international très présent dans le journal de TF1 depuis plusieurs années. Il se compose de 47 plans, d'une durée moyenne de 3,7 secondes. Ces plans montrent par ordre d'importance des secouristes portant de l'aide aux victimes du séisme, puis déblayant des gravats, ils montrent également des victimes seules, l'importance des dégâts causés par le tremblement de terre, des victimes secourues par des civils et des membres du gouvernement italien. L'accent est ainsi porté sur le travail important réalisé par les secouristes en premier lieu, puis ensuite sur les conséquences dramatiques pour la population du séisme. Les plans d'une durée supérieure à la moyenne de

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

3,7 secondes donnent à voir des secouristes qui interviennent dans un univers proche du chaos, mais également des victimes secourues par ces mêmes secouristes, des interviews de sinistrés et de secouristes, une vue des dégâts par hélicoptère et une conférence de presse donnée par le gouvernement italien. Ce dernier plan, d'une durée de 6,56 secondes est parmi les plus longs du reportage, on peut noter en arrière plan la présence du drapeau européen qui était associé au drapeau italien lors du plan précédent, cela confère alors à l'image et au propos qui y est développé une dimension plus officielle et sérieuse.

Le reportage débute sur un plan montrant une situation de fin du monde, un univers assez chaotique, tandis que le dernier plan montre une vue aérienne de l'étendue des dégâts.

Le commentaire situe dans un premier temps le cadre spatio temporel de l'événement avec les mots « Aquila » et « trois heures trente deux du matin ». La gravité de la situation est très fortement accentuée par un réseau sémantique qui quadrille tout ce commentaire « désastre », « abattu », « choc », « stupeur », « violence », « effondrent », « pris au piège », « terrible », « dramatique ». Les mots « détruit », « dévasté » et « mort » sont notamment répétés chacun à deux reprises. D'autre part, ces mots sont fréquemment dotés d'amplificateurs tels que « ampleur du », « sans précédent », « dizaines », « centaines », « au moins ». La situation donne même lieu à des descriptions qui la comparent à un phénomène hors du possible ou du réel : « fin du monde », « cauchemar ».

L'urgence de la situation et la rapidité du travail des secouristes viennent néanmoins nuancer ce tableau apocalyptique avec des locutions ou des mots comme « très vite », « immédiatement », « course contre la montre », et les propositions suivantes « les secouristes ont continué à chercher », « les camps de tentes qui ont commencé à se monter ». En outre, l'action du gouvernement qui est décrite vient renforcer cette idée d'une situation qui a déjà entamé son processus de remédiation, tout en lui accordant une grande gravité : « Silvio Berlusconi qui a décrété l'état d'urgence ». Il ressort de ce reportage une impression de très lourde gravité des faits, les sons présents en continu composés de pleurs, de cris, de sirènes renforcent d'autant plus cette impression, ainsi que le commentaire utilisé pour la décrire. Néanmoins, il est aussi noté que les secouristes et le gouvernement sont réactifs face à cette situation

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

VI.2.2.3- Comparaison.

Ces deux reportages présentent de nombreuses similitudes, des images provenant de sources communes et les mêmes interviews y étant notamment réalisées. Cependant, de nombreuses différences se font dans le choix et l'utilisation de ces sources communes. Au niveau de l'annonce du reportage, France 3 place sa présentatrice dans un environnement complexe, elle est filmée de façon moins proche que sur TF1 et semble être moins mise en valeur. Le vocabulaire utilisé dans cette introduction est scientifique sur France 3 tandis que TF1 utilise un mélange de vocabulaire scientifique et plus simple pour décrire l'événement. Le ton d'Audrey Pulvar est neutre quand Laurence Ferrari accentue fortement certains mots, accompagnant ces emphases de mouvements de tête. Il en ressort une impression générale d'une plus grande objectivité dans le lancement de France 3, néanmoins, il convient de noter que TF1 donne beaucoup plus de précisions concernant le cadre spatio-temporel de l'événement rapporté.

La durée moyenne d'un plan du reportage est plus courte sur TF1 que sur France 3, on peut considérer de cette façon que France 3 traite de façon moins superficielle les sujets qu'elle aborde, au détriment du dynamisme de l'information ainsi fournie. Cette chaîne présente trois aspects essentiels de la catastrophe : sa gravité, mais aussi la forte prégnance du catholicisme en Italie et les réactions peu efficaces du gouvernement. TF1, au contraire, insiste plus lourdement sur la gravité des faits, et surtout, montre le gouvernement et les secours officiels comme agissant avec efficacité.

VI.3- Mise en perspective à l'aide du corpus IntUne.

Cette étude montre que la télévision française tend à présenter un visage quelque peu stéréotypé de l'Italie, que ce soit en portant l'accent sur la désorganisation des secours ou sur le recours spontané des habitants de ce pays à une aide divine. Mais cette représentation se retrouve-t-elle dans l'ensemble du corpus IntUne, ou n'est-elle au contraire qu'un phénomène épisodique lié à la situation décrite dans le reportage ? D'autre part, cette représentation est-elle particulière à la télévision française, ou se retrouve-t-elle également dans la presse écrite ? Pour répondre à ces questions, nous allons tout d'abord étudier les

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

reportages traitant de l'Italie dans les journaux de TF1 en 2007 dans un premier temps. Puis dans un second temps, nous regarderons la représentation de différents mots clefs issus des reportages étudiés en 2009 dans les journaux de presse écrite.

VI.3.1- Les reportages de TF1 sur l'Italie en 2007.

TF1 dispose d'un reporter dévolu à l'Italie, Maurice Olivari, à la différence de France 3, cela montre d'emblée l'importance que la chaîne accorde à ce pays. C'est pourquoi nous allons travailler plus particulièrement les reportages effectués par ce journaliste, qui est présent en permanence dans le pays étudié. Six reportages sont effectués sur l'Italie seule dans les journaux de TF1 au sein du corpus IntUne en 2007. Cinq sont effectués par Maurice Olivari, et un dernier par Nicolas Escoulan. Les cinq reportages de Maurice Olivari traitent, dans un ordre chronologique, du carnaval de Venise, d'une déclaration du pape Benoît XVI, d'un nouveau système pour faciliter le recueil des bébés abandonnés, de l'arrestation de Cesare Battisti, et de la célébration du Jeudi saint. Le reportage de Nicolas Escoulan traite d'un transfert de peintures, dont un Picasso de l'Italie (Venise) vers l'Allemagne, ces peintures étant initialement en provenance de la France. Cinq des six reportages concernent donc des clichés généralement associés à l'Italie (fête, Église, archaïsme, art). L'analyse de la représentation de l'Italie dans ces cinq reportages nous permettra de mieux voir si ce pays a une représentation télévisée cohérente et marquée entre 2007 et 2009.

Les textes de ces reportages sont présents dans l'annexe 14. Nous allons maintenant nous attarder sur les constructions discursives que l'on y aperçoit pour rendre compte de l'Italie reportage par reportage. Nous mettrons en parallèle les premier et quatrième reportages qui se déroulent tous deux à Venise, et le second ainsi que le cinquième reportage qui traitent précisément de l'Église catholique. Le troisième reportage présente quant à lui un thème particulier.

Le premier reportage montre avant tout une Italie intemporelle, qui fait rêver, que l'on associe au voyage, à la fête et au plaisir. Ces dimensions se manifestent essentiellement dans le choix de la lexie, ainsi que dans l'utilisation d'hyperboles. Au niveau lexical, le domaine du rêve évoqué par la ville de Venise est présent à travers ses grands symboles stéréotypes : « Venise », « place Saint Marc », « palais » sont cités deux fois, et on relève

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

également « grand canal » en une occurrence. De plus, le côté intemporel du pays est souligné avec des formules comme « authentique », « historiques », « dix-septième et dix-huitième siècles », « pur dix-huitième », « il y a deux siècles et demi ». L'accent dans ce reportage est porté sur la légèreté, la fantaisie. Le thème de la fête, du plaisir est présent : « soirées » (deux fois), « soirée » (deux fois), « fête », « costumes de lumière », « décolletés », « sourires ». Le thème de l'esthétique et de la beauté est aussi présent : « beau », « style », « élégance », « élégante », « esthétique », « parures », « coiffures », « raffinée ». Enfin, des hyperboles viennent accentuer cette abondance de mots renvoyant le spectateur à des idées de légèreté de d'intemporalité du rêve, on retrouve notamment la construction « l'un des plus grands et des plus beaux palais de Venise » ou « six mois par an » à la fin du reportage, qui montre que la fête est alors démesurée dans sa beauté, mais qu'elle l'était également il y a quelques siècles, dans sa durée. L'idée principale qui se dégage de ce premier reportage, à travers l'emploi d'une multitude de mots et formules se référant tous aux mêmes thématiques, est que l'Italie semble être un pays de fête, d'intemporalité, de rêve, de beauté et de légèreté.

Le quatrième reportage, qui montre une certaine similitude avec le premier, puisqu'il traite d'art à Venise, vient renforcer cette idée d'un pays où tout est beauté et rêve. La Venise mythique est toujours présente, avec exactement les mêmes mots stéréotypes « grand canal » (qui revient à trois reprises), « Venise » (deux fois), « touristes », et « galions de pirates ». De plus, de nombreux adjectifs hyperboliques viennent renforcer le côté exceptionnel de la beauté qui y prend place : « exceptionnelle », « rares », « chefs d'œuvre », « œuvre unique », « œuvre irremplaçable ». Néanmoins, la modernité entre dans ce reportage via le rappel d'une réalité financière de plus en plus problématique « primes d'assurances exorbitantes », « centaines de millier d'euros », « deux cents caméras » rappellent que Venise est également une ville qui est ancrée dans le monde réel, quand bien même cet ancrage est à nouveau sujet à une description hyperbolique. Ce reportage donne alors à voir une Italie où tout n'est que beauté et art, néanmoins, cette Italie subit également l'inscription dans le monde réel. Il est à noter qu'il n'est pas réalisé par le correspondant en place en Italie, Maurice Olivari, qui propose le même jour un reportage sur l'Église catholique.

Les second et cinquième reportages traitent de l'Église catholique, et plus particulièrement des agissement du pape Benoît XVI. Le second reportage développe trois

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

thèmes autour d'une argumentation principale qui vise à contrer le discours du pape. Ces trois thèmes sont présents dans le lexique employé. On retrouve tout d'abord la rigidité qui s'exprime à travers les mots « refermer la porte », « rigide », « ferme » (deux fois), « conservateur », « pas négociables ». Le thème de l'interdit est aussi présent à travers des mots comme « rappel à l'ordre », « obligatoire », « interdit », « condamnation ». Enfin, un troisième thème, qui est celui de l'intrusion dans la vie de la cité est présent avec les formules « adresse avant tout aux hommes politiques tentés par certaines réformes » ou l'accumulation « avortement, l'euthanasie, et l'union homosexuelle ». Ces trois thématiques montrent alors l'image d'une Église rétrograde, ne discutant pas, repliée sur elle-même, mais qui veut cependant peser de tout son poids dans les décisions de la cité. Le journaliste met alors en place un système argumentatif qui vise à contrer la position décrite. D'une part, il met en relief le côté extrême de cette position, cela se fait à travers l'emploi d'un montage en parataxe des différents interdits ou de l'expression « troisième rafale » qui introduit une nouvelle suite d'interdits. D'autre part, il met en place une ironie en confrontant à la lexie d'extrême rigueur et d'austérité qui est développée par le pape Benoît XVI son propre discours rapporté avec les mots « maturité, joie et dévouement » ou encore en confrontant des relations de couple et de fratrie en reprenant, toujours au discours rapporté, l'expression : « à moins que les remariés ne s'engagent à vivre comme frère et sœur ». L'argumentation qui s'appuie sur la mise en relief d'un caractère extrême puis sur l'ironie se solde alors sur un argument cette fois plus rationnel. Le journaliste note en effet que le devoir du pape est de rappeler le dogme, mais que la plupart des interdits rapportés n'ont pas de rapport avec ce dogme, sauf dans le domaine de l'euthanasie : « *le pape ne fait que rappeler le dogme, sauf qu'ici le dogme, mis à part peut-être la question de l'euthanasie, n'a rien à voir à l'affaire* ». Ce second reportage montre donc une Italie où la présence du pape se fait fortement sentir, mais surtout, il donne à voir un pape dont les discours et décisions sont fortement critiqués, plus que de l'Église, c'est en effet le discours du pape qui est contré par le journaliste : le dogme catholique a de la valeur, c'est même sur ce dogme que se fonde le dernier argument, ce qui pose problème est alors l'interprétation qui est faite de ce dogme par le pape, évêque de Rome.

Le dernier reportage traite également de la religion en évoquant les cérémonies de la semaine sainte qui ont pris place à Rome. La relation forte entre l'Italie et l'Église y est

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

plus marquée que dans le second reportage, puisqu'il est rappelé à plusieurs reprises que ce sont des « prêtres de Rome » qui participent aux cérémonies. Le mot « Rome » est présent à trois reprises en incluant le lancement du reportage, sont également présents les noms « Venise » et surtout, le très symbolique « Colisée ». D'autre part, le lexique fait intervenir une pléthore de mots référant à la religion catholique « St Pierre », « pape », « Église », « Christ », « crucifixion », « disciples », « jeudi saint », « prêtres », « sacrement », « consacre », « huile », « baptêmes », « confirmations », « ordination », « extrême onction », « chemin de croix » sont autant d'expressions qui associent les concepts d'Italie et de religion catholique dans ce reportage. Une deuxième remarque peut être faite quant à ce reportage. Il compare les deux plus récents papes de l'Église catholique (Jean-Paul II et Benoît XVI), et cette comparaison se fait à la défaveur du second. En effet, il est précisé que le lavement des pieds de ce dernier est « très symbolique », tandis que Jean-Paul II pratiquait un lavement rigoureux, même complété par l'apposition d'un baiser sur les pieds : « *Autrefois, Jean-Paul deux apposait même un baiser sur le pied ainsi lavé, pour mieux souligner l'acte de modestie qui s'en dégage. Rien à voir en tout cas avec la cérémonie de ce matin* ». Un manque d'humilité et de modestie est alors implicitement reproché au pape Benoît XVI, alors que ces deux qualités sont considérés comme des piliers de la foi catholique.

Lorsque l'Italie est présentée en association avec la foi catholique différentes choses sont ainsi mises en relief. D'une part, l'accent est porté sur le côté fortement ancré voire parfois extrême de cette foi au sein du pays. D'autre part, l'association entre les deux entités est souvent forte : qu'il s'agisse d'une volonté de l'Église de participer à la vie de la cité, ou de son ancrage géographique concret dans cette région.

Enfin, dans les reportages de TF1 du corpus IntUne 2007, un sujet traite du recours à un ancien système : les boîtes à bébés. Ce reportage est retranscrit dans l'annexe 14. Il est d'autant plus intéressant qu'il semble traiter d'un sujet particulier en rapport aux quatre autres reportages qui faisaient intervenir des thèmes présents à plusieurs reprises. Il semble en fait agir comme une synthèse des reportages que nous venons d'analyser. En effet, la dimension intemporelle de l'Italie, voire archaïque y est soulignée via des mots comme « antique », « deux siècles » ou « immuable ». Le lien avec l'Église catholique est également présent avec le mot « couvent ». Cependant, c'est aussi une Italie qui est confrontée au monde moderne que l'on donne à voir dans ce reportage « moderne », « plein cœur de Rome », « reliées vingt

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

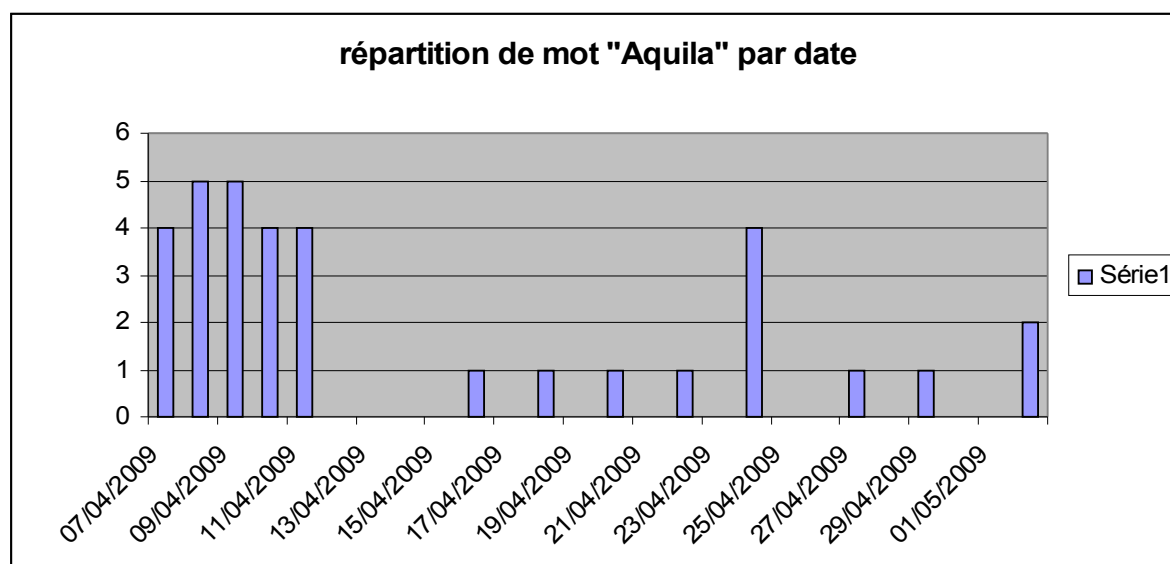
quatre heures sur vingt quatre par vidéo ». Bien que le sujet puisse prêter au pathos, les formules employées y confèrent une certaine légèreté. Le journaliste utilise notamment deux métaphores, la mère qui abandonne son enfant devient une « cigogne en perdition » et le nourrisson abandonné « un petit paquet », ces métaphores contribuent à alléger le sujet évoqué, mais également à en atténuer la réalité. Nous pouvons enfin noter que ce reportage fonctionne en partie comme un récit voire un conte de fée, inscrivant implicitement l'Italie dans un monde irréel et intemporel. Nous avons un héros, le docteur Stefano Vitali, qui est présenté dans une situation initiale : celui-ci a tout préparé pour que d'éventuels enfants abandonnés puissent être recueillis avec soin. Survient un élément modificateur : « *Il était ce soir là vingt et une heure trente lorsque sur l'écran de contrôle, le docteur Vitali a aperçu une paire de mains anonymes déposer délicatement un petit paquet bien enveloppé dans le berceau.* ». Le héros agit et propose une phase de résolution du problème ou élément modificateur : « *J'ai couru avec mes instruments de réanimation* ». La situation finale peut alors prendre place : « *L'enfant a été baptisé Stefano, du même prénom que son médecin favori, et il est en cours d'adoption.* ». Tout est bien qui finit bien dans ce conte italien.

Les cinq reportages que nous venons d'analyser montrent une très forte cohérence dans la représentation qu'ils donnent de l'Italie. Cette cohérence peut s'expliquer par le fait que c'est essentiellement le même journaliste qui les propose, Maurice Olivari. Cependant, même quand un autre journaliste de la presse télévisée française accorde tout un reportage à ce pays, l'image fait appel aux mêmes clichés. L'Italie est alors un pays irréel, intemporel, qui fait appel aux rêves de beauté et d'esthétisme. Mais c'est également un pays au passé très présent, qui montre parfois certains archaïsmes. Ce côté archaïque de l'Italie est d'autant plus présent qu'elle est fortement associée à l'Église catholique qui est décrite comme rétrograde dans les sujets qui y sont alloués. Cette représentation stéréotypée voire folklorique de l'Italie dans la presse télévisée apparaît de façon d'autant plus forte que ce n'est pas l'absence d'actualité qui donne ce traitement. En effet, lors de la collecte du premier corpus IntUne, l'Italie connaissait de nombreux revirements politiques, et avait une actualité étrangère chargée, que ce soit en rapport avec le Brésil ou l'Afghanistan, ces sujets n'ont cependant quasiment pas été traités par la télévision française dans le corpus IntUne. D'autres études menées sur le corpus IntUne ont par ailleurs montré une tendance des médias à représenter les différents pays de l'Union Européenne de façon stéréotypée (Williams 2008).

VI.3.2- La représentation de l'événement dans le corpus de presse écrite.

Maintenant que nous avons vu comment la télévision collectée dans le corpus IntUne représentait l'Italie en 2007, nous allons nous intéresser à une autre perspective pour voir si la représentation étudiée dans les deux reportages de TF1 et de FR3 du séisme de l'Aquila est cohérente ou non avec une représentation plus globale de l'Italie dans le corpus IntUne. Nous allons donc analyser un mot-clef de l'événement dans le corpus presse en 2009.

Ce mot clef est « Aquila », en effet, celui-ci intervient dans le corpus uniquement dans le contexte du séisme qui s'y est déroulé, ce qui n'est pas le cas des mots « séisme » ou « Abruzzes ». Celui-ci apparaît à 139 reprises dans le corpus, et il est présent dans 34 textes (un texte correspondant à un jour pour un journal). Il est exclusivement présent à partir de la date de la catastrophe, et sa répartition dans le corpus IntUne pourrait se résumer comme suit. On voit alors nettement que ce mot est particulièrement cité lors de la catastrophe qui s'est produite, et le 24/04/09, alors qu'une proposition est faite pour y organiser le prochain G8.



(Figure 27, répartition du mot « Aquila » dans les cinq journaux du corpus 2009 par date)

Ce mot apparaissant dans le cadre du séisme qui s'est produit dans cette ville, nous pouvons nous intéresser au traitement qui en est fait, au contexte dans lequel il apparaît. La figure 27 nous montre que ce mot apparaît pendant treize jours dans la presse, cela est sans

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

commune mesure avec sa présence dans les journaux télévisés qui s'est limitée à deux journaux télévisés du soir. La presse télévisée et la presse écrite ne proposent donc pas le même traitement de ce sujet en ce qui concerne l'importance qui y est accordée. Nous allons maintenant procéder à l'analyse des lignes de concordances dans lesquelles il apparaît.

A la lecture de ces lignes, il semble que les journaux associent la description de cet événement à différentes thématiques. Celle qui est de loin la plus présente est le côté dramatique de ce qui vient de se produire. Cela est d'autant plus marquant que même le journal Les Échos, peu enclin au sensationnalisme, montre cette dimension dramatique du sujet. Les morts sont nombreux, les villes détruites, les foules hagardes, et perdant même toute rationalité.

La secousse qui a détruit de nombreux bâtiments à	Aquila	(photo) a été ressentie jusqu'à Rome	ec070409
se pressaient hier encore autour des ruines à l'	Aquila	espérant contre toute logique qu'ils puissent	ec090409
agardes, sans un salut ni un sourire, les habitants de l'	Aquila	ont entrepris d'évacuer le centre ville.	fi070409
cercueils étaient alignés dans une caserne de l'	Aquila	.	fi110409
éplique d'une certaine intensité risque de mettre toute l'	Aquila	par terre. Regardez les maisons.	li070409
Dans la nuit de dimanche à lundi a ravagé l'	Aquila	et sa région, a fait plus de 150 morts,	li070409
néralles nationales des victimes auront lieu demain à l'	Aquila	, épiceutre de la catastrophe.	ou090409
	A l' Aquila	, ville la plus durement touchée par le séism	ou090409
colère de la terre a donc frappé la ville de l'	Aquila	et ses environs avec une intensité jamais	su080409
	L' Aquila	, capitale de la douleur d'un pays en deuil	su110409

Si le côté dramatique de l'événement qui s'est produit est accentué par tous les journaux, les deux nationaux politiques et les deux régionaux associent également cet événement à d'autres thématiques. Les problèmes de l'Église catholique sont alors abordés : les prêtres ou évêques doivent sacrifier certaines des multiples églises présentes dans cette zone, et touchées par le séisme. D'autre part, au niveau culturel, la perte d'un grand patrimoine est également déplorée par les journaux.

C'est la mort dans l'âme que l'évêque de l'	Aquila	, Mgr Giussepe Molinari, réfugié dans	fi080409
garde des finances à l'	Aquila	. L'homélie sera prononcée par le cardinal	fi100409
Maria des âmes saintes, sur la place centrale de l'	Aquila	s'est effondrée. Dans la maison à côté	li070409
Sur la place centrale de l'	Aquila	l'église Santa Maria des âmes sainte, déjà	li080409
bâtiments déjà endommagés de l'	Aquila	, dont la basilique	su080409
L'archevêque du diocèse de l'	Aquila	a annoncé hier que les manifestations religi	su080409
pour restaurer le patrimoine culturel. L'	Aquila	, capitale des Abruzzes, fondée au XIIIè siè	ou080409

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

A l'Aquila, la destruction du palazzo del governo, av li090409
ministre de la culture, Salvo Bondi, est attendu à l'Aquila pour un inventaire détaillé des dégâts fi080409

Cet événement semble donc faire surgir le même type de description dans la presse écrite que dans les journaux télévisés de l'Italie. Ce pays y est en effet fortement associé à l'Église catholique, et à son patrimoine culturel. Ces deux dimensions étant d'ailleurs reliées.

La mise en perspective de cette analyse de deux reportages de la presse télévisée avec l'ensemble du corpus IntUne semble donc montrer que lorsque l'on représente l'Italie, et d'autant plus lorsqu'on la représente dans les journaux télévisés, cette représentation est avant tout stéréotypée.

VI.4- Conclusion.

La présentation d'un même sujet par deux chaînes françaises différentes peut montrer deux visages présentant des divergences sensibles de l'Italie secouée par une catastrophe. Si ces deux chaînes utilisent les mêmes images pour présenter les faits, elles n'en font pas le même usage, et le commentaire diffère radicalement sur la présentation de la situation italienne. Le journal de France 3 porte une importance moindre à sa présentatrice que celui de TF1, les plans sur cette dernière n'y sont pas nécessairement rapprochés et elle apparaît en mouvement dans un environnement complet, multidimensionnel. La présentatrice du journal de TF1, quant à elle, est filmée en plan rapproché, de façon centrée dès qu'elle apparaît, ses intonations accentuent certains mots et son expressivité est plus importante que sur France 3, néanmoins, la sobriété de sa mise en scène minore cette présence importante.

Les reportages donnent à voir différentes facettes de la situation, tandis que TF1 accentue très fortement la gravité des faits qui se sont déroulés en Italie et la rapide réaction des gouvernants et des secouristes, France 3, bien que présentant également une réalité catastrophique, montre l'Italie encline à la croyance religieuse, et disposant de gouvernants dont les réactions ne sont pas abordées, mais faisant amende honorable de façon implicite sur une première réaction probablement peu pertinente. Le seul symbole européen présent dans ces deux reportages est le drapeau de l'union, affiché en arrière plan dans un des plans

6- La télévision française : quelle représentation d'un événement au retentissement européen ?

du reportage de TF1. Il apparaît donc que l'Italie, telle qu'elle est présentée dans ces journaux télévisés est éloignée de l'image de la France, elle peut notamment sembler désorganisée face à la catastrophe qui s'est produite. En outre, elle est montrée comme recourant à des solutions parfois obsolètes (prières). La description d'une catastrophe qui s'y produit donne alors lieu à montrer un stéréotype d'Italie.

7- Conclusion.

En conclusion, il serait intéressant de revenir sur le titre de ce travail, et sur les questions que celui-ci sous-tendait. L'intitulé de la démonstration proposait trois grands axes : la représentation et la construction d'une identité européenne, l'approche disciplinaire utilisée pour répondre à cette question, qui est la linguistique de corpus, et enfin, le fait que ce travail prenait place dans un projet européen. Le traitement de ces axes a été inversé par rapport à leur présentation dans le titre de la thèse, et cela, pour une meilleure cohérence de la démonstration.

Concernant le premier axe, une réflexion a été menée dans le premier chapitre. Les corpus sont utilisés depuis très longtemps dans la réflexion sur le langage, ils ont montré leur utilité dans des domaines aussi divers que la lexicographie (avec le dictionnaire anglais du docteur Johnson, par exemple) ou l'enseignement d'une langue étrangère. Néanmoins, un phénomène nouveau est apparu, qui a donné naissance à la discipline linguistique de corpus. Il s'agit de l'avènement de nouvelles technologies qui ont permis de créer des corpus de taille de plus en plus importante (Bank of English est un corpus d'environ 450 millions de mots) et surtout de traiter ces corpus de façon automatique. Ignorer les possibilités induites par de telles avancées technologiques dans le domaine de la linguistique reviendrait en quelque sorte à penser que l'invention du microscope électronique n'a en rien modifié la pratique des chercheurs en biologie. Ces changements majeurs ont introduit une nouvelle discipline, la linguistique de corpus.

Avant l'apparition des outils informatique, la linguistique a utilisé les corpus, mais elle s'est aussi énormément basée sur l'introspection du linguiste. C'est ce qui a donné

7- Conclusion.

naissance à la grammaire, des premières réflexions d'Aristote dans la *Poétique* jusqu'aux développements de notre siècle. La grammaire, telle qu'on l'entend traditionnellement se base sur l'introspection du grammairien, sur son expérience langagière de locuteur natif. Elle a un but prescriptif, normatif, c'est-à-dire qu'elle montre la façon dont la langue doit être utilisée, le bon usage que l'on doit en faire. Cette discipline est ainsi quelque peu éloignée du premier but de la linguistique qui est d'observer et de décrire la langue telle qu'elle fonctionne effectivement. A partir de ces observations, il s'agit d'identifier des régularités dans la langue et de les expliquer.

Saussure a donné ce but à la linguistique : trouver comment fonctionne le système langagier à travers le repérage de certaines unités régulières, mené de façon scientifique. L'interprétation du *Cours de Linguistique Générale*, reconstruction posthume de sa pensée, a mené à une linguistique de plus en plus introspective et détachée des faits de langue authentiques. Néanmoins, certains linguistes ont continué à travailler sur des corpus, notamment dans le domaine de l'enseignement des langues étrangères en Angleterre. Le développement des nouvelles technologies ainsi que le constat des échecs auxquels l'introspection pure semblait mener ont conduit à une synthèse se fondant à la fois dans les écrits de Saussure ou des premiers structuralistes, et dans l'utilisation de corpus.

Si les corpus permettent d'appréhender davantage la réalité, il est nécessaire de rester vigilant par rapport à leurs limites. Le corpus IntUne a été collecté selon différents critères expliqués dans le second chapitre, sa réalisation a été soumise à un certain nombre d'impératifs. Beaucoup étaient d'ordre scientifique : pour une meilleure comparabilité, on a ainsi exclu des journaux pourtant signifiants dans les différents pays de collecte. D'autres étaient d'ordre plus pratique : le Télégramme a été exclu pour des raisons de coûts du corpus français, au profit de Ouest France. Ces différentes contraintes ont permis d'une part d'obtenir un corpus comparable, comme nous l'avons vu, dans une certaine mesure. Elles ont néanmoins également conduit à appauvrir parfois certaines parties du corpus, que ce soit au niveau des textes collectés, ou parfois même des balises utilisées. Pour que le corpus soit comparable, les informations concernant les locuteurs du corpus oral ont ainsi été retranscrites sous forme de catégories ne pouvant rendre compte de toute la diversité des informations initialement disponibles lors de la diffusion du texte. Ces catégories ont évidemment été élaborées de façon collégiale, et permettent de rendre compte de points

7- Conclusion.

importants par rapport aux buts de recherche du projet IntUne, mais il est important d'avoir à l'esprit les limites qu'elles ont. Les différentes étapes selon lesquelles le corpus IntUne s'est constitué, et les difficultés ou réussites qui ont eu lieu lors de ce travail sont mises en relief dans cette thèse. En cela, un des buts du projet IntUne a été atteint, des techniques nouvelles (création et interrogation de corpus) ont été employées, et de nouvelles données ont été collectées (un corpus diachronique et multilingue regroupant des textes issus des médias).

Le corpus permet d'approcher davantage la réalité, c'est aussi le cas, pour ce qui est de la signification des concepts, des prototypes sémantiques. Nous avons ainsi cherché à montrer sur différents exemples français comment ces prototypes pouvaient se construire, et comment ils pouvaient par la suite être utilisés. L'étude ici développée les a utilisés comme aunes auxquelles mesurer les différentes représentations réelles des concepts dont ils rendaient compte. En élaborant un prototype sémantique qui rendait compte de la diversité des sens qu'un même mot complexe pouvait recouvrir, il a été possible de vérifier de façon structurée et signifiante comment la presse française actualisait et construisait le sens des concepts étudiés. Cependant, les prototypes sémantiques, tels qu'ils ont été décrits dans cette thèse, peuvent avoir de nombreuses autres applications.

Les prototypes se constituent de plusieurs sens potentiels qui sont des énoncés simples et traduisibles dans plusieurs langues. Ils peuvent donc présenter des avantages indéniables dans le domaine de la traduction. A partir du moment où les sens potentiels peuvent être dégagés et traduits dans plusieurs langues, traduire le sens de tel ou tel mot d'une langue A vers une langue B peut se faire si de nombreux prototypes sémantiques sont disponibles dans un domaine. Au niveau de la pédagogie, et plus particulièrement de l'enseignement des langues étrangères, les prototypes sémantiques sont également pourvus d'intérêt. Lorsqu'un apprenant est confronté à un mot dont il ignore le sens, il peut se tourner vers une définition classique dans laquelle il trouvera plusieurs sens bien distincts pour interpréter le mot rencontré. Il peut également faire appel à un prototype sémantique qui aura été construit à partir d'une grande matière lexicographique. Mais surtout, ce modèle permettra de rendre compte de l'actualisation réelle dans la langue de tous les jours du mot inconnu. La vérification de cette actualisation ayant notamment été rendue possible grâce à l'utilisation de grands corpus de langue.

Si l'utilisation des prototypes sémantiques, combinée à l'analyse de grands corpus

peut présenter de nombreux avantages dans les domaines de l'apprentissage des langues étrangères ou de la traduction, ils ont essentiellement été utilisés pour mieux rendre compte de la façon dont la presse contribuait à l'émergence d'une ou plusieurs identités en Europe.

Une des questions de base du projet IntUne est de savoir comment l'identité européenne se construit, comme elle émerge, et comment différentes identités peuvent-elles coexister au sein d'un système complexe, l'Europe. A travers l'analyse menée dans cette recherche, différentes réponses peuvent être apportées à ces questions. D'une part, les représentations de l'ensemble dans les médias français peuvent être favorables à l'entité interrogée, et contribuer à la construction d'un sentiment d'une identité européenne. Ainsi, quand l'Europe est présente à travers son nom, elle montre des signes encourageants d'une certaine proximité entre les différents états membres. Ceux-ci sont capables de s'entendre pour organiser des manifestations sportives bien représentées au niveau médiatique. Ils sont fréquemment comparés les uns aux autres, présentant de fait de nombreux points communs, tant au niveau de leur économie que de leur situation sociale. Leurs cultures sont riches et diverses, mais soit elles partagent des racines communes, soit leurs différences sont une source permanente d'enrichissement et d'inspiration pour les artistes européens. Quand bien même cette entité est parfois présentée comme vieillissante, ou en difficulté d'un point de vue économique, la proximité des états membres, leurs nombreux points communs sont évoqués.

Nonobstant, sous cette dénomination « Europe », une réalité sensiblement plus problématique est également dépeinte. Il s'agit d'un ensemble qui vit une grave panne institutionnelle, avec un grave déficit de légitimité : l'organe autocratique de pouvoir de l'entité, la Commission européenne semble agir tel un despote qui impose selon son bon vouloir des règles à tous ses États membres quel que soit le domaine concerné : pêche, agriculture, téléphonie, etc. En outre, elle peut avantager et handicaper ces états membres, la presse française montre ainsi le plus souvent une commission européenne favorisant d'autres pays au détriment de la France (aides aux pêcheurs espagnols, aides aux agriculteurs polonais, aides aux pays de l'est, amende record pour l'agriculture française, etc.). Ce portrait moins flatteur est d'autant plus présent qu'il recouvre une vaste majorité des occurrences du corpus.

En fait, plus que d'une Europe, l'étude de ce corpus nous amènerait plutôt à parler

de plusieurs « Europes » dont la géométrie variable n'aurait d'égal que la multiplicité. Des dénominations basées sur des critères géographiques apparaissent en effet avec une grande fréquence à la suite du mot « Europe » lorsque l'on souhaite préciser l'entité dont on parle. A une Europe du Nord viendra s'opposer une Europe du Sud, à une Europe de l'Ouest son pendant symétrique de l'Est ; s'opposent également occidentale et orientale, tandis que l'adjectif centrale est de plus en plus fréquemment utilisé pour désigner une entité qui se situerait à cheval entre Est et Ouest. Lorsque ces différentes « Europes » sont présentes, on note alors une concurrence certaine entre elles, tandis que le sentiment de proximité précédemment évoqué s'atténue sensiblement. D'autre part, la présence de cette rivalité, du développement ou de l'aide d'une Europe au détriment d'une autre prend de plus en plus d'importance au fil du temps, notamment du fait de la crise économique secouant la zone Europe en 2009.

L'étude de deux pays appartenant à deux de ces « Europes », l'Italie et la Pologne permet de voir comment les médias français représentent nos voisins européens. Cette représentation, si elle montre une certaine proximité culturelle ou économique et sociale entre la France et l'Italie, en dresse néanmoins un portrait souvent caricatural (Williams 2008), ne cherchant pas à aller dans la profondeur des réalités sociologiques de ce pays, et préférant souvent présenter un biais défavorable des italiens. Concernant la Pologne, l'image qui en est donnée en 2007 est résolument caricaturale, à travers un gouvernement « ubuesque », et des salariés migrants en nombre suffisant pour investir l'ensemble des pays d'Europe de l'Ouest. En 2009, sa dimension politique semble s'être améliorée, mais au détriment de la composante économique. La Pologne est alors un pays qui demande toujours plus de solidarité de la part de ses voisins européens quand ceux-ci connaissent déjà de nombreuses difficultés. Ainsi, l'Europe et ses différents États membres ne semblent-ils pas complètement intégrés et unis à travers la représentation qu'en donne la presse française. Si certains traits caractéristiques sont partagés, si l'on reconnaît le besoin d'Europe de ces différents pays, lorsque celle-ci est décrite à travers des comportements concrets et non plus sous la forme d'une idéalisation, ses actions sont le plus souvent négatives pour les citoyens français auxquels elles sont décrites. Le fonctionnement institutionnel semble poser un problème certain de légitimité, les lecteurs français de presse ne se reconnaissant plus dans un organe de pouvoir autocratique, contre lequel leurs propres gouvernants nationaux disent négocier.

7- Conclusion.

Ces différents constats suite au travail d'analyse et de recherche effectués permettent de répondre à l'un des objectifs du projet IntUne. Une base de connaissances a été élaborée sur la façon dont l'identité européenne se comportait dans les médias français d'un point de vue scientifique. Les résultats de ces recherches peuvent être utilisés par un large public, allant des masses aux acteurs clefs ou décideurs politiques.

Cette thèse répond donc aux objectifs fixés dans le cadre du projet IntUne. Mais elle répond également à des applications plus larges, comme nous l'avons vu précédemment, à travers l'emploi des méthodologies ici expérimentées dans le domaine de l'apprentissage des langues étrangères ou de la traduction. Notre tâche est de poursuivre conjointement l'expérimentation de ces méthodologies dans différents domaines, ainsi que leur application, notamment dans l'enseignement des sciences du langage.

Nous avons déjà expérimenté ce type d'applications lors d'un cours portant sur l'exploitation de ressources audiovisuelles qui s'est fondé sur entre autres sur l'utilisation du corpus télévisé IntUne. Les étudiants ont alors pu utiliser cet outil pour voir ce que signifiait un thème de leur choix selon les médias français. La mise en place d'un cours de français langue étrangère (FLE) faisant un grand emploi des corpus est également envisagée à court terme. D'une part, la réalisation d'exercices basés sur des lignes de concordances, ou la mise en situation pour des étudiants en FLE peut être réalisée. Cela peut passer par le choix de différents mots qui ont une forte proximité de sens dans différents énoncés, le choix de ces mots pouvant se faire à l'aide de lignes de concordance en format KWIC du mot étudié. D'autre part, la linguistique de corpus peut être utilisée pour réaliser des dictionnaires pour apprenant avancés. Ces dictionnaires peuvent se fonder sur les prototypes sémantiques évoqués plus haut, et présente surtout l'avantage de montrer des usages réels et authentiques de la langue. En outre, ces dictionnaires peuvent se fonder sur des corpus de spécialités, et concerner des apprenants qui ont des besoins spécifiques. C'est le cas de nombre d'étudiants dans diverses disciplines scientifiques, les dictionnaires qui leur sont proposés pour l'instant ne proposent que trop peu souvent des définitions qui correspondent à leur champ d'étude ou qui soient suffisamment étayées par rapport à leur niveau de connaissance.

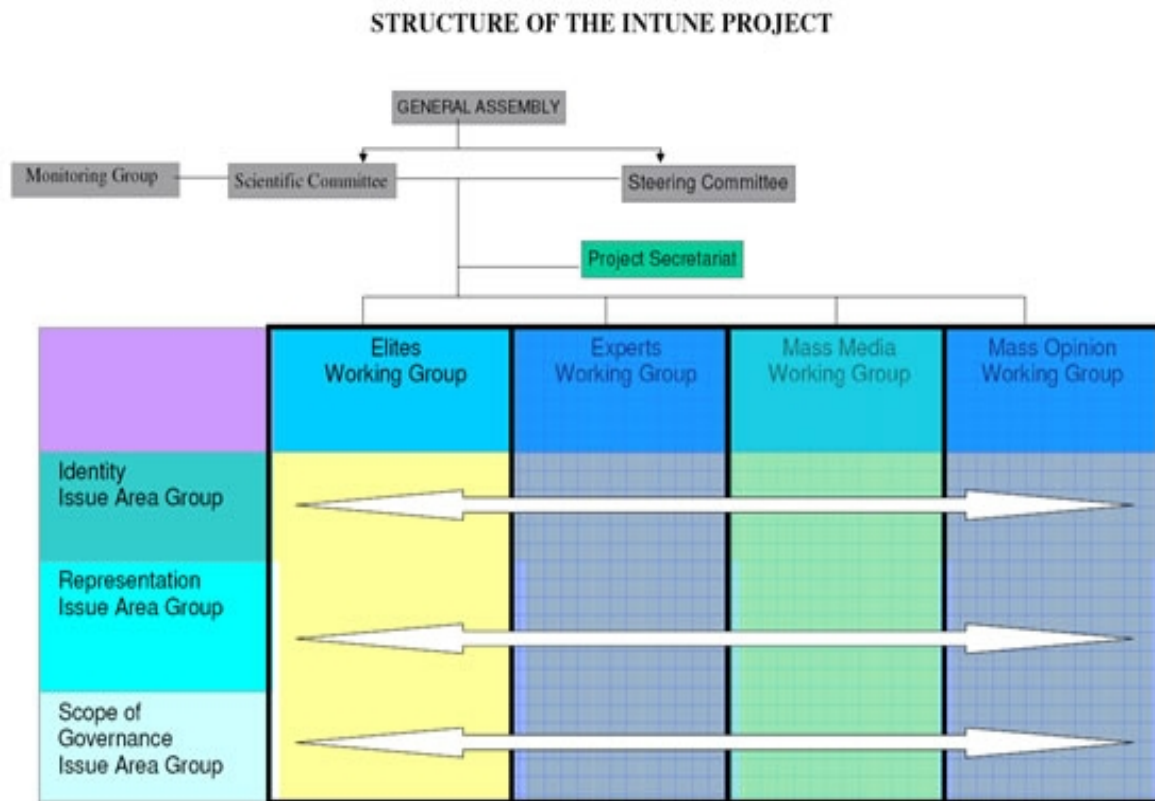
La réalisation, puis l'analyse du corpus IntUne peut ainsi mener à l'élaboration d'un dictionnaire pour sociologues, cet ouvrage donnant une représentation fine de concepts tels que l'identité, la nation, l'Europe, etc. dont l'importance est fondamentale en sociologie ou

7- Conclusion.

même en sciences politiques. La constitution d'autres corpus spécifiques pourrait également mener à la création de dictionnaires répondant avec davantage d'acuité aux besoins d'apprenants en langue spécialisés dans les domaines de la biologie, des affaires et des finances, de la chimie, du droit, ou encore de la médecine. La linguistique de corpus présente en conséquence de nombreuses ressources dans différents domaines, allant de la lexicographie à l'enseignement qu'il conviendra d'exploiter dans les recherches à venir.

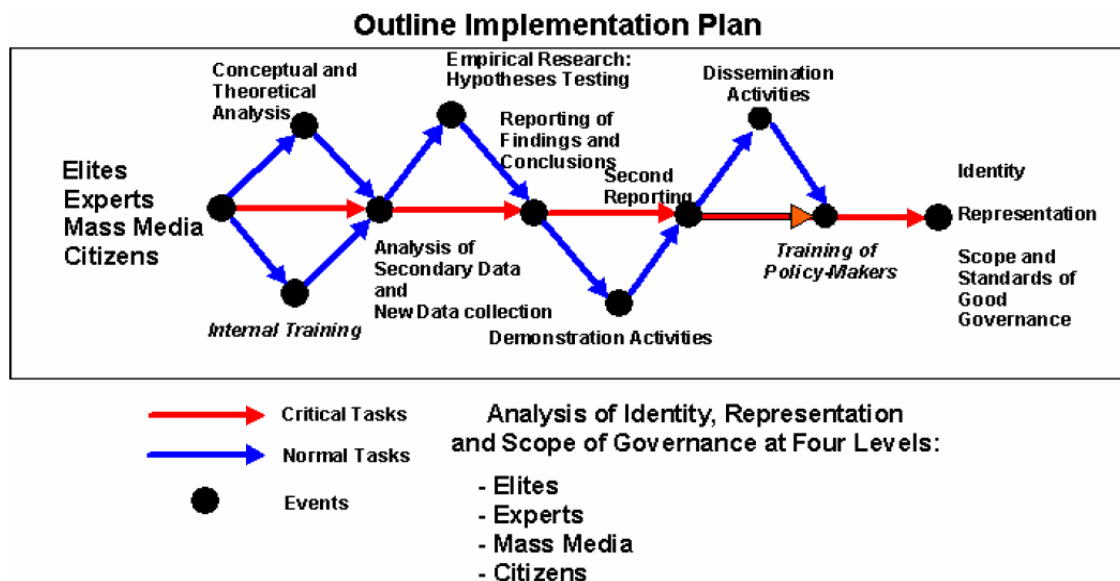
Annexe 1. Organigramme du projet IntUne.

<http://www.intune.it/structure/welcome>



Annexe 2. Fonctionnement des quatre groupes du projet IntUne.

<http://www.intune.it>



Annexe 3. Extrait du corpus presse brut.

```
- <TEI xmlns="http://www.tei-c.org/ns/1.0">
```

```
<teiHeader>
- <fileDesc>
- <titleStmt>
  <title>Le Figaro</title>
- <respStmt>
  <resp>data capture and transcription</resp>
  <name>Lorient Intune Media Group</name>
</respStmt>
</titleStmt>
- <editionStmt>
  <p>INTUNE Pilot Corpus French language press texts</p>
</editionStmt>
- <publicationStmt>
  <distributor>Lorient Intune Media Group</distributor>
- <address>
  <addrLine>Maison de la Recherche, Universite de Bretagne Sud, 4 rue Jean Zay, BP
92116, 56321 LORIENT CEDEX, FRANCE .</addrLine>
  <addrLine>Telephone: +33 2 97 87 65 31</addrLine>
</address>
</publicationStmt>
- <sourceDesc>
- <bibl>
  <title>Le Figaro</title>
  <publisher>Société du Figaro</publisher>
  <pubPlace>Paris</pubPlace>
  <date value="2007-02-05" />
</bibl>
</sourceDesc>
</fileDesc>
- <profileDesc>
- <textClass>
  <catRef target="Fr wr con nt fig" />
</textClass>
</profileDesc>
</teiHeader>
- <text>
- <body>
- <div1>
- <div2 xml:id="fig070205_1" type="LE FIGARO - Une-FIG" n="1">
- <div3 type="Editorial">
  <byline />
  <head type="main">Sarkozy creuse l'écart avec Royal</head>
  <head type="sub">La percée de Bayrou se confirme</head>
```

Annexe 3. Extrait du corpus presse brut.

<dateline />

<p>AVEC 13 % d'intentions de vote au premier tour, le président de l'UDF poursuit sa progression et gagne 4 points. Il devance Jean-Marie Le Pen (12,5 %) et conforte ainsi sa stature de possible « troisième homme » de la présidentielle. Page 8</p>

</div3>

</div2>

Annexe 4- Extrait du corpus télévisé brut.

```
- <TEI>
```

```
- <teiHeader>
- <fileDesc>
- <titleStmt>
  <title>tfr320070212</title>
</titleStmt>
- <publicationStmt>
  <authority>IntUne Media Group</authority>
- <pubPlace>
  <name>Siena</name>
  </pubPlace>
</publicationStmt>
- <sourceDesc>
- <bibl>
  <title>tfr320070212</title>
</bibl>
- <recordingStmt>
- <recording type="audio">
- <broadcast>
- <bibl>
  <distributor />
  <date value="" />
</bibl>
</broadcast>
</recording>
</recordingStmt>
</sourceDesc>
</fileDesc>
- <encodingDesc>
  <p>_</p>
</encodingDesc>
- <profileDesc>
- <textClass>
  <catRef target="Fr sp reg news fr3" />
</textClass>
- <particDesc>
- <listPerson>
  <person xml:id="bayrou_françois" sex="1" role="candidat udf à l'ump" />
  <person xml:id="courtellemont_marie-pierre" sex="2" role="reporter" />
  <person xml:id="delloye_amélie" sex="2" role="reporter" />
  <person xml:id="pulvar_audrey" sex="2" role="présentatrice" />
  <person xml:id="berrou_laure-anne" sex="2" role="reporter" />
  <person xml:id="de-villiers_philippe" sex="1" role="candidat mpf à l'élection
présidentielle" />
```



```

<person xml:id="lepage_corinne" sex="2" role="candidate cap 21 à l'élection
présidentielle et avocate partie civile" />
<person xml:id="bougrain-dubourg_alain" sex="1" role="président de la ligue de
protection des oiseaux" />
<person xml:id="soulez-larivière_daniel" sex="1" role="avocat du groupe total" />
<person xml:id="priou_christophe" sex="1" role="maire ump du croisic" />
<person xml:id="devillairs_marcel" sex="1" role="propriétaire du camping l'océan" />
<person xml:id="bonnel_benoît" sex="1" role="conchyliculteur, collectif anti marée
noire" />
<person xml:id="dumas_dominique" sex="1" role="reporter" />
<person xml:id="charroy_myriam" sex="2" role="fille de victimes" />
<person xml:id="oudard_yannick" sex="1" role="rescapé" />
<person xml:id="petitcolas" sex="1" role="lieutenant colonel dirigeant les secours" />
<person xml:id="chollet_jean-marc" sex="1" role="reporter" />
<person xml:id="denis_sylvain" sex="1" role="président de la fédération nationale des
associations de retraités" />
<person xml:id="karniewicz_danièle" sex="2" role="présidente de la caisse nationale
de l'assurance vieillesse" />
<person xml:id="leclercq_gérard" sex="1" role="reporter" />
<person xml:id="allenbach_jean-philippe" sex="1" role="candidat fédéraliste à
l'élection présidentielle" />

</listPerson>
</particDesc>
</profileDesc>
</teiHeader>
= <body>
= <text xml:lang="fr">
= <div1>
= <div2 type="newsitem">
= <div3 type="headlines">
  <u who="#pulvar_audrey">Bonsoir, dans cette édition, le début du procès du naufrage de
  l'erika. Plus de soixante dix parties civiles entendent faire de ce procès un exemple, quinze
  personnes morales et physiques dont la compagnie total sont sur les bancs des accusés, en
  décembre quatre vingt dix neuf, l'erika avait déversé vingt mille tonnes d'un fioul
  particulièrement polluant sur près de quatre cents kilomètres de côte. Alors qu'une bonne
  partie de l'ouest du pays est frappée aujourd'hui par une tempête, nous retournerons dans
  l'est où une tornade a soufflé des dizaines de maisons hier, les habitants de Soulosse sous St
  Elof sont encore sous le choc. Ségolène Royal rencontrait des enseignants dans un lycée de
  Seine St Denis aujourd'hui, elle a évoqué avec eux la modification de la carte scolaire, nous
  reviendrons sur les propositions faites par la candidate socialiste, et Nicolas Sarkozy, et sur
  leur financement, pour le moment très imprécis. Après la polémique suscitée par le départ
  de Johnny Halliday en Suisse, nous consacrerons notre dossier ce soir au forfait fiscal dont
  bénéficient les Français les plus aisés en Suisse, à peine un millier de ces contribuables
  riches y ont droit. Et puis le champix fait son apparition en pharmacie pour aider les
  fumeurs à décrocher, il ne diffuse pas de nicotine dans le sang, mais agit directement sur le
  cerveau. </u>
</div3>

```

<div3 type = "intro">

<u who="#pulvar_audrey">

C'est le premier procès en France s'attachant à déterminer les responsabilités dans une grande catastrophe écologique. Celle provoquée par l'erika, ce navire vétuste dont l'avarie en pleine tempête en décembre quatre vingt dix neuf avait provoqué la pollution de quatre cents kilomètres de littoral souillé par vingt mille tonnes de fioul lourd. Soixante dix parties civiles entendent faire aujourd'hui de ce procès un exemple. Laure-Anne Berrou, Claude Pfafvman.</u>

</div3>

= <div3 type="report">

<u who="#berrou_laure-anne">Marée de robes noires au palais de justice de Paris. Une centaine d'avocats pour quinze prévenus et soixante parties civiles. Six cent mille euros de budget, une traduction simultanée en trois langues, une deuxième salle pour les journalistes. Le procès de l'erika est à la hauteur de l'émotion suscitée par la catastrophe de Noël quatre vingt dix neuf. Hasard du calendrier, l'élection présidentielle s'est largement invitée à cette audience.</u>

<u who="#de-villiers_philippe">Moi, je suis là, euh, pour défendre un département qui a été souillé, une côte atlantique qui a été blessée, et aussi pour euh, pour penser à l'avenir parce que depuis que nous nous parlons, il y a déjà euh, deux bateaux poubelle qui sont passés sur le rail d'Ouessant.</u>

<u who="#lepage_corinne">J'en attends que la justice se fasse, et aussi que le coût soit élevé, non pas par plaisir de faire payer, mais parce que nous vivons dans un monde économique, où seul compte, malheureusement, la finance, et il faut que polluer les côtes, ça coûte très cher, sinon, ça continuera.</u>

<u who="#berrou_laure-anne">Elus de tous bords contre multinationale, l'occasion était trop belle. Dans les couloirs du palais, on attend même Ségolène Royal, mais la présidente de Poitou Charentes, partie civile, viendra un autre jour. Les associations écologistes sont là pour défendre la nature.</u>

<u who="#bougrain-dubourg_alain">Actuellement, on indemnise, et c'est normal, aussi bien les ostréiculteurs, les conchyliculteurs, que les marins pêcheurs, mais on ne tient pas compte des espèces sauvages, des oiseaux, des tortues, des phoques qui n'ont pas de valeur, euh, aux yeux du délit.</u>

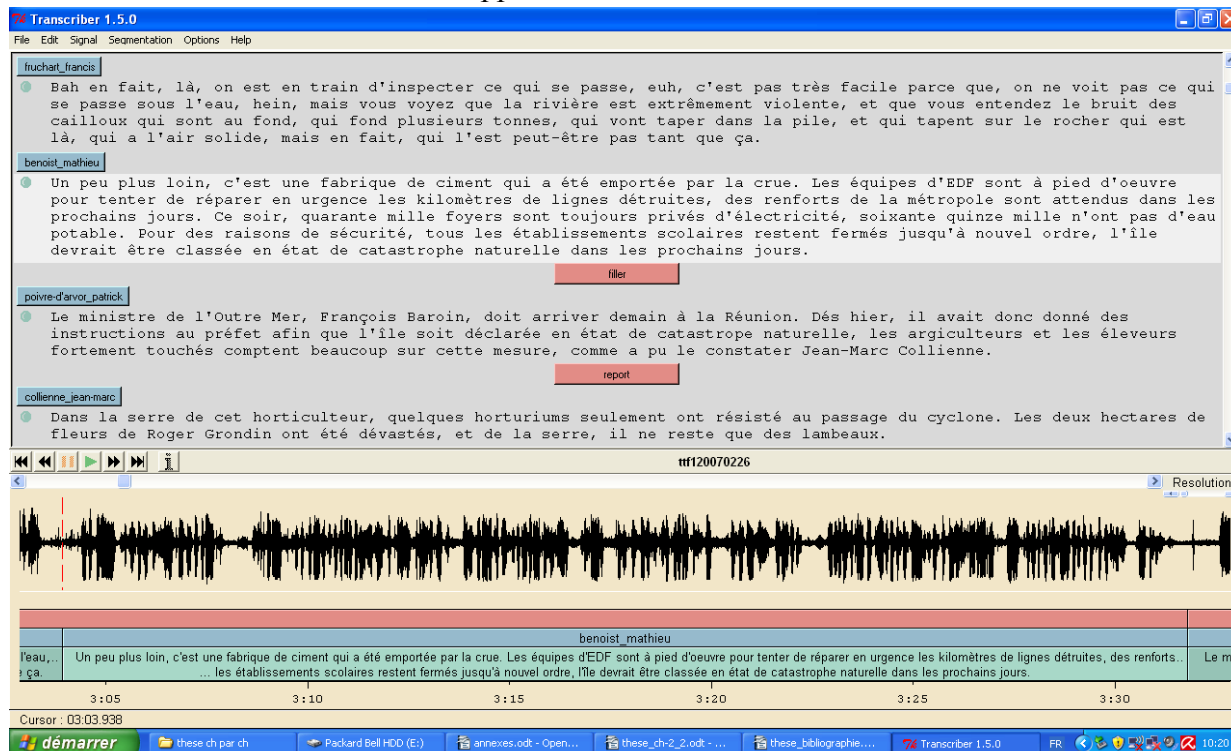
<u who="#berrou_laure-anne">Côté prévenu, seul le propriétaire de l'erika a fait une entrée remarquée. Ce golden boy italien est soupçonné d'avoir violé délibérément les règles de sécurité. Le capitaine Indien n'a pas fait le déplacement, le groupe total fait donc figure de principal accusé.</u>

<u who="#soulez-larivière_daniel">La compagnie prend des navires qui sont certifiés, qui disposent de certificats. En l'espèce, il les avait tous, il était fort bien coté par les ports, il venait de passer en grande visite quinze mois plus tôt, alors en conséquent il n'y avait aucune raison qu'il ne soit pas affrété.</u>

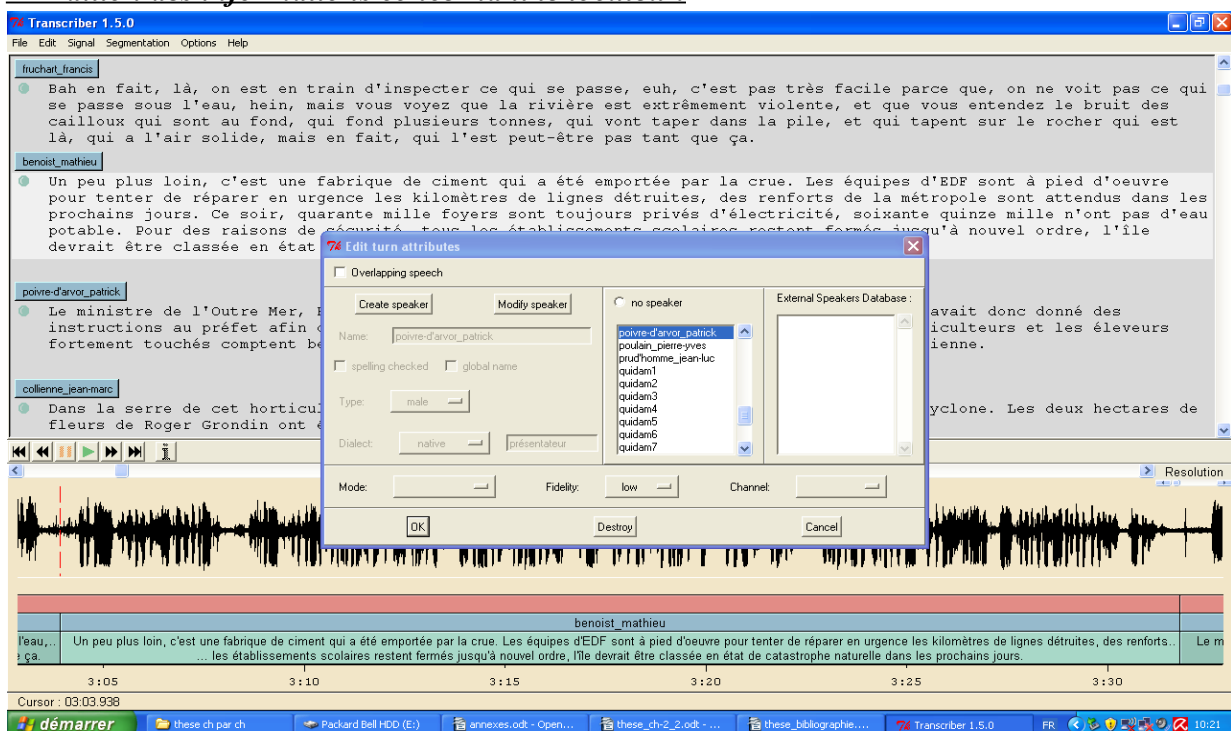
<u who="#berrou_laure-anne">Dans la mesure où le navire battait pavillon étranger, Malte, en l'occurrence, aucun prévenu ne risque la prison. En revanche, le procès aura une vertu pédagogique : comprendre, et juger de pratiques parfois malhonnêtes du monde maritime, et puis une nouvelle donnée économique pour total, combien peut coûter une marée noire.</u> </div3>

Annexe 5- Captures d'écran du logiciel transcriber

A- Saisie du traitement de texte en rapport avec le fichier son.



B- Édition des informations concernant le locuteur.



Annexe 6- Feuille de style xslt utilisée pour le corpus télévisé.

```
<?xml version="1.0" encoding="UTF-8" ?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="2.0">
  <xsl:output method="xml" encoding="UTF-8" indent="yes" />
  <xsl:template match="Trans">
    <xsl:for-each select="/Speakers/Speaker">
      <xsl:element name="person">
        <xsl:for-each select="/@name">
          <xsl:attribute name="xml:id">
            <xsl:value-of select="." />
          </xsl:attribute>
        </xsl:for-each>
        <xsl:if test="@type='male'">
          <xsl:attribute name="sex">1</xsl:attribute>
        </xsl:if>
        <xsl:if test="@type='female'">
          <xsl:attribute name="sex">2</xsl:attribute>
        </xsl:if>
        <xsl:if test="@type='unknown'">
          <xsl:attribute name="sex">9</xsl:attribute>
        </xsl:if>
        <xsl:for-each select="/@accent">
          <xsl:attribute name="role">
            <xsl:value-of select="." />
          </xsl:attribute>
        </xsl:for-each>
        <xsl:if test="@check='no'">
          <xsl:attribute name="n">quidam</xsl:attribute>
        </xsl:if>
        <xsl:if test="@check='yes'">
          <xsl:choose>
            <xsl:when test="@dialect='native'">
              <xsl:attribute name="n">politique-liberal</xsl:attribute>
            </xsl:when>
            <xsl:otherwise>
              <xsl:attribute name="n">politique-conservative</xsl:attribute>
            </xsl:otherwise>
          </xsl:choose>
        </xsl:if>
      </xsl:element>
    </xsl:for-each>
    <xsl:for-each select="/Episode">
      <xsl:element name="div2">
        <xsl:attribute name="type">newsitem</xsl:attribute>
        <xsl:for-each select="/@program">
          <xsl:attribute name="xml:id">
            <xsl:value-of select="." />
          </xsl:attribute>
        </xsl:for-each>
        <xsl:for-each select="/Section">
          <xsl:element name="div3">
            <xsl:if test="@type='filler'">
              <xsl:attribute name="type">intro</xsl:attribute>
            </xsl:if>
          </xsl:element>
        </xsl:for-each>
      </xsl:element>
    </xsl:for-each>
  </xsl:template>
</xsl:stylesheet>
```

Annexe 6- Feuille de style xslt utilisée pour le corpus télévisé.

```
</xsl:if>
= <xsl:if test="@type='report'">
  <xsl:attribute name="type">report</xsl:attribute>
</xsl:if>
= <xsl:if test="@type='nontrans'">
  <xsl:attribute name="type">exchange</xsl:attribute>
</xsl:if>
= <xsl:for-each select="./Turn">
= <xsl:element name="u">
= <xsl:if test="@speaker='spk1'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[1]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk2'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[2]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk3'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[3]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk4'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[4]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk5'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[5]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk6'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[6]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk7'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[7]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk8'">
= <xsl:attribute name="who">
#
<xsl:value-of select="ancestor::Trans/Speakers/Speaker[8]/@name" />
</xsl:attribute>
</xsl:if>
= <xsl:if test="@speaker='spk9'">
```

Annexe 6- Feuille de style xslt utilisée pour le corpus télévisé.

```
- <xsl:attribute name="who">
#
  <xsl:value-of select="ancestor::Trans/Speakers/Speaker[9]/@name" />
</xsl:attribute>
</xsl:if>
- <xsl:if test="@channel='studio'">
  <xsl:attribute name="type">camera</xsl:attribute>
</xsl:if>
- <xsl:if test="@channel='telephone'">
  <xsl:attribute name="type">voiceover</xsl:attribute>
</xsl:if>
  <xsl:value-of select="." />
</xsl:element>
</xsl:for-each>
</xsl:element>
</xsl:for-each>
</xsl:element>
</xsl:for-each>
</xsl:template>
</xsl:stylesheet>
```


Annexe 7- Carte physique de l'Europe.



La courbe bleue continue représente une délimitation selon ce qui est donné par le Larousse comme des informations certaines, la partie pointillée représente une délimitation possible qui est effectuée selon les informations prises en compte "traditionnellement".

Annexe 8- Collocations de "Europe" dans le corpus IntUne 2007 et 2009.

Rg	2007			2009		
	collocat	fréquence	z-score	collocat	fréquence	z-score
1	l'	4811	116,5	orientale	80	137,9
2	orientale	68	110,2	centrale	200	99
3	continentale	55	108,7	en	2803	96,3
4	coupe	321	105,7	l'	3344	93,5
5	en	3589	103,3	occidentale	61	89,3
6	championnats	121	90,7	coupe	188	80,2
7	airpost	15	73,1	continentale	31	77,1
8	centrale	151	67,6	aew	9	62,1
9	occidentale	43	57,9	championnats	61	53,6
10	ettu	9	53,7	xxl	16	52,9
11	champion	109	48,7	écologie	49	51,9
12	cers	8	47,7	seine-nord	10	51,7
13	odéon-théâtre	6	46,3	émergente	11	48,5
14	asie	83	45,7	ouest	99	42,2
15	ouest	129	45,3	ecologie	24	41,9
16	ios	7	44	antimissile	24	41
17	nord	147	43,6	nord	103	39,7
18	etats-unis	141	42,9	etats-unis	109	38,8
19	d'	1910	40,8	asie	56	37,4
20	conventionnelles	12	39,9	états-unis	104	37,3
21	championne	36	34,1	d'	1445	35,9
22	aew	4	33,8	pays	227	35,3
23	l/tv5	3	32,7	carl-peter	6	34,4
24	avr-05	3	32,7	chi-x	6	32
25	streettracks	3	32,7	protège »	3	31,1
26	états-unis	91	32,1	niederösterreich	3	31,1
27	chrétiennes	16	32	défense	99	30,9
28	sud	121	31,9	moyen-orient	27	30,4
29	moyen-orient	35	31,8	osce	6	30
30	l	286	31	champion	50	29,6
31	netjets	4	28,5	sud	94	29,3
32	amérique	58	28,4	forster	6	29,2

Annexe 8- Collocations de "Europe" dans le corpus IntUne 2007 et 2009.

33	snc-lavalin	5	28,4	l	227	29,1
34	schott	3	28,3	1664	3	27,8
35	express	34	27,1	lipietz	3	27,8
36	msci	7	25,7	amérique	52	25,7
37	saint-groux	3	25,3	stx	10	25,2
38	seine-nord	4	25,1	arca	4	24,9
39	pays	206	24,8	gm	27	23,5
40	antimissile	14	24,3	surfrider	4	22,9
41	vieille	36	23,6	lycos	3	21,9
42	infiniti	5	23,5	quiksilver	5	21,4
43	partout	48	23,1	fogiel	4	21,3
44	ace	5	22,8	vieille	26	21,2
45	toute	123	22,3	frigide	3	20,6
46	france	287	21,9	« l'	35	20,5
47	patries	3	21,3	civilisations	10	20,5
48	élargie	12	21,2	turquie	25	20,1
49	et	1496	20,9	championne	14	20
50	osce	4	20,8	forte »	4	19,4
51	supranationale	4	20,8	et	1100	19,1
52	fce	4	20,8	partout	33	18,8
53	tv5	7	20,4	championnat	38	18,7
54	lyle	5	20,4	indoor	3	18,6
55	cfe	3	19,9	afi	3	18,6
56	vice-champion	9	19,6	schweppes	5	17,5
57	championnat	50	19,4	hauts-fourneaux	5	17,5
58	unie	10	18,7	champions	23	16,8
59	passoire	4	18,7	proche-orient	13	16,5
60	ixis	10	18,6	rendez-vous »	3	16,5
61	aig	6	18,6	vice-champion	7	15,9
62	rodamco	7	18	faugère	3	15,9
63	nations	28	17,8	marc-olivier	3	15,9
64	pacifiée	3	17,8	est	643	15,4
65	adobe	5	17,3	unie	6	14,9
66	sociale	62	17,2	osons	3	14,9
67	référendum	30	17	devises	11	14,6
68	libénautes	5	17	protège	9	14,5
69	rink-hockey	3	17	kronenbourg	3	14,5
70	civilisations	10	16,4	voudrions	3	14,5
71	toyal	5	16,4	entière	15	14



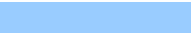





Annexe 8- Collocations de "Europe" dans le corpus IntUne 2007 et 2009.

72	rugby	32	16,2	route	68	13,8
73	afrique	41	16,1	expressions	7	13,8
74	relance	21	16	défense »	3	13,7
75	panne	17	15,8	hudson	4	13,6
76	mutuel	13	15,7	japon	24	13,4
77	pharmaceuticaux	3	15,6	foundation	4	13,2
78	02-avr	3	15,6	snc-lavalin	3	13,1
79	boycott	10	15,3	visa	7	12,8
80	coupes	11	15,2	colony	6	12,8
81	méridionale	4	15,2	libérale	8	12,7
82	val	12	14,9	atlantique	16	12,6
83	emea	5	14,6	voulons	17	12,4
84	champions	29	14,5	coupes	10	12,4
85	anoeta	4	14,3	asie-pacifique	4	11,8
86	chatignoux	4	14,3	computer	3	11,7
87	restrictives	5	14,1	solidaire	9	11,4
88	désamou	3	14	filiales	13	11,3
89	scouts	8	13,8	opel	12	11,3
90	ps3	7	13,7	vice-président	19	11,1
91	gros-verheyde	3	13,6	antimissiles	3	10,9
92	relancer	19	13,5	afrique	23	10,6
93	bouclier	15	13,4	construire	19	10,6
94	vice-président	27	13,1	« grand	5	10,6
95	budapest	6	12,8	delors	4	10,6
96	championnes	4	12,5	liste	30	10,5
97	hors	37	12,3	record	22	10,5
98	protectrice	4	12	lgy	6	10,5
99	est	761	11,7	endettés	4	10,3
100	tv	16	11,7	implantés	4	10,2

Ce tableau présente les groupes collocationnels suivants en 2007 et 2009 :

couleur	thème	2007	2009
	Mot lexicalement vide	5	4
	Entité présentant différentes zones géographiques	9	8
	Entité sportive	15	14
	Entité géographique et physique	4	3
	Entité économique	22	28
	Entité culturelle (religion, civilisation, média, culture, etc.)	8	5

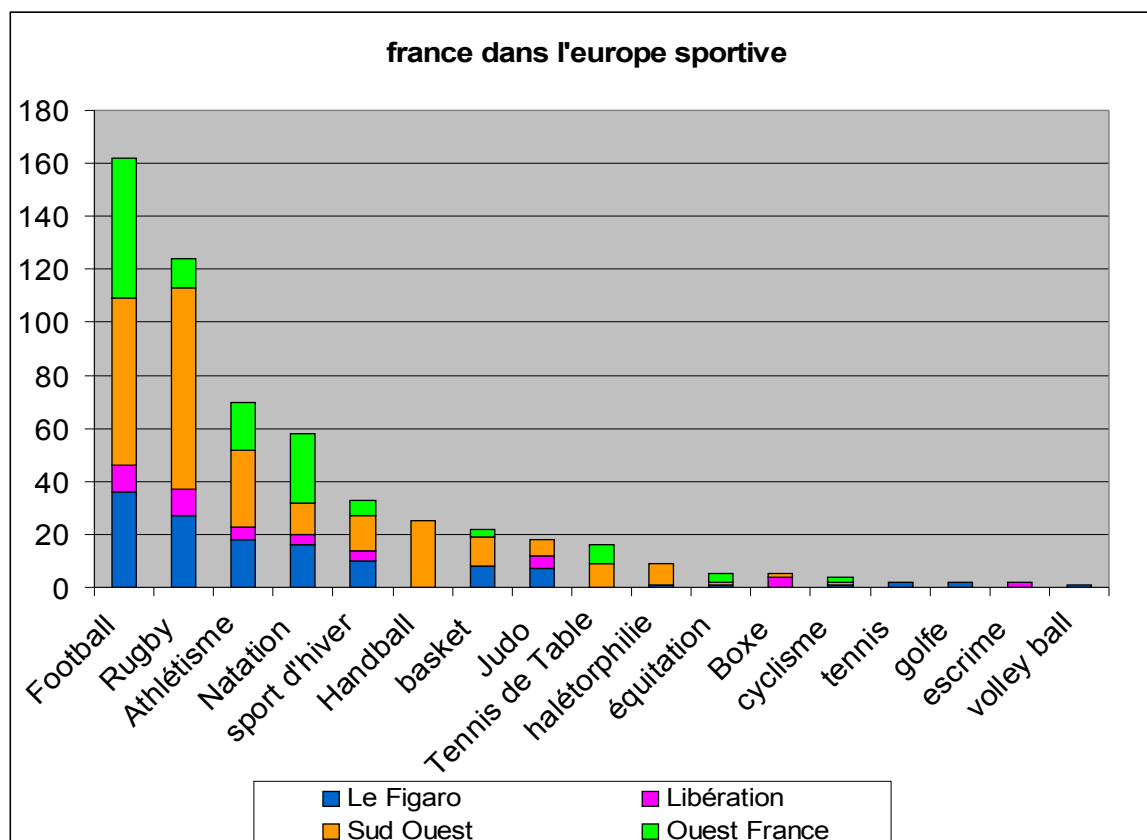
Annexe 8- Collocations de "Europe" dans le corpus IntUne 2007 et 2009.

	Entité ayant des rapports avec d'autres zones dans le monde	7	10
	Entité qui se défend	6	5
	Entité qui est un nom de marque	3	1
	Entité composée de membres	5	1
	Entité symbolique	8	10
	Entité politique	7	5
	Entité à l'égard de laquelle une volonté existe	0	3
	Entité avec une réflexion écologique	0	2
	Turquie	1	1
TOTAL		100	100

Annexe 9- La France dans l'Europe sportive.

Nombre d'occurrences par sport et par journal où "France" est associée à "Europe" dans un contexte sportif.

	Le Figaro	Libération	Sud Ouest	Ouest France	TOTAL
Football	36	10	63	53	162
Rugby	27	10	76	11	124
Athlétisme	18	5	29	18	70
Natation	16	4	12	26	58
sport d'hiver	10	4	13	6	33
Handball	0	0	25	0	25
basket	8	0	11	3	22
Judo	7	5	6	0	18
Tennis de Table	0	0	9	7	16
halétophilie	1	0	8	0	9
équitation	1	0	1	3	5
Boxe	0	4	1	0	5
cyclisme	1	0	1	2	4
tennis	2	0	0	0	2
golfe	2	0	0	0	2
escrime	0	2	0	0	2
volley ball	1	0	0	0	1
TOTAL	130	44	255	129	558



Annexe 10- Les différents pays de l'Europe sportive.

Ce tableau présente les différents pays prenant part à des compétitions européennes, en fonction des sports dans lesquels ils s'illustrent.


Pays	Fréquence	Sport (occurrences)
Royaume Uni	65	-Football (30) -Rugby (30) -Athlétisme (3) -Glisse (1) -F1 (1)
Italie	30	-Football (20) -Cyclisme (4) -Rugby (4) -Boxe (1) -Patinage (1)
Allemagne	23	-Football (10) -Tennis de Table (9) -Patinage artistique (4)
Espagne	21	-Football (12) -Voile (3) -Basket (3) -Corss (1) -F1 (1) -Handball (1)
Irlande	13	-Rugby (13)
Lituanie	10	-Football (5) -Basket (5)
Pologne	7	-Tennis de table (4) -Football (2) -Basket (1)
Serbie	5	-Athlétisme (3) -Football (1) -Judo (1)
Portugal	4	-Football (3) -Judo (1)
Grèce	4	-Football (3) -Tennis de table (1)
Pays Bas	3	-Nautisme (2) -Football (1)
Suisse	3	-Patinage artistique (3)
Danemark	2	-Tennis de table (1) -Basket (1)
Arménie	2	-Haltérophilie (2)
Russie	2	-Halétophilie (1) -Basket (1)
République Tchèque	2	-Patinage artistique (2)
Suède	1	-Athlétisme (1)
Israël	1	-Basket (1)

Annexe 11- Contexte de « du pays ».





Journal	Premier sujet	Deuxième sujet	Troisième sujet
Le Figaro	Guerre ou terrorisme	Economie	Tensions politiques
Libération	Guerre ou terrorisme	Environnement	Droits de l'homme
Ouest France	Guerre ou terrorisme	Environnement	Economie
Sud Ouest	Guerre ou terrorisme	Solidarité internationale	Environnement

Ce tableau indique les sujets les plus importants dans lesquels la colligation « du pays » apparaît dans les différents journaux du corpus IntUne 2007 en fonction de leur fréquence.





Annexe 12 – Tableau récapitulatif du reportage de France 3.

Pla n	durée	Description et commentaire.	image
Introduction par la présentatrice			
1	16,08	Plan fixe. <i>Le bilan est encore provisoire, mais on parle déjà d'une centaine de morts dans le centre de l'Italie touché par un séisme de magnitude six virgule trois sur l'échelle de Richter. Une vingtaine de villes et de villages ont été partiellement ou totalement détruits.</i>	
2	5,18	Une petite animation fait apparaître le nom de l'Aquila dans un second temps. <i>Cinquante mille personnes, au moins, sont considérées comme sans abris. Stéphanie Pérez.</i>	
Reportage			
1	10,45	Homme sauvé des décombres, en pleur dans les bras de proches, des secouristes passent. Zoom sur l'homme quand il pleure <i>« Il était prisonniers sous les décombres de sa maison. Il s'en sort indemne. Il est l'un des miraculés de la catastrophe. »</i>	


Annexe 12 – Tableau récapitulatif du reportage de France 3.

2	2,12	<p>Le même homme, escorté par les secouristes dans une rue, en sous vêtements. La caméra fait un zoom arrière, donnant l'impression de marcher devant l'homme. <i>« Comme cet homme, les habitants de l'Aquila, épiceutre ...</i></p>	
3	3,72	<p>Des secouristes s'affairent autour d'une maison. Plan fixe. <i>... du séisme ont tous été surpris dans leur sommeil par la violence...</i></p>	
4	0,9	<p>Même chose, image fixe <i>... de la secousse. Depuis ce matin,</i></p>	
5	2,16	<p>La caméra amorce un mouvement latéral de gauche à droit en descendant peu à peu. On voit toujours des secouristes s'affairer. <i>... secouristes et volontaires s'acharnent sur le tas de gravats...</i></p>	





Annexe 12 – Tableau récapitulatif du reportage de France 3.

6	2,49	<p>La caméra suit le mouvement du bras du civil qui dégage les décombres, de bas en haut. <i>...qu'est devenue leur ville. Ici et là,</i></p>	
7	2,04	<p>La caméra suit le mouvement du chien : mouvement latéral de gauche à droite. <i>...les chiens signalent encore des traces de vie.</i></p>	
8	7,63	<p>Plan fixe rapproché sur l'interviewé. <i>« C'est une femme qui est coincée, mais on n'arrive pas à comprendre où elle se trouve, parce que ce sont des tous petits trous par lesquels on peut à parler. »</i> <i>Pendant l'interview on entend un peu la voix originale couverte par la traduction.</i></p>	
9	2,25	<p>Image fixe. <i>Il faut faire vite, sauver ceux qui peuvent l'être encore.</i></p>	





Annexe 12 – Tableau récapitulatif du reportage de France 3.

10	2,4	Image fixe. <i>Derrière le barrage pudique de ces secouristes,</i>	
11	2,07	Image Fixe. <i>...une jeune femme vient d'être extraite des ruines.</i>	
12	8,53	Plan fixe <i>« Tout doucement, avec de la patience, et malgré la fatigue, on a réussi à creuser avec nos mains, jusqu'à arriver à cette jeune fille qui nous appelait. »</i>	
13	10,28	Plan fixe <i>« On creuse avec les mains, parce qu'avec les machines, l'état des immeubles, ce serait trop dangereux. »</i>	

Annexe 12 – Tableau récapitulatif du reportage de France 3.

14	4	<p>Mouvement latéral de la droite vers la gauche, qui suit la marche des secouristes. <i>« - Il y a quelqu'un ? »</i></p>	
15	5,35	<p>La caméra va du haut vers le bas, puis de gauche à droite, pour suivre la jeune femme. <i>Les secours sillonnent la ville pour évacuer des habitants terrorisés.</i></p>	
16	3,35	<p>Mouvement de caméra qui suit le déplacement de l'hélicoptère, de droite vers gauche <i>Les dégâts sont considérables,</i></p>	
17	2,57	<p>Mouvement de caméra qui suit le déplacement de l'hélicoptère de la droite vers la gauche. <i>Au moins dix mille bâtiments se sont effondrés...</i></p>	

Annexe 12 – Tableau récapitulatif du reportage de France 3.

18	3,02	<p>Plan fixe <i>... dans cette région riche d'édifices moyen-âgeux. Les Eglises ont...</i></p>	
19	2,73	<p>Mouvement de droite à gauche latéral <i>... perdu leurs clochers,</i></p>	
20	3,27	<p>Zoom avant <i>... les routes se sont coupées en deux.</i></p>	
21	4,74	<p>Zoom arrière qui part de la voiture pour arriver à la totalité de la rue. <i>Il faudra beaucoup de temps pour effacer les traces de la catastrophe, a reconnu ce matin le chef du gouvernement,</i></p>	




Annexe 12 – Tableau récapitulatif du reportage de France 3.

22	2,98	Plan fixe, un secouriste aide quelqu'un au bord du gouffre. <i>...Silvio Berlusconi.</i>	
23	1,43	Plan fixe <i>Hagards,</i>	
24	4,12	Zoom avant sur des blessés. <i>... les survivants sont prostrés depuis ce matin.</i>	
25	6,37	Plan fixe. <i>« -Cette nuit, c'était la fin du monde. L'immeuble tanguait, on ne pouvait pas marcher. »</i>	

Annexe 12 – Tableau récapitulatif du reportage de France 3.

26	4,36	Zoom sur l'interviewée « - Ce que j'ai vécu cette nuit, il n'y a pas de mots pour le raconter. »	
27	3,76	Plan fixe. Une couverture sur les épaules, une valise à la main...	
28	3,22	Plan fixe ...une grosse partie des cinquante mille habitants de la ville a quitté la zone...	
29	1,96	Zoom avant de la caméra qui donne l'impression de suivre les marcheurs. ... ravagée, par peur des répliques.	





Annexe 12 – Tableau récapitulatif du reportage de France 3.

30	2,41	Plan fixe <i>Les autres...</i>	
31	4,45	Mouvement descendant puis zoom arrière sur un blessé transporté en civière. <i>... ceux qui sont hospitalisés ou hébergés dans des structures d'accueil redoutent désormais la nuit qui s'annonce...</i>	
32	2,77	Plan fixe. <i>... et cherchent toutes les protections possibles.</i>	





Annexe 13- Tableau récapitulatif du reportage de TF1.

Plan	Durée	Description	Image
Introduction par la présentatrice			
1	20	<p>Plan fixe sur la présentatrice qui s'exprime sans quitter la caméra des yeux. Son nom est présent en bas de l'écran quelques instants puis disparaît. Un léger zoom est effectué vers elle au début du plan.</p> <p><i>Le séisme a donc surpris la population en pleine nuit. Il était trois heures et demies du matin lorsque le centre de l'Italie a été frappé par un violent tremblement de terre. La secousse, d'une magnitude supérieure à six degrés, a été ressentie dans presque tout le pays, même jusqu'à Rome. Le bilan, encore provisoire, n'a cessé de s'alourdir</i></p>	
2	6,72	<p>Image animée. Les cercles autour de l'Aquila se répandent de façon concentrique, comme une onde sismique.</p> <p><i>au fil de la journée. Il est ce soir d'au moins cent cinquante morts et mille cinq cents blessés. Récit de Michel Scott.</i></p>	
Reportage			
1	7,08	<p>La caméra bouge beaucoup, les secouristes se fraient un passage compliqué à travers les gravats.</p> <p><i>Il faut attendre les premières lueurs du jour pour réaliser l'ampleur du désastre qui s'est abattu sur l'Aquila.</i> [bruits de gravats qui s'entrechoquent]</p>	





Annexe 13- Tableau récapitulatif du reportage de TF1.

2	4,27	<p>Zoom avant, même scène, mais au bord d'un gouffre, les secouristes s'y affairant. [voix féminine et masculine italiennes non traduites + gravats].</p>	
3	2,57	<p>Plan légèrement ascendant sur une femme qui pleure dans les bras d'un homme. <i>Au détour des ruelles, c'est d'abord le choc et la stupeur.</i> [Voix féminine]</p>	
4	4,08	<p>Zoom arrière sur deux hommes dans une rue puis mouvement latéral vers la droite, la caméra filme alors des secouristes devant des décombres <i>En pleine nuit, à trois heures trente deux du matin, très précisément,</i> [voix mélangées, cris, gravats]</p>	
5	2,48	<p>Intérieur d'un bâtiment éventré par le séisme. Plan fixe <i>la terre tremble, avec une violence sans précédent.</i> [Voix mélangées]</p>	

Annexe 13- Tableau récapitulatif du reportage de TF1.

6	3,52	<p>Secouristes civils et officiels devant une faille dans les décombres. Plan fixe.</p> <p><i>Des dizaines de maisons s'effondrent</i> [gravats, voix mélangées]</p>	
7	4,04	<p>Une femme blessée est évacuée sur une civière. La caméra suit le mouvement latéral de gauche à droite légèrement descendant de la civière.</p> <p><i>sur des habitants pris au piège.</i> [Gravats, voix mélangées, cris]</p>	
8	2,84	<p>Zoom arrière sur une rue avec plusieurs personnes qui portent une autre civière, amorce de mouvement latéral droite gauche à la fin.</p> <p><i>Très vite, les secours s'improvisent. On sauve ceux qui sont immédiatement accessibles,</i> [voix mélangées]</p>	
9	5,52	<p>Zoom avant vers un groupe de secouristes avec un secouriste en uniforme extrayant des décombres un homme en sous vêtements.</p> <p><i>comme cet homme, surpris en plein sommeil, qui seulement une fois sorti des décombres comprend</i> [voix mélangées, gravats (faible)]</p>	


Annexe 13- Tableau récapitulatif du reportage de TF1.

10	7,36	<p>Plan sur le même homme. Mouvement latéral de gauche à droite puis léger zoom, puis plan fixe. Les civils s'affairent autour de lui, il pleure dans les bras d'un homme. <i>qu'il a frôlé la mort.</i> [pleurs et cris de l'homme + bruits habituels]</p>	
11	6,2	<p>Plan fixe, femme interviewée avec des béquilles <i>-Cette nuit pour moi, c'était la fin du monde, tout volait autour de moi.</i> [voix féminine non traduite]</p>	
12	4,2	<p>Plan fixe, autre femme interviewée. <i>-Ce que j'ai vécu, je crois que je ne pourrai jamais le raconter.</i> [voix féminine non traduite]</p>	
13	3,2	<p>Un homme saute d'un muret dans une rue, mouvement latéral gauche droite de la caméra et descendant pour filmer un gouffre au milieu de cette rue, avec un scooter dans le gouffre et une voiture au bord. <i>Le même cauchemar a surpris des dizaines de milliers</i> [bruit de vent, de moteur]</p>	

Annexe 13- Tableau récapitulatif du reportage de TF1.

14	1,76	<p>Mouvement latéral de la droite vers la gauche dans une rue sans décombres, des victimes ont des couvertures sur eux. <i>d'habitants de l'Aquila et des villages alentours.</i> [klaxon, bruits de circulation]</p>	
15	2,28	<p>Plan fixe, accueil d'une civière dans un centre apparemment hospitalier (personnel en uniforme). <i>Les blessés se comptent par centaines.</i> [diverses sirènes]</p>	
16	2,12	<p>Plan fixe sur une ambulance feux allumés, un ambulancier court vers elle. <i>L'hôpital de la ville</i> [sirène de l'ambulance]</p>	
17	2,84	<p>Mouvement latéral gauche droite de la caméra pour suivre un groupe de personnes aidant un blessé à marcher dans la pénombre <i>touché lui aussi,</i> [voix mélangées, diverses sirènes]</p>	





Annexe 13- Tableau récapitulatif du reportage de TF1.

18	2,24	<p>Plan fixe. Personne se faisant soignée allongée sur des chaises par une professionnelle. <i>Est débordé</i> [voix mélangées]</p>	
19	8,28	<p>Interview d'une femme assise qui a des pansements sur le nez. Léger zoom au milieu de l'interview. <i>-Ma mère n'a qu'une fracture à la main et des points sur la tête ; mon mari, lui, s'en est moins bien sorti, mais c'est lui qui a réussi à nous libérer des ruines.</i> [voix féminine non traduite]</p>	
20	5,52	<p>Interview d'une femme, mouvement latéral gauche droite léger, puis zoom léger sur la femme. <i>-C'est terrible, tout est détruit, on ne sait plus quoi faire.</i> [voix féminine non traduite]</p>	
21	3	<p>La caméra bouge un peu, un homme apparaît à terre sur une civière au milieu de personnes lui portant secours. <i>Toute la journée, les sauvetages se poursuivent,</i> [voix mélangées, cris]</p>	

Annexe 13- Tableau récapitulatif du reportage de TF1.

22	3,96	<p>Zoom avant sur un blessé à terre sur une civière, entouré de secouristes en uniformes <i>dans le centre ville de l'Aquila, dévasté, une course contre la montre s'organise.</i> [voix mélangées, gravats]</p>	
23	2,12	<p>Mouvement latéral de gauche à droite qui suit d'abord un berger allemand sur des décombres puis stoppe sur un groupe de secouristes en uniformes. <i>Il faut tenter de porter secours</i> [bruits habituels]</p>	
24	2,5	<p>Mouvement descendant de la droite vers la gauche, des civils déblayent à la pelle les décombres <i>à ceux qui sont prisonniers, parfois sous</i> [sirènes d'ambulances]</p>	
25	2,08	<p>Plan fixe sur un tas de gravats devant un immeuble avec des secouristes. <i>plusieurs tonnes de gravats.</i> [gravats]</p>	

Annexe 13- Tableau récapitulatif du reportage de TF1.

26	7,44	<p>Interview d'un secouriste avec uniforme en plan fixe. <i>-Là, on a repéré une jeune femme, mais on n'arrive pas à la localiser avec précision on l'aperçoit de loin par un tout petit trou.</i> [voix masculine non traduite]</p>	
27	9,72	<p>Interview d'un autre secouriste avec uniforme en plan fixe. <i>-On creuse, on creuse avec nos mains pour l'instant, sinon, avec des moyens plus lourds, ça pourrait être dangereux pour les victimes.</i> [voix masculine non traduite]</p>	
28	2,4	<p>Mouvement latéral de la gauche vers la droite filmant deux hommes qui portent une civière. <i>Au fil des heures, le bilan s'alourdit</i> [voix mélangées]</p>	
30	1,84	<p>Deux personnes se prennent dans les bras, plan fixe <i>de manière dramatique. Cent cinquante</i> [voix mélangées]</p>	

Annexe 13- Tableau récapitulatif du reportage de TF1.

31	3,36	Zoom avant sur des secouristes en uniformes dégageant des gravats. <i>morts au moins sont à dénombrer, bilan provisoire, comme le confirme</i> [voix mélangées + gravats]	
32	2,88	Bas du corps de personnes en uniforme, à l'arrière plan on aperçoit trois corps enveloppés de linges. <i>le responsable de la protection civile italienne aux côtés d'un Silvio Berlusconi</i> [voix mélangées + gravats]	
33	2,16	Plan fixe, les mêmes personnes en uniforme autour d'une ambulance. [voix mélangées + gravats + radio]	
34	3,12	Léger zoom arrière sur une table avec le chef de gouvernement italien et divers hommes publics. On aperçoit en arrière plan les drapeaux italien et européen. <i>qui a décrété l'état d'urgence.</i> [Voix masculine non traduite]	





Annexe 13- Tableau récapitulatif du reportage de TF1.

35	6,56	<p>Plan fixe sur un homme d'État italien. <i>-On aura des chiffres définitifs dans 24 heures. Mais ils vont probablement augmenter d'ici là.</i> [voix masculine non traduite]</p>	
36	1,8	<p>Mouvement latéral droite gauche suivant un homme emportant des effets. <i>Dés l'aube,</i> [sirènes]</p>	
37	1,84	<p>Plan fixe sur deux jeunes filles enveloppées d'une couverture <i>ceux qui l'ont pu ont quitté les quartiers</i> [moteurs]</p>	
38	1,8	<p>Zoom sur des blessés secourus <i>dévastés, laissant tous leurs biens à l'abandon.</i> [sirène d'un engin de manutention]</p>	

Annexe 13- Tableau récapitulatif du reportage de TF1.

39	2,8	Secouriste dans une rue marchant avec une échelle, mouvement de caméra latéral droite gauche [voix italienne masculine]	
40	3,27	Léger mouvement latéral droite gauche, puis zoom sur un secouriste. <i>Dans le centre historique de la capitale des Abruzzes, les secouristes</i> [voix masculine non traduite]	
41	8,17	Mouvement descendant de la caméra pour suivre une femme secourue. <i>ont continué à chercher les rescapés qui n'ont pas osé descendre leurs escaliers, tels cette femme, encore sous le choc plusieurs heures après le séisme.</i> [voix mélangées, cris]	
42	2,52	Zoom avant sur une tente et des secouristes en uniformes. <i>Les uns et les autres vont rejoindre les camps de</i>	

Annexe 13- Tableau récapitulatif du reportage de TF1.

43	2,04	Zoom avant sur d'autres secouristes un uniforme montant une tente. <i>tentes qui ont commencé à se monter alors que</i>	
44	3,48	Mouvement rotatif de la caméra sur une route coupée en deux. <i>des répliques sismiques ont eu lieu tout au long de la journée.</i>	
45	5,92	Vue aérienne de l'Aquila, la caméra bouge en suivant le mouvement de l'hélicoptère. <i>Les dégâts sont considérables dans toute cette partie des Abruzzes, au cœur de cette cité historique où de nombreuses églises,</i> [bruit d'hélicoptère]	 <p data-bbox="927 1272 1385 1339">TF1 COMMENTAIRE : MICHEL SCOTT MONTAGE : C. MARCHINA</p>
46	3,52	Même chose, mais d'une autre ville alentours. <i>et des joyaux architecturaux ont également été détruits.</i> [bruit d'hélicoptère]	 <p data-bbox="927 1621 1385 1682">TF1 COMMENTAIRE : MICHEL SCOTT MONTAGE : C. MARCHINA</p>

Annexe 13- Tableau récapitulatif du reportage de TF1.

47	1,48	Même chose. [bruit d'hélicoptère]	
----	------	--------------------------------------	--

Annexe 14 : Textes des reportages télévisés de TF1 concernant l'Italie en 2007.

A- Reportage du 20/02/2007 :

<div3 type="intro">
<u who="#roselmack_harry">

Et l'un de ces carnivals qui fait **rêver** : celui de **Venise**. On connaît les défilés de masques sur la **place St Marc**, mais plus confidentielles, ces **somptueuses soirées** dans les **palais** de la ville auxquelles Maurice, Maurice Olivari et Florence Lozach ont pu assister.

</u>

</div3>
<div3 type="report">
<u who="#olivari_maurice">

Pour donner le ton d'une soirée, rien ne vaut une étape au Florian. Le plus **authentique** et le plus **ancien** café de Venise, c'est ici en général que se donnent rendez-vous les **historiques**. Ceux qui ont choisi les **XVII et XVIIIè siècles** comme parure. Ce soir, c'est dans un palais qui borde le grand canal que se donne la **fête**. Alors, la plupart des invités s'y rendent en taxi. La **soirée** aura un thème. L'Espagne. C'est dire si les **costumes de lumière** et toutes ses déclinaisons vont pouvoir s'en donner à cœur joie. En fait, ce sera un repas de 250 couverts, dans **l'un des plus grands et des plus beaux palais de Venise**. Les invités sont accueillis par l'hôtesse qui n'est pas propriétaire du palais, mais qui organise depuis des années cette **soirée**.

</u>

<u who="#quidam8">

Les invités sont toutes des personnes qui apportent, en plus d'un **très beau costume, du style et de l'élégance** à ces **soirées**.

</u>

<u who="#olivari_maurice">

Ici, le menu n'a pas beaucoup d'importance, l'**esthétique** des lieux, 4 salles **pur dix huitième**, les **parures, les coiffures, les décolletés, les gestes, les ports de têtes, et les sourires entendus** sont plus importants que tout le reste.

</u>

<u who="#quidam9">

On doit garder un attitude **élégante, raffinée**, jusqu'au bout, même si l'on souffre parfois ! On a trop chaud, les perruques sont trop lourdes, mais, on doit garder, on joue notre rôle jusqu'au bout.

Annexe 14 : Textes des reportages télévisés de TF1 concernant l'Italie en 2007.

</u>

<u who="#olivari_maurice">

Et pendant ce temps là, là-bas, au-delà du **grand canal**, sur la **place St Marc**, d'autres masques continuent la parade. Ce sera d'ailleurs la dernière de ce carnaval. Et dire qu'**il y a deux siècles et demi**, la mascarade à Venise durait **six mois par an**.

</u>

</div3>

B- Reportage du 13/03/2007.

<div3 type="intro">

<u who="#poivre-d'arvor_patrick">

Restons dans ces questions de société, en Italie, le pape Benoît seize a réaffirmé aujourd'hui le caractère obligatoire du célibat pour les prêtres, un rappel à la doctrine qui met un terme au changement souhaité par certains au sein même de l'Eglise. Reportage de Maurice Olivari et Florence Lozach.

</u>

</div3>

<div3 type="report">

<u who="#olivari_maurice">

Pour beaucoup de non catholiques, c'est un peu comme si l'Eglise venait soudain de **refermer la porte**. Et presque deux ans après son élection, Benoît seize apparaît à leurs yeux fidèle au cardinal Ratzinger qu'il était, **rigide** sur la doctrine, **ferme** sur la morale, et **conservateur** sur les mœurs. Premier **rappel à l'ordre**, les prêtres, dont le célibat reste **obligatoire**. Célibat que les intéressés sont invités à vivre avec **maturité, joie et dévouement**. **Deuxième interdit**, le divorce, ou plutôt, le droit pour les remariés d'accéder au sacrement de la communion, à moins, dit le pape, que ces **remariés ne s'engagent à vivre comme frère et soeur**. Troisième rafale, **l'avortement, l'euthanasie, et l'union homosexuelle**, là encore, la **condamnation** est **ferme** et les changements ne sont **pas négociables**. C'est un message qui **s'adresse avant tout aux hommes politiques tentés par certaines réformes**. Et voilà l'Eglise une nouvelle fois à **contre courant** de la sensibilité ambiante, certains diront qu'en fait, le pape ne fait que rappeler **le dogme**, sauf qu'ici le dogme, mis à part peut-être la question de l'euthanasie, **n'a rien à voir à l'affaire**.

</u>

</div3>

C- Reportage du 26/03/2007.

<div3 type="intro">

<u who="#poivre-d'arvor_patrick">

Au début de la vie, à présent, l'Italie est en train de renouer avec les antiques boîtes à bébés, c'est un système qui permet d'éviter aux mères en détresse d'abandonner leur nouveau-né dans la rue. On en compte désormais une dizaine à travers tout le pays. Reportage de Maurice Olivari et Florence Lozach.

</u>

</div3>

<div3 type="report">

<u who="#olivari_maurice">

Ca ressemble à une grande boîte aux lettres, réservée à l'usage, réservée à l'usage exclusif des cigognes en perdition, celle-ci, ouverte il y a deux siècles dans le mur d'un couvent a été la première en Italie à reprendre du service. Le principe est immuable même si le confort est moderne. Une trappe qui garantit l'anonymat, un berceau confortable, et un signal d'alarme pour annoncer la livraison. Et voici la dernière née de ces boîtes à bébé en plein cœur de Rome, ultra moderne et polyglotte en fonction de la population immigrée, y compris de langue française, qui peuple le quartier.

</u>

<u who="#paolilio_michel">

En trois ans, sur trente bébés abandonnés dans le quartier, l'un d'eux est mort de froid sur le trottoir, et l'autre a été sauvé de justesse.

</u>

<u who="#olivari_maurice">

Et cette maisonnette ne sauverait la vie que d'un seul enfant qu'elle serait déjà très utile. Pour plus d'efficacité, l'installation est reliée vingt quatre heures sur vingt quatre par vidéo au service des urgences de l'hôpital.

</u>

<u who="#vitali_stéfano">

La caméra n'est pointée que sur le berceau, et on a installé ici une couveuse pour pouvoir transporter l'enfant vers le service des urgences néo-natales.

</u>

<u who="#olivari_maurice">

Et trois semaines seulement après son ouverture, cette boîte à bébé ultra moderne allait prouver son efficacité. Il était ce soir là vingt et une heure trente lorsque sur l'écran de contrôle, le docteur Vitali a aperçu une paire de mains anonymes déposer délicatement un petit paquet

Annexe 14 : Textes des reportages télévisés de TF1 concernant l'Italie en 2007.

bien enveloppé dans le berceau.

</u>

<u who="#vitali_stéfano">

J'ai couru avec mes instruments de réanimation, mais l'enfant était en bonne santé, sa maman s'en était bien occupé.

</u>

<u who="#olivari_maurice">

L'enfant a été baptisé Stefano, du même prénom que son médecin favori, et il est en cours d'adoption.

</u>

</div3>

D- Reportages du 05/04/2007.

<div3 type="intro">

<u who="#poivre-d'arvor_patrick">

Autre trésor, les oeuvres d'art, il est souvent difficile de les transférer d'un pays à un autre, Nicolas Escoulan et Jean-Etienne Mach ont pu assister, et c'est rare, au départ d'une collection de Picasso de Venise vers l'Allemagne.

</u>

</div3>

<div3 type="report">

<u who="#escoulan_nicolas">

Jusqu'à présent, tout va bien. A l'abri du tumulte et du grand canal, l'une des plus exceptionnelles collections d'art au monde s'apprête à vivre ce que tout conservateur redoute par dessus tout, un voyage. Plus de deux cents Picasso période Antibes, dix neuf cent quarante six, dix neuf cent quarante huit, vont quitter le palais Grassi à Venise où ils viennent de passer l'hiver. Des pièces rares comme la joie de vivre que Benoît Dagron, restaurateur de chefs d'œuvre ausculte avant le décrochage.

</u>

<u who="#dagron_benoît">

Bah quand on a un doute, il y a le, la photo numérique, hein.

</u>

<u who="#escoulan_nicolas">

Chaque craquelure, chaque croûte de Ripolin a son importance.

</u>

<u who="#dagron_benoît">

Parce qu'on a une œuvre unique l'or de Lascaux, c'est une œuvre irremplaçable, donc c'est vrai qu'on peut pas perdre une petite élément

Annexe 14 : Textes des reportages télévisés de TF1 concernant l'Italie en 2007.

de matière, c'est irremplaçable.

</u>

<u who="#escoulan_nicolas">

Irremplaçable et source d'angoisse pour les rares musées qui prennent le risque d'emprunter une œuvre. Le moindre accroc peut ternir une réputation à jamais.

</u>

<u who="#andral_jean-louis">

Faut, faut pas se louper comme on dit hein, c'est, c'est, c'est bon, on a pris toutes les précautions qu'on, on estimait devoir prendre, mais le risque zéro il existe pas.

</u>

<u who="#escoulan_nicolas">

Et pour éviter les drames, ce sont toujours les mêmes transporteurs, spécialistes de l'art qui se chargent du décrochage.

</u>

<u who="#dagron_benoît">

Ca s'est passé sans problèmes !

</u>

<u who="#escoulan_nicolas">

Mais le palais Grassi responsable des Picasso pendant les quatre mois d'exposition paie surtout des primes d'assurance exorbitantes.

</u>

<u who="#andral_jean-louis">

Et puis le problème, c'est que comme les valeurs d'assurance euh, ont tendance à être de plus en plus importantes, pour les, pour les établissements qui empruntent, ça entraîne des frais aussi euh, très conséquents.

</u>

<u who="#escoulan_nicolas">

Rien que pour cette exposition, plusieurs centaines de milliers d'euros. Venise s'endort, les Picasso emmaillotés vont passer leur dernière nuit au calme, en sécurité. Car le lendemain matin, la collection va connaître l'instant le plus délicat du transfert. Tout se passe désormais sur le quai, le long du grand canal, à l'extérieur, donc, sous haute surveillance.

</u>

<u who="#andral_jean-louis">

On attend pas de gallions de pirates mais euh, on fait attention.

</u>

Annexe 14 : Textes des reportages télévisés de TF1 concernant l'Italie en 2007.

<u who="#escoulan_nicolas">

Deux cents caméras scrutent le palais sous toutes ses moulures et dans ses moindres secrets, à sa manière, la police flottante maintient un cordon de sécurité, et des guetteurs que l'on ne verra pas observent à distance cet étrange ballet aquatique. Chaque minute compte, il faut passer le moins de temps possible à l'extérieur de l'enceinte.

</u>

<u who="#andral_jean-louis">

En général, c'est encore plus anonyme parce que vous n'êtes pas là, il n'y a pas une équipe de télé qui tourne, là, on attire un peu l'attention. Mais sinon, c'est vrai que plus c'est discret et mieux c'est.

</u>

<u who="#escoulan_nicolas">

Sur ce grand canal que Picasso n'a jamais immortalisé, et au milieu de touristes ignorant tout des dizaines de millions d'euros qui leur passent sous les yeux, la collection rejoint le port de Venise. Chargement express, cette fois dans des camions presque anonymes, les chefs d'oeuvres seront suivis par satellites jusqu'à leur berceau, au musée d'Antibes, ainsi qu'en Allemagne, pour une nouvelle exposition.²

</u>

</div3>

<div3 type="intro">

<u who="#poivre-d'arvor_patrick">

De Venise, nous passons à **Rome** où en ce jeudi saint, le pape Benoît seize a procédé en fin d'après midi à la cérémonie du lavement des pieds qui évoque les derniers moments de... du Christ avec ses, les disciples.

Reportage de Maurice Olivari et Florence Lozach.

</u>

</div3>

<div3 type="report">

<u who="#olivari_maurice">

C'est un acte symbolique d'humilité que s'imposent depuis des siècles les successeurs de Saint Pierre. Le pape est le chef de l'Eglise, et il se devait de faire au moins autant que celui qu'il représente sur terre, le Christ, qui selon la tradition, la veille de sa crucifixion a lavé les pieds de ses disciples. Alors chaque jeudi saint, douze prêtres de **Rome** sont choisis pour incarner ces douze apôtres et exposer au pape l'un de leur deux pieds. Lavement effectivement très symbolique, quelques gouttes d'eau, une serviette qui effleure pour essuyer. Autrefois, Jean-Paul deux aposait même un baiser sur le pied ainsi lavé, pour mieux souligner l'acte de modestie qui s'en dégage. Rien à voir en tout cas avec la cérémonie de ce matin, elle aussi pourtant intimement liée à la fête de Pâques. Cette fois, le pape, en tant qu'évêque de **Rome**, comme tous les autres évêques

Annexe 14 : Textes des reportages télévisés de TF1 concernant l'Italie en 2007.

à travers le monde, consacre les huiles qui serviront tout au long de l'année à administrer les sacrements. Baptêmes, confirmations, ordination des prêtres, bénédiction des malades, et même, d'où son nom, au moment de la mort, l'extrême onction. Et demain soir, vendredi saint, Benoît seize parcourra le chemin de croix du **collisée**.

</u>

</div3>

Bibliographie.

Aarts J. (1991) *Intuition-based and Observation-based Grammars* in Aijmer et al. (eds) London : Longman. 44-62.

Aitchinson J., (1987), *Words in the Mind*, Oxford : Blackwell.

Alonso A. (2009), *Características del lexico del medio ambiente en español y pautas de representacion en el diccionario general*, Universidad Pompeu Fabra.

Armstrong S.L., Gleitman L.R. & Gleitman H. (1983), « What Some Concepts Might Not Be », *Cognition. International Journal of Cognitive Psychology*. Paris 13 : 263-308.

Arnheim R. (1973). *Vers une psychologie de l'art*. Paris : Segehrs

Atkins B.T.S., Clear J. & Osler N., (1992) *Corpus Design Criteria. Literary and Linguistic Computing* Journal of the Association for Literary and Linguistic Computing 7.1 : 1-16.

Aumont J., Bergala A, Marie M., Vernet M., (1983) *Esthétique du film*. Paris : Nathan.

Aumont J., Marie M., (1988) *L'analyse des films*. Paris : Nathan.

Banks D., (2005), *Introduction à la linguistique systémique et fonctionnelle*. Paris : L'harmattan.

Baker P. (ed.) (2009), *Contemporary Approaches to Corpus Linguistics*. London : Continuum.

Baker P., Gabrielatos C. Krzyzanowski M., McEnery T. & Wodak R. (2008) A Useful methodological synergy ? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press., *Discourses and Society* 19(3) : 273-306

Baude O., (2006), *Corpus Oraux : Guide des bonnes pratiques*. Orléans : CNRS Éditions et Presses Universitaires d'Orléans.

Barnbrook G. (2007), Sinclair on collocation, in *International journal of Corpus Linguistics*, Amsterdam/ Philadelphia.

Benveniste E. (1966), *Problèmes de linguistique générale*. Tome 1. Paris : Gallimard.

Biber D. (1998), *Corpus Linguistics : Investigating Language Structure and Use*, Cambridge University Press. Cambridge

Bloomfield L. (1933) *Language*, éd. Henry Holt, New York.

Bouquet S. (2005), Après un siècle, les manuscrits de Saussure reviennent bouleverser la linguistique. *Texte*, Juin 2005. Disponible sur http://www.revue-texto.net/Saussure/Sur_Saussure/Bouquet_Apres.html . Consultée le 10/08/09

Burnard L., (2007), Une introduction au British National Corpus dans son édition xml, *Journées de la Linguistique de Corpus*, Lorient.

Brown G. & Yule G. (1983) *Discourse Analysis*. Cambridge : Cambridge University Press.

Catford J.C., (1983), « Insects are free »: reflections on meaning in linguistics. In *Language Learning*, vol. 33 n°5.

Chomsky N., (1957), *Syntactic Structures*, The Hague ; Paris : Mouton.

Chomsky N., (1962), Paper in 3rd Texas Conference on Problems of Linguistics Analysis in English, Austin, University of Texas.

Chomsky N., (1965) *Aspects of the Theory of Syntax*, Cambridge : MA:MIT Press.

Crystal D. (1985), *Linguistics, second edition*, London : Penguin Language and linguistics.

De Cillia R., (1999), *The discursive construction of national identities*. Edinburgh : Edinburgh University Press.

Déjean H. et Gaussier E. (2002) Une nouvelle approche à l'extraction de lexiques bilingues à partir de corpus comparables. In *Lexicometrica, Electronic journal, Special issue on aligned corpora*, 1-22 (<http://www.cavi.univ-paris3/Lexicometrica/thema/thema6/Dejean.pdf>)

Dugalès N. (2008) « Une Europe introuvable ? Constitution et analyse d'un corpus de presse quotidienne française dans le cadre du projet de recherche européen *IntUne* », *Texte et corpus*, n°3, 2008 (disponible sur <http://web.univ-ubs.fr/corpus/publi.html>).

Eco U. (1985) *Lector in fabula*. Paris : Grasset.

Escal F. (1979) *Espaces sociaux, espaces musicaux*. Paris : Payot.

Firth J. R., (1957) *Papers in Linguistics 1934-1951*, Oxford : Oxford University Press.

Francis G., (1982) Problems of Assembling and Computerizing Large Corpora. In Johansson S. ed.) *Computer Corpora in English Language Research* 7-24.

Goldsmith J., (2004). The legacy of Zellig Harris, review article. *Language*, vol 81 n°3, 2005. Consulté sur : <http://hum.uchicago.edu/~jagoldsm/Papers/ZelligHarris.pdf> le 04/09/2009.

Grice P. (1975) « Logic and Conversation », in Cole, P. and Morgan, J. (eds) *Syntax and Semantics*, vol.3, New York : Academic Press

Hanks P., (1986) Typicality and Meaning potentials. *Euralex Internation Congress*. Zurich. 37-47

Hanks P., (1994) "Linguistic Norms and pragmatic exploitations, or why lexicographers need prototype theory and vice versa" in F. Kiefer, G. Kiss, and J. Pajzs (eds.), *Papers in Computational Lexicography : Complex 94*, Research Institute for Linguistics, Hungarian Academy of Sciences.

Hanks P., (2000) « Do Words Meanings exist ? » in *Computers and the humanities*, n°34-2000, London : Kluwers Academic Publishers.

Hanks P., (2006) 'Metaphoricity is a Gradable' in A. Stefanowitsch and S. Gries (eds.) *Corpora in Cognitive Linguistics. Vol. 1: Metaphor and Metonymy*. Berlin and New York: Mouton de Gruyter

Higgins M., (2004) Putting the nation in the news : the role of location formulation in a selection of Scottish Newspapers, *Discourse and Societies*. London : Thousand Oaks, SAGE publications, CA and New Dehli.

Hoey M., (2005) Lexical Primings and the properties of text. In Partington A., J. Morley & L. Haarman (eds) *Corpora and discourse*. Bern: Peter Lang, 385-412.

Hoey M.(1983), Persuasive Rhetorics in Linguistics : A stylistic study of some features of the language of Noam Chomsky. In Hunston and Thompson (eds) *Evaluation in Texts : Authorial stance and the construction of Discourse*. Oxford : Oxford University Press. 28-37.

Holec H., (1990) Qu'est-ce qu'apprendre à apprendre, in *Mélanges pédagogiques*, 1990.

Honeybone P., (2005) *Key thinkers in linguistics and the Philosophy of language*, Edinburgh : Edinburgh University Press. 80-86.

Hudson R. A., (1980) *Sociolinguistics*. Cambridge : Cambridge University Press.

Hunston S., (1993) “Professional Conflict — Disagreement in Academic Discourse”. In *Text and Technology*, Baker, Mona, Gill Francis and Elena Tognini-Bonelli (eds.) Amsterdam : John Benjamins..

Hunston S., (2007). Semantic prosodies revisited, *International journal of corpus linguistics*, ISSN 1384-6655, Vol. 12, N° 2, 2007 , 249-268

Jacques M.-P. (2005) Pourquoi une linguistique de corpus ? In *La linguistique de corpus*, G. Williams, Rennes : Presses Universitaires de Rennes. 21-31

Jakobson R. (1962) *Selected Writings*, I. The Hague; Paris : Mouton Publishers.

Jakobson R. (1973). *Questions de poétique*. Paris : Seuil

Joly M. (1994) *L'image et le sens*. Paris : Nathan.

Kennedy G., (1998) *An Introduction to corpus Linguistics*, Studies in Language linguistics. London : Longman.

Kenny D. (2000) Translators at play : exploitations of collocational norms in German-English translation. In B. Dodd (ed.) *Working with German Corpora*. Birmingham : University of Birmingham Press, 143-160.

Kerbrat-Orecchioni C. (1980). *L'énonciation. De la subjectivité dans le langage*. Paris : Armand Colin.

Kleiber G., (1990) *La sémantique du prototype*, Collection Linguistique nouvelle. Paris : Presses Universitaires de Paris.

Langelier R. E. (2006) L'influence des médias électroniques sur la formation de l'opinion publique : du mythe à la réalité. In *Lex electronica*, volume 11 n° 1; consulté sur le site http://www.lex-electronica.org/docs/articles_56.pdf le 11/09/2009.

Lakoff G. (1973) Hedges and meaning criteria. In Mc David and Audrey R. Duckert (eds), *Lexicography in English, Annals of the New York Academy of sciences*, 144-153.

Lakoff G., (1987), *Women, Fire and dangerous things : What categories reveal about the Mind*. Chicago : University of Chicago Press.

Leech G. (2002) Corpora. In K. Malmkjaer (ed.), *The Linguistics Encyclopaedia*, London : Routledge, 84-93

Leech G. (2004) Developing linguistic corpora: A guide to good practice adding linguistic annotations, consulté sur la page web suivante : <http://ahds.ac.uk/creating/guides/linguistic-corpora/index.htm> le 31/08/2009.

Léon J., (2007) « Empiricism versus Rationalism revisited. Current Corpus Linguistics and Chomsky's arguments against corpus, statistics and probabilities in the 1950-1960s », *Linguistische und epistemologische Konzepte - diachron*, eds. S. Matthaios & P. Schmitter, Münster, Nodus Publikationen 157-176.

Levinson S. C., (1983), *Pragmatics*, Cambridge : Cambridge University Press.

Lévi-Strauss C. (1964) *Le cru et le cuit*. Paris : Plon.

Louw B., (1993) 'Irony in text or insincerity in the writer? The diagnostic potential of semantic prosodies' in M. Baker, G. Francis, and Tognini-Bonelli (eds), *Text and Technology*. Amsterdam : John Benjamins.

Maingueneau D. (1981) *L'énonciation en linguistique française*. Paris : Hachette.

Malinowski B., (1968) *Une théorie scientifique de la culture*, Paris : Point Essais.

Mauntner G. (2009). Corpora and Critical Discourse Analysis. In Paul Baker (ed.) *Contemporary corpus linguistics*, London 32-47.

McEnery T. & Wilson A. (1996) *Corpus Linguistics*, Edinburgh : Edinburgh University Press.

Moon R., (2007) “Sinclair, lexicography, and the COBUILD project : The application of theory” , in *International Journal of corpus linguistics*. Amsterdam : John Benjamins.

Partington A., (2006) “Metaphors, motifs and similes across discourse types : Corpus-Assisted Discourse Studies (CADS) at work”, *Trends in Linguistic and Monographs*, vol. 171, Netherlands : Mouton de Gruyter, 267-304.

Pearson J. (2003) Using parallel texts in the translator environment. In F. Zanetti, S. Bernardini and D. Stewart (eds.), *Corpora in Translation Education*. Manchester : St Jerome, 15-24.

Petitclerc A. et Schepens P. (2009) Critical Discourse Analysis I., Les notions de contexte et d'acteurs sociaux, in *SEMEN, revue de sémio-linguistique des textes et discours n° 27*. Besançon : Presses Universitaires de Franche-Comté.

Rastier F. (2005a) Enjeux épistémologiques de la linguistique de corpus, in *La linguistique de corpus*, G. Williams. Rennes : Presses Universitaires de Rennes. 31-47.

Rastier F. (2005b) Saussure au futur : écrits retrouvés et nouvelles réceptions. *Texte !* mars 2005 [en ligne]. Disponible sur : <http://www.revue-texto.net/Saussure/Saussure.html> (consultée le 10/08/09)

Retzlaff S. and Gänzle S., (2008) Constructing the European Union In Canadian News, in *Critical Approaches to Discourse Analysis across Disciplines*, <http://cadaad.org/ejournal> Vol 2(2): 67-89 ISSN : 1752-3079

Rosch E., (1975), Cognitive representation of semantic categories, in *Journal of experimental Psychology : General*, 104, 192-233.

Saussure F., (1916) *Cours de linguistique générale*. Paris : Payot.

Sinclair J. (1987) *Looking Up*, Oxford : Collins ELT

Sinclair J., (1991) *Corpus, Concordance, Collocation*. Oxford : Oxford University Press.

Sinclair J., et Renouf A., (1988) « A Lexical Syllabus for Language Learning ». In *Vocabulary and Language Teaching*. Carter and McCarthy (eds). Harlow: Longman 140-158.

Sinclair J., (2003) *Reading Concordances*, London : Longman.

Sinclair J., (2004) *English Collocation Studies, The Osts Report*. Continuum International Publishing Group.

Sinclair J. (2005) Meaning in the framework of corpus linguistics. In Teubert (ed.) *Lexicographica*, Tubingen : Max Niemeyer 20-32.

Smith R. C., (1999) *The writings of Harold Palmer, an overview*, hon-no Tomosha (ed), Tokyo.

Stubbs M., (1993) British Traditions in Text Analysis in Baker et al. (eds) 1-33.

Stubbs M., (1995) Collocations and semantic profiles : on the cause of the trouble with quantitative studies. In *Functions of Language*, 2:1 23-55.

Stubbs M., (1996) *Text and Corpus Analysis : Computer Assisted Studies of Language and Culture*. Oxford : Blackwell.

Swales J. M., (1990), *Genre Analysis, English in academic and research settings*. Cambridge : Cambridge University Press.

Taylor J., (1995) *Linguistic Categorization*, Oxford : Oxford University Press.

Tognini-Bonelli E., (2001) *Corpus Linguistics at work, Studies in Corpus Linguistics*, Amsterdam : John Benjamins.

Teubert W. (2001) "A Province of a Federal Superstate, Ruled by an unlected Bureaucracy - keywords of the Euro-Sceptic Discourse in Britain" in *Attitudes towards Europe : language in the unification process*. Eds Andreas Musloff, Colin Good, Petra Points, Ruth Wittlinger, Ashgate, Aldershot.

Teubert W. (2007) Sinclair, Pattern grammar and the question of hatred, in *International Journal of Corpus Linguistics*. Amsterdam.

Véronique D. (1992) Sweet et Palmer, précurseurs de la didactique des langues ? In *Cahiers Ferdinand de Saussure* n° 46/1992, Genève : Droz 173 à 192.

Viallon P. (1996). *L'Analyse du discours de la télévision*. Paris : Presse Universitaires de

France.

Wierzbicka A., (1995). Universal semantic primitives as a basis for lexical semantics. In *Folia Linguistica* 29 (1-2) 149-169.

Wierzbicka A., (1996) *Semantics Primes and Universals*. Oxford : Oxford University Press.

Williams G. (1998). Collocational Networks : Interlocking Patterns of Lexis in a Corpus of Plant Biology Research Articles, in *International Journal of Corpus Linguistics*.3(1) 151-171

Williams G. (2002). In search of representativity in specialised corpora, in *International Journal of Corpus Linguistics*. 7 (1) 43-63.

Williams G.(2006) La linguistique de corpus, une affaire prépositionnelle, *revue Texto*, consultée le 10/08/09 sur <http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/Williams.pdf>

Williams G., (2008) Sweden, France, Europe : Identity in the Corpus. In *Nog Ordat ? Festschrift till Sven-Göran Malmgren den 25 april 2008*, Götteborg.

Williams G. (2009) Exploring Wider Networks : Identity. *IntUne Paper (Lisbonn)*

Williams G., Piazza R., Giuliani D. (2009, *forthcoming*) Nation and Supernation : a tale of three Europes, in Bayley P. Williams G. (eds) *European Identity, what media says ?* Oxford : Oxford University Press.

Wittgenstein L. (1953) *Philosophical Investigations*. London : Blackwell

Wittgenstein L., (1953/2004) *Recherches Philosophiques*, Paris : Galimard (Titre original, *Philosophische Untersuchungen*, Blackwell publishers, 1953.

Résumé

La linguistique de corpus connaît un intérêt croissant dans le domaine des sciences humaines. Elle est de plus en plus utilisée dans des domaines aussi variés que l'histoire, la sociologie, la psychologie, la linguistique ou la didactique. L'objet corpus est dans cette discipline entendu comme une collection informatisée et de grande taille de textes. Selon l'objectif poursuivi, le corpus utilisé ne sera pas conçu de la même façon. Sa structuration et les textes sélectionnés différeront en fonction des buts de recherche. Pour étudier l'émergence et la construction d'une identité européenne, l'utilisation d'un corpus de presse française semble pertinente

Après avoir expliqué le cadre théorique dans lequel elle se situe, l'étude exploite un corpus de presse française à l'aide des logiciels d'extraction de texte et de concordances Xaira et WordSmith. L'analyse du corpus se veut à la fois quantitative et qualitative. A travers l'extraction de collocations pour différents mots-clefs en rapport avec la thématique étudiée (Europe, Italie, Pologne), des prototypes sémantiques sont construits, puis leur activation dans les médias en 2007 et 2009 est analysée et comparée. Une fois cette analyse menée, il devient possible d'avoir un aperçu de la façon dont les médias français relaient la construction et la coexistence de différentes identités au sein de l'ensemble européen.

Discipline : Sciences du Langage.

Spécialité : Linguistique de corpus.

Mots clés : Corpus, Collocation, Europe, Identité, Lexicographie, Linguistique de corpus, Médias, Nation, Pays.

UFR LSHS.

4 rue Jean Zay.

56 100 Lorient.