

Mario Marcon

Apprendre les genres du secteur wallon de l'énergie. Un regard croisé sur les professionnels et le grand public

Résumé

L'apprentissage des discours de spécialité demande la maîtrise de lexiques, mais aussi la reconnaissance d'une situation de communication, à savoir les conventions textuelles d'une communauté de discours. La linguistique de corpus appliquée à l'étude des genres, a montré sa validité et sa pertinence pour l'étude du lexique et d'autres caractéristiques linguistiques et textuelles de genres différents. Cet article présente la construction du corpus spécialisé ENERWAL, consacré au secteur de l'énergie en Région wallonne et visant deux publics (professionnels vs. grand public). Nous illustrons quelques exemples concernant des caractéristiques morphologiques, syntaxiques, phraséologiques et discursives des genres d'ENERWAL. Chaque exemple présente des suggestions pratiques et des aspects critiques qui concernent ces caractéristiques ainsi que l'emploi d'un corpus spécialisé en vue d'une prise de conscience de la part des apprenants en salle de cours.

Mots-clés: discours de spécialité, genre, linguistique de corpus, secteur de l'énergie, situations de communication.

1. Genre : entre situations de communication et corpus

La notion de *genre* a fait couler beaucoup d'encre. Il sera question de citer ici quelques contributions des milieux anglophone et francophone pour encadrer notre étude.

Swales (1990) et Bhatia (1993) décrivent le genre en fonction de :

1. l'appartenance à une (ou plusieurs) communauté(s) de discours (*discourse community*), à savoir des « *sociorethorical networks that form in order to work towards sets of common goals* » (Swales, 1990: 9) ;
2. la reconnaissance d'une situation de communication sous une forme de discours conventionnalisée ;
3. les intentions et les finalités d'une forme de discours.

Plus récemment, Bhatia (2006: 387) a proposé une synthèse très claire à cet égard :

« Genre is viewed as an instance of language use in a conventionalized social setting requiring an appropriate response to a specific set of communicative goals of a disciplinary or social institution, and thus giving rise to stable structural forms by imposing constraints on the use of lexico-grammatical as well as discursal resources. »

La même attention aux particularités lexico-grammaticales et discursives est partagée par Condamines (2007: 42) :

« [...] le genre, parce qu'il fait intervenir une classification des situations de communication (c'est-à-dire aussi des textes) pourrait servir de palier de description qui autoriserait des classifications de mots (de sens de mots). Ces classifications pourraient être faites à partir d'attestations réelles. C'est bien sûr, donner un rôle très important au genre textuel dont on sait bien qu'il peut être lui-même très fluctuant. Mais il peut jouer un rôle pour comprendre l'articulation entre textes et mots, entre introspection et attestations ».

Le genre textuel fonctionne de guide pour une compréhension des sens des mots. Cette compréhension passe par la mise en relation entre mots et textes, d'une part, et par l'analyse d'exemples concrets d'usage qui renvoient à des situations de communication, d'autre part.

La classification de ces situations de communication ainsi que la reconnaissance de leurs conventions (intratextuelles, intertextuelles et extratextuelles) débouche sur la création de taxinomies des genres. Force est de constater que les finalités qui sous-tendent leur création influencent le(s) résultat(s) d'une étude et des applications éventuelles (Aussenac-Gilles et al., 2002: 295). Ces taxinomies se révèlent partielles et adaptées aux besoins d'une communauté restreinte de professionnels : linguistes, traductologues, professionnels d'un domaine et, bien sûr, enseignants. D'un côté, ce travail d'étiquetage des genres et de catégorisation des situations de communication est maintenant facilitée par la linguistique de corpus (Hunston, 2002 ; Sinclair, 1991). D'un autre côté, l'apprentissage d'un discours de spécialité et de ses genres est de plus en plus supporté par l'emploi de corpus *ad hoc* (Gavioli, 2005). Dans cette perspective, l'enseignant joue un rôle de guide et de filtre pour l'identification des phénomènes linguistiques spécifiques ainsi que pour l'acquisition des techniques de traitement et d'observation des données textuelles (Gavioli, 2005: 29-30). En ce

qui concerne l'étude des genres, la compilation d'un corpus offre un cadre favorable (Gavioli, 2005: 56-60), malgré le fait que :

« a genre [...] and general language are not discrete categories. Even though we collect a huge number of texts representing a genre, frequent characteristics emerging from the analysis of such a corpus may not be characteristic features of that genre [...] »
(Gavioli, 2005: 63).

À ce propos, Gavioli continue :

« it is necessary to compare frequent features derived from that corpus with frequent features of other [...] genres and with those of the language in general »
(ibid.)

Autrement dit, l'identification d'un genre se réalise par la comparaison avec d'autres genres, et plus précisément par la comparaison de phénomènes linguistiques récurrents. L'analyse des lexiques de spécialité et de leurs comportements constitue un point de départ (et un ancrage) pour essayer de dessiner les contours des genres textuels d'un discours de spécialité. En linguistique de corpus, l'extraction des lexiques de spécialité par une comparaison statistique de deux corpus, est une pratique largement diffusée. Notre expérience d'extraction du lexique dans le secteur de l'énergie en Wallonie suit cette ligne d'investigation (Marcon, 2010). Outre à obtenir une liste alphabétique d'unités lexicales, nous avons également décrit des comportements lexicaux spécifiques et/ou similaires aux genres textuels et à des situations de communication professionnelle et non professionnelle, en tenant compte du fait que dans les études récentes des genres :

« emphases on text and context have been almost reversed, with context attracting more serious attention, though without undermining the importance of linguistic form, in the description of specialized genres »
(Gotti & Bhatia, 2006: 10).

Après une présentation de notre corpus ENERWAL (§ 2) et des genres qui le composent (§ 3), nous passons brièvement en revue les techniques employées pour l'extraction du lexique du secteur de l'énergie wallon (§ 4). Le restant de l'article est consacré à la présentation des particularités morphologiques, lexico-grammaticales et discursives liées à la fois aux genres et aux situations de communication du secteur de l'énergie wallon, en esquisant un parcours didactique pour leur enseignement/apprentissage (§§ 5-11).

2. Le corpus ENERWAL

Le corpus ENERWAL (ENERgie WALLonne, 1.171.042 occurrences) regroupe des textes gratuits disponibles aux mois d'avril-mai 2009 sur le site *Portail de l'énergie en Région wallonne* (<http://energie.wallonie.be> > sections *Pratique* > *Astuces au quotidien, Construire, rénover, Mon électricité – mon gaz, Demander conseil, Se documenter*), à savoir le site institutionnel d'interface avec les citoyens du *Département de l'Énergie et du Bâtiment durable* de la Région wallonne. Nous avons ainsi ciblé la communication institutionnelle entre la Région wallonne et ses citoyens. Force est de constater que cet émetteur s'exprime par les études de centres de recherche agréés, de départements universitaires ou d'organismes wallons, à savoir d'une communauté scientifique d'experts dans le secteur de l'énergie.

ENERWAL se compose de deux sous-corpus : ENERWAL-XP (*ENERgie WALLonne eXPerts* – EWXP ; 821.004 occurrences) et ENERWAL-GP (*ENERgie WALLonne Grand Public* – EWGP ; 350.038 occurrences). Comme leurs noms l'indiquent, la distinction de ces sous-corpus se fonde sur la notion de *situation de communication* (*communicative settings*, Pearson, 1998: 35-39), c'est-à-dire de l'émetteur (des émetteurs) et du (des) public(s) expressément visé(s) par les textes : textes destinés aux professionnels pour EWXP, et textes destinés au grand public pour EWGP. Cette distinction est reprise par le *Portail*, qui trie la documentation en deux collections : *Collections pour le grand public* et *Collections pour les professionnels*. En même temps, nous avons tiré parti d'autres suggestions de Pearson (1998: 35-61), et opérationnalisé ses critères.

3. Les genres d'ENERWAL

Nous avons gardé la nomenclature proposée par le *Portail* pour attribuer des étiquettes aux genres d'ENERWAL. Le tableau ci-dessous reprend la liste des genres de notre corpus avec leurs quantités respectives :

brochure	32
cadre de référence	1
étude de cas	11
feuille de calcul	20
fiche	2
guide	16
manuel	2
mode d'emploi	3
page web	43
plan d'action	3
syllabus	6
trimestriel	18
vadémécum	6
TOTAL	163

Composition en genres d'ENERWAL

À première vue, l'équilibre (McEnery et al., 2006: 16) d'ENERWAL est compromis par la quantité des pages web et des brochures qui représentent, à elles seules, environ la moitié des textes de notre corpus.

Si l'on regarde de plus près la distribution des genres dans nos deux sous-corpus :

<i>Genres EWGP</i>	<i>Qt.</i>
brochure	21
fiche	2
guide	1
page web	43
trimestriel	18
	85

<i>Genres EWXP</i>	<i>Qt.</i>
brochure	11
cadre de référence	1
étude de cas	11
feuille de calcul	20
guide	15
manuel	2
mode d'emploi	3
plan d'action	3
syllabus	6
vadémécum	6
	78

Répartition de genres ENERWAL-XP et en ENERWAL_GP

on constate l'emploi (quasi) exclusif de la plupart des genres dans une seule situation de communication. Sur un total de 13 genres pour ENERWAL, 3 genres (*fiche, page web et trimestriel*) caractérisent EWGP, alors que 8 genres (*cadre de référence, étude de cas, feuille de calcul, guide, manuel, mode d'emploi, plan d'action, syllabus et vadémécum*) caractérisent EWXP. La priorité accordée au public visé permet de mieux comprendre la spécificité de certains genres dans le secteur wallon de l'énergie, et de l'explicitier aux apprenants de ce discours de spécialité. De plus, la collection des textes et leur classification peut constituer une activité pour sensibiliser les apprenants vis-à-vis de l'utilisation d'un genre par rapport à une situation de communication.

Deux genres, à savoir la *brochure* et le *guide*, sont partagés par EWXP et EWGP : ils ont le statut de genres « hybrides ». La brochure se prête à satisfaire les finalités les plus disparates. Elle révèle une souplesse telle qu'elle peut être aisément destinée à un grand public ou à des professionnels, encore que les données quantitatives de nos deux sous-corpus témoignent d'une préférence pour le grand public. Quant au guide, de par sa nature encyclopédique et didactique, il remplit les besoins d'apprentissage de (futurs) membres d'une communauté de discours, d'une part, et d'autre part il vulgarise des contenus techniques à des novices. D'ailleurs, il semble qu'en Région wallonne, le guide soit privilégié plutôt en tant qu'outil de formation (15 guides en EWXP) que de vulgarisation (1 seul guide en EWGP).

Quant au genre *bilan énergétique* d'ENERWAL, il a fait l'objet du corpus ENERWAL-BE (*ENERgie WALLonne – Bilans Energétiques – EWBE – 401.397 occurrences*) pour peaufiner le lexique extrait d'ENERWAL (Marcon, 2009: 89-94).

4. Traitement des données

À l'aide d'un simple script Perl, nous avons ajouté une balise pour identifier chaque texte de notre corpus :

```
<EWXP-00045-guide_elements_cahier_de_charge_commander_audit_solaire.txt>
```

Exemple de balise identifiant les textes en ENERWAL-XP et en ENERWAL_GP

Comme le montre l'exemple, une balise indique le sous-corpus d'appartenance (<EWXP>), numérote tout texte (<00045>) et reprend le genre (<guide>) et quelques mots du titre du texte. De cette façon nous avons retracé le contenu informationnel des textes où les concordances sont encadrées. Ce balisage supporte l'enseignant qui souhaite cibler son syllabus en fonction du niveau de langue de ses apprenants.

Pour extraire notre lexique du secteur wallon de l'énergie, nous avons fait recours à la suite *WordSmith Tools* (Scott, 2008), notamment à son programme *KeyWords*. Nous avons calculé les mots graphiques centraux (*mots-clés*) sur base statistique (chi-carré, $p = 0,000001$, fréquence par mot-clé = 3 ; nb max mots-clés = 1.000). La liste des mots de chaque sous-corpus d'ENERWAL a été comparée à la liste des mots du corpus de référence ACTUINT (1.455.549 occurrences), qui est composé d'articles parus en 2003 dans la rubrique 'actualité internationale' du journal belge *Le Soir*. Le choix de ce corpus est motivé par :

1. la représentation d'une même variété diatopique de la langue française (variété belge) ;
2. les résultats d'une étude pilote où le changement de la thématique de rubrique n'a pas joué un rôle déterminant pour l'identification des mots-clés ;
3. la sélection d'un modèle de langue qui représente (pour autant que possible) l'usage commun, d'où l'exclusion de corpus littéraires en tant que corpus de référence ;
4. la comparaison d'un corpus à plusieurs genres (ENERWAL) et d'un corpus à genre unique (ACTUINT, *article de presse*) (§ 12).

Nous avons obtenu deux listes de mots-clés : 947 mots-clés pour EWXP, 933 mots-clés pour EWGP (534 mots-clés communs aux deux sous-corpus). Ces mots-clés ne sont pas lemmatisés, et ce pour faciliter :

- l'identification de comportements similaires ou divergents des mots-clés communs aux sous-corpus ;
- la description des comportements les plus fréquents de chaque mot-clé.

Dans les paragraphes suivants, nous présentons quelques exemples concernant des noms, des adjectifs et des verbes tirés des listes de mots-clés d'EWXP et d'EWGP. Chaque exemple analyse des phénomènes :

- morphologiques, comme la récurrence de certains préfixes et suffixes (§ 5) ;
- lexico-grammaticaux, comme les collocations et les préférences sémantiques (Sinclair 2004: 24-47) (§§ 6, 7 et 8) ;
- syntaxiques, concernant surtout l'emploi des verbes (§§ 9, 10 et 11) ;
- et discursifs, comme la structuration de l'information et la manifestation de l'auteur et de ses intentions (§§ 6, 8, 9, 10 et 11)

des genres textuels du secteur de l'énergie en Wallonie. Tout exemple est conçu comme un parcours d'observation guidée par l'enseignant, et propose des suggestions et des aspects critiques à propos des phénomènes linguistiques traités ainsi que de l'emploi d'un corpus spécialisé et d'un corpus de référence.

5. Affixations

Tout d'abord, l'enseignant peut soumettre les listes de mots-clés aux apprenants et se concentrer, par exemple, sur l'affixation. Il est notoire que de nombreuses unités lexicales des sciences naturelles sont de formation savante à forte composante dénotative. Ces unités lexicales, qui reviennent dans la liste d'EWXP, sont repérées aussi dans la liste d'EWGP, comme celles au préfixe (très productif) *bio-* : *biocarburants, biodiesel, biomasse, biométhanisation*. Ces deux derniers sont partagés par EWXP, tout comme d'autres unités lexicales au préfixe *poly-* : *polyéthylène, polystyrène, polyuréthane*. Par cette constatation, les apprenants peuvent prendre conscience de la perméabilité des frontières entre lexique de spécialité et lexique commun (Depecker, 2002: 63).

En parcourant encore la liste des mots-clés d'EWXP, les adjectifs au suffixe *-ble* sont également nombreux : *disponible, disponibles, durable, habitable, indispensable, navigables, négligeable, perméables, possible, possibles, préférable, réglable, réglables, renouvelables, variable*. Il en va de même pour EWGP, avec quelques superpositions entre les deux listes d'adjectifs : *confortable, disponible, disponibles, durable, possible, perméable, perméables, programmable, renouvelable, renouvelables, rentable*. À ce point, il sera convenable de guider les apprenants pour qu'ils comprennent que la dimension de la possibilité apportée par le suffixe s'explique :

1. par la description des propriétés de quelques matériaux et moyens employés dans le secteur de l'énergie ;
2. par l'évolution de ce secteur, ainsi que par son importance stratégique sur les plans économique et politique pour un avenir éco-compatible ;
3. et par les genres, à savoir par la finalité didactique-prescriptive des manuels et des guides pour EWXP, et des brochures et des trimestriels pour EWGP, qui justifie également les verbes modaux dans les listes de mots-clés (§ 9).

6. Polyéthylène

À ce stade, l'enseignant propose l'analyse de *polyéthylène*, qui est une unité lexicale à formation savante commune aux deux sous-corpus, pour vérifier des comportements similaires et/ou divergents. Dans ce cas, le support d'un concordancier est indispensable [programme *Concord* de *WordSmith Tools* (Scott, 2008), ANTCOINC (Anthony, 2011) ou le concordancier d'UNITEX (Paumier, 2002)], où notre unité lexicale sera le nœud des concordances que les apprenants verront affichées à l'écran. Cette lecture servira à mettre en relief les préférences phraséologiques d'une même unité lexicale, et leurs variations d'un public à l'autre, à savoir d'un ensemble de genres à un autre.

Cette unité compte 25 occurrences en EWXP, 16 en EWGP. En EWXP, il est souvent (13 occurrences) précédé du co-occurent *feuille*, suivi de la préposition *de* :

1. [...] les pare-vapeur dont la résistance μd est comprise entre 5 et 25 m (type E2), tels que : une feuille de polyéthylène (d'au moins 0,2 mm d'épaisseur) ; une feuille d'aluminium plastifié sur une face. (EWXP – guide)

Le co-occurent *film* ne revient qu'une seule fois, dans le nom composé Nom+Préposition+Adjectif *film en relief*, suivi de la préposition *en* :

2. Dans le cas 5 (panneaux préfabriqués), on peut prévoir le même support des feuilles de Zn-Cu-Ti qu'avec le cas 1 ou encore utiliser un film en relief en polyéthylène haute densité [...] (EWXP – guide)

Au contraire, 13 occurrences d'EWGP (sur un total de 16) révèlent le rapport privilégié du co-occurent *film* suivi de la préposition *de*. Pourtant, il faut préciser que 11 occurrences paraissent dans des tableaux ou dans des listes. Seules 2 occurrences attestent l'emploi dans une proposition, les deux dans une seule brochure d'EWGP :

3. La 1ère mesure a eu lieu entre la pose du film de polyéthylène et la pose des panneaux de plâtre enrobé de carton. (EWGP – brochure)
4. Cela s'est traduit en pratique par : [...] la pose d'un film étanche à l'air (film de polyéthylène) sous l'isolant de la toiture. (EWGP – brochure)

Les 3 autres occurrences de *polyéthylène* en EWGP montrent d'autres co-occurents :

5. Les lés de polyéthylène ont été collés entre eux au moyen d'une bande adhésive. (EWGP – brochure).

7. Photovoltaïque

L'enseignant pose l'attention sur les différences entre les deux listes de mots-clés. Par exemple, la liste d'EWGP inclut l'adjectif *photovoltaïque*, alors qu'il est absent de la liste d'EWXP. De cette donnée, on pourrait conclure que *photovoltaïque* serait censé faire partie du discours grand public sur l'énergie. On peut lancer une recherche pour comprendre de quelle façon cet adjectif est employé dans les textes d'EWGP :

6. Lors de l'installation d'un système photovoltaïque, un compteur comptabilise la production d'électricité verte. (EWGP – trimestriel)
7. Preuve que les Wallons sont concernés et motivés par les énergies renouvelables et en particulier l'énergie solaire photovoltaïque, un créneau qui n'avait pas encore été exploité chez nous à sa juste valeur. (EWGP – trimestriel)
8. La durée de vie des panneaux photovoltaïques est de plus de 25 ans et leur bon fonctionnement (80% de la puissance crête initiale) est généralement garanti sur 20-25 ans. (EWGP – trimestriel)

Un regard aux co-occurents sélectionnés parmi les concordances du genre trimestriel, montre que cet adjectif suit un nom [*système photovoltaïque* (6), *panneaux photovoltaïques* (8)], ou un nom composé Nom+Adjectif [*énergie solaire photovoltaïque* (7)].

Presque toutes les concordances encadrent ces noms composés dans un environnement lexical relevant du discours général, qui favorise la compréhension autant des apprenants de la langue que des novices du secteur de l'énergie solaire. Cet exemple montre comment un corpus basé sur la distinction des situations de communication professionnelle et non professionnelle, peut simplifier l'apprentissage du lexique d'un domaine ou d'un secteur de spécialité par la lecture de genres textuels de vulgarisation scientifique. Au préalable d'un cours, à l'aide d'un corpus pareil, l'enseignant aussi peut familiariser avec un domaine/secteur dont il/elle n'est pas censé être un spécialiste, ce qui lui permet de mieux cibler la création d'un syllabus et de guider les apprenants à la découverte de nouvelles connaissances linguistiques et encyclopédiques.

8. Développement durable

L'attention peut s'adresser aux unités lexicales qui relèvent de (sous-)thématiques spécifiques. Par exemple, les listes de mots-clés d'EWXP et d'EWGP se concentrent sur l'économie de l'environnement. Nous soulignons (l'enseignant le fera à notre place) le partage de l'adjectif *durable* à sa forme singulière (§ 5). Malgré cette présence dans les deux listes, la fréquence de cet adjectif varie de la communication

professionnelle (140 occurrences) à la vulgarisation (47 occurrences). Il sera question de réfléchir un instant avec les apprenants sur les tailles de nos sous-corpus. La différence d'occurrences de chaque sous-corpus (§ 2) influence en effet les résultats quantitatifs, ce qui invite à faire attention à toute sorte de généralisation. Cette remarque sera exprimée de la part des apprenants les plus attentifs aux chiffres ou, le cas échéant, proposée par l'enseignant.

Dans les deux sous-corpus, cet adjectif sélectionne deux co-occurents en première position à gauche : *développement* et *bâtiment*. Par exemple, *développement durable* compte 54 occurrences en EWXP, 18 occurrences en EWGP.

9. Simultanément, la conférence de Rio a diffusé le vocable de « développement durable », qui tend à désigner une synthèse des différents buts que l'on vient de distinguer, en donnant un poids particulier au but écologique d'aval et, pour les pays pauvres, au but macroéconomique. (EWXP – brochure)
10. La construction durable est un processus qui intègre les trois aspects du développement durable : • efficacité économique ; • équité sociale et bien être ; • préservation de l'environnement. (EWXP – guide pratique)
11. Et c'est dans ce même souci de promouvoir l'innovation et de valoriser les filières favorables au développement durable que j'ai lancé dès le 1er janvier de cette année le plan SOLWATT [...] (EWGP – brochure)
12. Mais la Région wallonne a voulu que la concurrence s'accompagne également de mesures favorables au développement durable (notamment en soutenant la production d'électricité verte et en promouvant la consommation rationnelle de l'énergie). (EWGP – page web – FAQ)

Dans le discours wallon, il paraît que *développement durable* n'est pas en concurrence avec la forme anglicisée *développement soutenable* (Paissa, 2008), un nom composé qui est absent de nos deux sous-corpus. De plus, l'enseignant constate que la *brochure* [(9) et (11)] satisfait plusieurs finalités communicatives à la fois. La finalité didactique et encyclopédique l'emportent en (9), où l'unité lexicale *développement durable* fait l'objet d'une véritable réflexion métalinguistique. Autrement dit, il est défini (comme le montrent les guillemets et 'vocalbe') par rapport à un événement historique (conférence de Rio de Janeiro en 1992) et par rapport aux enjeux économiques. Au contraire, (11) privilégie la promotion d'un plan régional (plan SOLWATT) qui subordonne et passe sous silence la nature du développement durable. Au plan discursif, l'enseignant pourra également signaler l'utilisation du pronom personnel sujet *je*, qui unit la promotion du plan énergétique pour la citoyenneté wallonne à la promotion politique individuelle.

À l'aide de toutes ces concordances, l'enseignant peut encore proposer une comparaison entre genres et situations de communication en ce qui concerne les modalités de présentation et de traitement d'une même thématique. Pour les concordances d'EWGP, la visée promotionnelle en (11) est mitigée par l'effort d'explicitation en (12), où le développement durable se dévoile par ses effets sur les citoyens, à savoir par les choix politiques de la Région wallonne en la matière (production d'électricité verte et consommation énergétique rationnelle). Tout cela fait écho à (10) en EWXP qui, sous forme de liste, reprend les buts d'une

construction durable qui s'insère dans le développement durable. Pour ce qui en est du genre, contrairement à ce qu'on aurait cru, la page web consacrée à la FAQ en EWGP préfère élaborer des propositions entières, alors que le guide d'EWXP adopte une liste, en faveur d'une schématisation du contenu informationnel. Toutes ces informations peuvent faire l'objet d'un débat ponctuel en classe dirigé par l'enseignant, ou d'un exercice individuel pour chaque apprenant, d'après le niveau de familiarité avec les techniques d'observation en linguistique de corpus. Dans les deux cas, ces informations sensibilisent les apprenants à prendre conscience des spécificités qui concernent l'organisation textuelle de chaque genre dans sa propre situation de communication.

9. Verbes modaux

L'enseignant peut consacrer un volet à part consacré aux verbes. Dans les deux listes de nos sous-corpus, nous observons la présence de verbes modaux à la troisième personne singulière et plurielle au présent de l'indicatif (*doit, doivent, peut, peuvent*). À ce moment, il est nécessaire de réfléchir sur la nature du corpus de référence (§ 4). Face aux articles de presse d'ACTUINT, qui décrivent des faits et des événements, les suggestions et les indications des textes d'EWXP et d'EWGP font que les verbes modaux soient des mots-clés de leurs listes. Dans les cas de *peuvent* et *doivent*, d'ailleurs, ce n'est pas l'absence de notre corpus de référence qui penche en faveur, par exemple, d'EWXP. Toutefois, rien que l'écart de la fréquence brute entre les deux corpus met en évidence une différence d'emploi.

	EWXP	Corpus de référence
<i>peuvent</i>	567	265
<i>doivent</i>	589	318

Fréquences absolues des unités *peuvent* et *doivent* en EWXP et dans notre corpus de référence

La lecture des concordances montre en outre que ces verbes précèdent systématiquement le verbe *être*, et rentrent dans le patron : <verbe modal + ÊTRE + adjectif/participe à valeur adjectivale>. Pourtant, si en EWXP ce patron marque surtout les guides :

13. Dans les grands réseaux, il peut être judicieux de subdiviser l'installation en plusieurs réseaux autonomes. (EWXP – guide)
14. [...] le débit d'air fourni par ce système doit être mesuré conformément à la norme. (EWXP – guide)
15. [...] les horaires d'occupation peuvent être adaptés, diminuant de façon importante le risque de voir des gens [...]. (EWXP – étude de cas)
16. L'audit, les objectifs et clauses URE doivent être revus périodiquement. [...]. (EWXP – guide)

en EWGP, le même patron revient dans plusieurs genres textuels, parfois avec insertion d'adverbes après le verbe *être* :

17. La chaleur peut être en outre réglée dans chaque pièce [...] (EWGP – dossier)
18. Le linge qui doit être repassé peut contenir une certaine humidité. (EWGP – page web)
19. Les granulés peuvent être achetés selon des conditionnements aussi divers que le sac de 15 kg ou le bateau cargo. (EWGP – dossier)
20. De plus, les installations doivent être conformes au Règlement Général des Installations Electriques. (EWGP – brochure)

10. Verbes à l'infinitif

Dans les cent premiers mots-clés d'EWXP, nous trouvons des verbes à l'infinitif, comme *construire* et *calculer*. À l'aide du concordancier, c'est l'occasion pour l'enseignant de poser des questions aux apprenants sur la nature de quelques genres d'EWXP, et en général sur les procédés de constitution (et de nettoyage éventuel) d'un corpus. Les apprenants s'apercevront que le verbe *calculer* apparaît (dans la plupart des cas, à savoir 276 occurrences sur un total de 398) dans les intitulés des diapositives employées comme supports de cours.

21. Calculer le niveau Ew 74 Construire avec l'Energie 2 - Formation - module 1 - Calculer le niveau Ew 75 Construire avec l'Energie 2 - Formation - module [...]. (EWXP – syllabus)

De la même façon, on justifie d'autres verbes à l'infinitif : *installer*, *améliorer*, *ventiler*, *chauffer*. D'une part, leur centralité statistique est fonction de la répétition des intitulés dans certains textes. D'autre part, elle remonte à la finalité didactique-prescriptive des genres, où les listes d'instructions sont assez fréquentes.

22. Guide de pré-faisabilité 27 Installer une Cogénération dans votre Établissement 1ère Phase [...] (EWXP – guide)
- 23.- Revoir l'occupation des locaux selon les périodes de la journée [...]
 - Installer des sas (portes entre parties chauffées et non chauffées). Isoler les conduites de chauffage. (EWXP – guide)

Les verbes à l'infinitif sont également distribués dans la liste de mots-clés d'EWGP. Entre autres, nous observons encore le verbe *installer* :

24. Dans les conditions actuelles de primes octroyées par la Région wallonne, de nombreux ménages ont intérêt à installer un chauffe-eau solaire! (EWGP – brochure)

et pour la première fois le verbe *économiser* :

25. Une amélioration de l'étanchéité à l'air d'une partie importante du parc immobilier est souhaitable, afin d'économiser l'énergie, d'améliorer le confort [...]. (EWGP – brochure)

L'enseignant peut encore interroger les apprenants sur la place des verbes dans les concordances de (22) à (25). En (22) et (23), les verbes à l'infinitif ouvrent des propositions, en position sujet/thème. Au contraire, en (24) et (25) ces verbes sont encadrés dans des propositions et ajoutent des nouvelles informations (rhème). Cet emploi des verbes à l'infinitif décèle des indices préférentiels sur l'agencement de l'information dans deux genres différents, à savoir dans le guide en situation de communication professionnelle (EWXP) et dans la brochure en situation de communication non professionnelle (EWGP).

11. Verbes à l'impératif en EWGP

La préférence d'EWGP pour les verbes à l'infinitif, qui implique souvent des suggestions, fait écho à la préférence pour les verbes à l'impératif. C'est le cas de *consultez* :

26. Pour une vue globale de tous les nouveaux mécanismes de soutien à vos projets de rénovation, construction et amélioration énergétique, consultez la nouvelle brochure « Habitat & Énergie – Les aides et primes de la Région wallonne, édition 2008 ». (EWGP – trimestriel)

Cette préférence relève uniquement du genre *trimestriel*, qui remplit non seulement sa fonction informative, mais se charge aussi d'inviter et de persuader son audience.

12. Conclusions

Un corpus basé sur les situations de communication et sur les publics visés peut être efficace pour distinguer les genres d'un domaine/secteur de spécialité, ainsi que pour mettre en évidence leurs spécificités linguistiques. De la morphologie à la syntaxe, des préférences de co-sélection des unités lexicales à l'agencement de l'information, jusqu'à la détection des fonctions communicatives, la lecture des listes de mots-clés et la lecture des concordances permettent d'enquêter ces genres de plus près et sous plusieurs angles. D'autres regards pourraient être jetés, surtout grâce à la comparaison de notre corpus ENERWAL avec d'autres corpus de référence. On pourrait suggérer de vérifier la reproductibilité de nos observations par la comparaison des genres du secteur wallon de l'énergie avec d'autres corpus à

plusieurs genres. Des analyses plus approfondies entre situations de communication et genres pourront également témoigner de la spécificité des caractéristiques mises en relief par la présente étude.

Dans l'attente de ces études à venir, nos exemples suggèrent quelques applications pratiques destinées à l'enseignement et à l'apprentissage des conventions qui gèrent la rédaction des genres d'une communauté de discours. Le potentiel des corpus spécialisés présentés dans notre article pourrait constituer un point de départ pour des descriptions et des classifications des genres à l'usage des enseignants, ainsi que pour l'enseignement/apprentissage des discours de spécialité.

Références bibliographiques

- Anthony, L. (2011). *AntConc Build 3.2.2.1*. Available: http://www.antlab.sci.waseda.ac.jp/software/README_AntConc3.2.2.1.pdf (25 May 2011).
- Aussenac-Gilles, N., A., Condamines & S., Szulman (2002). Prise en compte de l'application dans la constitution de produits terminologiques. In *Actes des 2es Assises Nationales du GDR I3*, Nancy, décembre 2002. Toulouse: Cépaduès Editions, 289-302. Available: <http://tia.loria.fr/TIA/IMG/pdf/produits-termino.pdf> (11 May 2011).
- Bhatia, V.K. (1993). *Analysing Genre. Language use in professional settings*. Harlow: Longman.
- Bhatia, V.K. & M., Gotti (Eds) (2006). *Explorations in Specialized Genres*. Bern: Peter Lang.
- Condamines, A. (2007). L'interprétation en sémantique de corpus: le cas de la construction de terminologies. *Revue Française de Linguistique Appliquée* 13(1), 39-52.
- Depecker, L. (2002). *Entre signe et concept: éléments de terminologie générale*. Paris: Presses de la Sorbonne Nouvelle.
- Gavioli, L. (2005). *Exploring Corpora for ESP Learning*. Studies in Corpus Linguistics 21. Amsterdam/Philadelphia: John Benjamins Publishing.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Marcon, M. (2009). *Extraction du lexique de l'énergie en Région wallonne. Des expériences sur corpus*. Final dissertation for MA in Linguistics. Louvain-la-Neuve: Université Catholique de Louvain.
- Marcon, M. (2010). Dépouillement terminologique et situations de communication. Une expérience dans le secteur de l'énergie en Wallonie. *Arena Romanistica* 7, 76-103.
- McEnery, T., R., Xiao & Y., Tono (2006). *Corpus-Based Language Studies. An Advanced Resource Book*. Abingdon: Routledge.
- Paissa, P. (2008). Parasyonymes et euphémismes dans la médiatisation de quelques notions de l'économie de l'environnement. In S. Cigada and M. Verna. (Eds). (2008). *La sinonima tra langue e parole nei codici francese e italiano. Atti del convegno dell'Università Cattolica del Sacro Cuore di Milano*, 24-27 ottobre 2007. (pp.551-571). Milano: Vita&Pensiero.

Paumier, S. (2002). *Unitex. Manuel d'utilisation*. Available: <http://igm.univ-mlv.fr/~unitex/UnitexManual3.0.pdf> (28 May 2011).

Pearson, J. (1998). *Terms in Context*. Amsterdam & Philadelphia: John Benjamins Publishing.

Portail de l'énergie en Région wallonne: <http://energie.wallonie.be/> (20 May 2011).

Scott, M. (2008). *WordSmith Tools version 5*. Liverpool: Lexical Analysis Software.

Sinclair, J. McH. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J. McH. (2004). The search for the units of meaning. In J. McH. Sinclair and R. Carter. (Eds.). (2004). *Trust the text. Language, corpus and discourse*. (pp.24-47). Londres and New York: Routledge.

Swales, J.M. (1990). *Genre Analysis. English in academic and research settings*. Cambridge: Cambridge University Press.