



ASp
la revue du GERAS

73 | 2018
Varia

Le corpus comme aide à la rédaction de résumés scientifiques pour des étudiants LANSAD : une approche comparative

Using a corpus for abstract writing in ESP classes: a comparative approach

Laura-May Simard



Édition électronique

URL : <http://journals.openedition.org/asp/5122>

DOI : 10.4000/asp.5122

ISSN : 2108-6354

Éditeur

Groupe d'étude et de recherche en anglais de spécialité

Édition imprimée

Date de publication : 1 mars 2018

Pagination : 75-104

ISSN : 1246-8185

Référence électronique

Laura-May Simard, « Le corpus comme aide à la rédaction de résumés scientifiques pour des étudiants LANSAD : une approche comparative », *ASp* [En ligne], 73 | 2018, mis en ligne le 01 mars 2019, consulté le 02 mai 2019. URL : <http://journals.openedition.org/asp/5122> ; DOI : 10.4000/asp.5122

Ce document a été généré automatiquement le 2 mai 2019.

Tous droits réservés

Le corpus comme aide à la rédaction de résumés scientifiques pour des étudiants LANSAD : une approche comparative

Using a corpus for abstract writing in ESP classes: a comparative approach

Laura-May Simard

Introduction

- 1 Les étudiants qui aspirent à intégrer le monde de la recherche doivent maîtriser l'anglais académique et ses codes. C'est l'avis de Swales (1997 : 374) quand il compare l'anglais à un carnivore puissant (*English as a Tyrannosaurus Rex*) qui engloutit les autres habitants des pâturages de la linguistique académique. Ce constat est toujours pertinent, si ce n'est plus, vingt années plus tard, alors qu'un nombre croissant de cursus universitaires incluent une formation en anglais (Millot 2017 : 70).
- 2 Si l'anglais scientifique est régulièrement enseigné, nous proposons d'explorer une méthode encore peu utilisée dans les salles de classe, mais reconnue comme ayant des résultats positifs (O'Sullivan 2007 ; Chambers 2010 ; Boulton & Tyne 2013) : le *data-driven learning*, tel que nommé par un des précurseurs de cette approche, Johns (1991). Cette méthode, aussi appelée « apprentissage sur corpus » (ASC) par Boulton et Tyne (2013), utilise des corpus pour l'enseignement langagier. L'étudiant se positionne alors typiquement comme un chercheur et le professeur comme un guide dans un apprentissage qui se fait au rythme des découvertes dans le corpus. Afin de décrire cette dynamique d'apprentissage, Chambers (2010 : 15) rappelle la métaphore du détective Sherlock Holmes utilisée par Johns (1997 : 101) qui « souligne le rôle plus actif et plus autonome de l'apprenant dans cette nouvelle approche ».

- 3 Dans le cadre d'un cours d'anglais scientifique qui vise notamment l'apprentissage de la rédaction du résumé de recherche (ou « abstract »), nous voulons essayer de mesurer l'apport qu'un corpus spécialisé peut avoir comparé à un outil plus usité des étudiants : le dictionnaire en ligne. Nous proposons ici une étude comparative de l'utilisation de ces deux ressources pour la rédaction d'un résumé.
- 4 Après une première partie dédiée au cadre théorique, nous présentons le cadre méthodologique et institutionnel de cette expérience. La troisième partie consiste en une analyse de l'utilisation des deux ressources par des étudiants en master. La dernière partie de cette étude propose une discussion et des applications pédagogiques des résultats.

1. Cadre de l'étude

1.1. Les recherches en apprentissage sur corpus

1.1.1. Un domaine en plein essor, mais peinant à s'exporter

- 5 Dans leur panorama rétrospectif de l'ASC en 2013, Boulton et Tyne parlent d'une « révolution corpus » en citant McCarthy (Boulton & Tyne 2013 : 97 ; McCarthy 2008 : 564). Il est vrai que depuis sa popularisation par Johns dans les années 1980, l'ASC n'a cessé de se développer (Chambers 2005 : 111) de concert avec la linguistique de corpus.
- 6 Même si la recherche en ASC ne cesse de se développer et de démontrer les avantages de l'outil corpus, les résultats publiés par les chercheurs ne semblent pas avoir de réel impact sur les politiques pédagogiques de l'apprentissage des langues en France et ailleurs. (Cazade 2001 ; Bernardini 2000 ; Römer 2006 ; McCarthy 2008 ; Boulton & Tyne 2013). L'ASC offre des perspectives intéressantes, mais trop peu explorées par les enseignants. Ce phénomène explique très certainement la véritable croisade entreprise par certains chercheurs, qui se sont donné une « mission corpus » (Römer 2009). En effet, on comprend cette position en lisant le raisonnement d'O'Sullivan :

An understanding of the implications and relevance of corpus use for pedagogy may help teachers and learners overcome this resistance, and hence accelerate the process of “percolation” (McEnery & Wilson 1997 : 5) or the “trickle down” (Leech 1997 : 2) of corpus research to language teaching and learning. (2007 : 269)

1.1.2. Des résultats positifs à nuancer

- 7 Nombre d'expériences effectuées en ASC montrent des résultats encourageants (Sun 2003 ; Chambers & O'Sullivan 2004 ; Chambers 2005 ; Koo 2006 ; Boulton 2010 ; Mueller & Jacobsen 2016 ; Boulton & Landure 2016 ; Kennedy & Miceli 2017) ; néanmoins on trouve également des critiques à leur égard.
- 8 Premièrement, malgré les résultats avancés, nombreux sont les chercheurs qui déplorent une trop grande distance entre recherche et pratique : « *the direct exploitation of corpora in the EFL classroom is unusual* » (Aijmer 2009).
- 9 De plus, la nature même de certaines de ces études, qui utilisent des questionnaires et entretiens avec des étudiants comme unique source afin de confirmer l'apport des corpus, ne permet pas de résultats quantitatifs autres que concernant l'attitude des étudiants, comme le soulignent Boulton et Tyne (2013 : 100) en se référant notamment à Sun (2003). Il faut noter que ces études purement qualitatives ne représenteraient que moins d'un

quart des études publiées en linguistique appliquée selon Richards (2009), et que le domaine est tout à fait abordable avec des outils quantitatifs tels que les statistiques. Les recherches quantitatives et avec des méthodes diverses sont en plein développement (Boulton & Cobb 2017), la présente expérience s'inscrit dans cet élan.

1.2. Apprentissage sur corpus : applications pédagogiques

- 10 Dans cette section, nous proposons un rapide aperçu des applications pédagogiques que peuvent avoir les corpus. Au vu du large choix de types de corpus et d'approches d'utilisation, plusieurs exercices, visant des compétences différentes, sont possibles avec un corpus. Chambers parle de l'éventail de choix et mentionne notamment « des projets [qui] ont vu le jour dans lesquels un corpus sert de base à des exercices de reconstitution (Mangenot 2000) ou à la révision d'un texte écrit par l'apprenant (O'Sullivan & Chambers 2006) » (2010 : 10). Nous commencerons par parler de l'enseignement et de l'apprentissage avec un corpus : comment ils informent les pratiques pédagogiques et comment ils peuvent aider les apprenants d'une langue. L'aide à la rédaction étant l'objet de notre étude, on s'attardera plus en détail sur l'utilisation des corpus dans ce contexte-ci. Sans proposer ni détailler des exercices précis à réaliser avec corpus, les applications possibles sont évoquées succinctement.

1.2.1. Enseignement et apprentissage avec un corpus

- 11 Les corpus peuvent être utilisés de manière indirecte pour l'enseignement ; ils informent notamment les créateurs de matériel pédagogique (Boulton & Landure 2016). De plus, les corpus d'apprenants ont également vocation à informer les pratiques des enseignants (Flowerdew 2001). De Cock et Tyne (2014 : 138-145) présentent plusieurs études visant à mieux comprendre les difficultés des étudiants et à informer les pratiques d'enseignement.
- 12 Le A de ASC, l'apprentissage, illustre les cas où le corpus sert à sensibiliser les étudiants à la manière dont la langue est utilisée. Harwood (2005) avance que le but premier de l'utilisation de supports pédagogiques (dont le corpus) en anglais de spécialité est de sensibiliser les étudiants aux principales caractéristiques du discours universitaire. En effet, les corpus peuvent être utilisés en classe pour faire de l'analyse de discours ou simplement pour explorer des faits de langue (appartenant à un genre particulier ou non selon si on utilise un corpus général ou spécialisé).
- 13 Chambers (2010 : 13) va dans ce sens, mais indique que le corpus peut également constituer une aide précieuse à la rédaction :
- La plupart des enseignants-chercheurs qui ont étudié l'emploi de corpus par leurs étudiants se sont concentrés sur l'apprentissage des langues en général plutôt que sur l'apprentissage de la compétence plus spécifique qu'est l'écriture.
- 14 C'est par ailleurs ce que nous cherchons à analyser dans la présente étude. Ces deux utilisations du corpus ne sont cependant pas mutuellement exclusives. Kennedy et Miceli suggèrent d'ailleurs que l'utilisation guidée d'un corpus pour un exercice de rédaction permet de développer la sensibilité linguistique des étudiants et de les entraîner à se poser les bonnes questions (2017 : 111).

1.2.2. Aide à la rédaction

- 15 De nombreuses études démontrent que le corpus peut être utilisé comme une ressource pour la rédaction (Kennedy & Miceli 2001 ; Lee & Swales 2006 ; Boulton & Landure 2016 ; Boulton & Cobb 2017), même avec des étudiants de niveau plus faible (Frankenberg-Garcia 2014 ; Mueller & Jacobsen 2016). De plus, il a été écrit que l'utilisation d'un corpus spécialisé pour enseigner la rédaction d'un genre spécialisé permet aux étudiants d'améliorer leurs compétences de rédaction (Landure 2004 ; O'Sullivan & Chambers 2006) et de disposer d'exemples directement liés à leur contexte de rédaction (Chambers 2010). La question qui se pose alors est celle de la comparaison du corpus à d'autres ressources plus classiques (notamment les dictionnaires).
- 16 Si le dictionnaire est une ressource plus pertinente pour vérifier l'orthographe d'un mot (Chambers & O'Sullivan 2004 : 166), il n'en est pas forcément de même pour l'usage des prépositions, qu'un corpus pourra éclairer (*ibidem* 167-8). Dans son étude qualitative de l'utilisation de corpus, dictionnaires et dictionnaires de synonymes, Koo (2006) demande à ses étudiants de paraphraser un article de presse en utilisant ces différentes ressources. L'exercice de rédaction est souvent réalisé en combinant les ressources, mais les étudiants utilisent néanmoins l'outil corpus plus fréquemment (47,4 % du temps), notamment pour chercher des collocations. Landure, dans une étude comparant dictionnaire bilingue, corpus général et corpus spécifique, conclut que les dictionnaires, tout comme les corpus, « ne peuvent répondre à tous les besoins » (2014 : 316) tout en précisant que les dictionnaires ne permettent pas de faire des recherches de collocations ou d'expressions idiomatiques spécifiques. Mueller et Jacobsen (2016 : 5) indiquent que des études récentes montrent que les données offertes par un corpus peuvent être plus utiles que des données plus classiques. Il faut toutefois préciser qu'un nombre croissant de dictionnaires incluent des exemples et des données linguistiques directement issus de corpus (Frankenberg-Garcia 2015 : 490), les dictionnaires ne sont donc plus aussi éloignés d'un contexte authentique qu'avant. Cependant, Frankenberg-Garcia (2014) souligne la nécessité pour les apprenants d'avoir accès à de multiples exemples (au moins trois) afin d'améliorer leur production. Ceci n'est pas le cas dans tous les dictionnaires incluant des exemples tirés de corpus. Enfin, la combinaison des ressources pour la production semble être la solution la plus bénéfique au vu des apports différents des dictionnaires et des corpus (Landure 2004 ; Koo 2006 ; Frankenberg-Garcia 2015).
- 17 Ces résultats positifs doivent néanmoins prendre en compte les « retours d'expérience » des étudiants, qui sont parfois négatifs (Pérez-Paredes et al. 2012 : 489). Certains déplorent des interfaces trop compliquées (Mueller & Jacobsen 2016 : 8), d'autres la difficulté à se positionner comme « chercheurs en linguistique » lors de l'utilisation du concordancier (Kennedy & Miceli 2017 : 93). Les difficultés lors de la prise en main du concordancier et de la formulation d'hypothèses sont en effet des freins courants à la bonne utilisation du corpus (Kennedy & Miceli 2001 ; Bernardini 2000 ; O'Sullivan & Chambers 2006). La présente étude cherche à confirmer, ou non, certains de ces résultats dans le cadre d'un exercice de rédaction d'un genre spécialisé (le résumé d'article de recherche) avec accès à une seule ressource ; on ne pourra ainsi se prononcer sur les bénéfices de la combinaison d'un dictionnaire et d'un corpus.

2. Méthodologie

- 18 L'expérience consiste à demander aux étudiants de faire un travail de rédaction (d'un résumé) avec accès à un corpus de résumés ou un dictionnaire bilingue en ligne. Les étudiants doivent également remplir un journal de bord dans lequel ils détaillent toutes leurs recherches. Deux questionnaires, un en début de semestre visant à préciser le profil des étudiants puis un en fin de semestre afin d'avoir un retour sur la formation et l'expérience, permettent de compléter les données présentées.

2.1. Cadre de l'expérience : contexte et public

- 19 L'étude s'inscrit dans un contexte universitaire d'apprentissage de la rédaction du résumé. L'expérience présentée a été mise en place dans le cadre de vacances à l'École normale supérieure Paris-Saclay (ENSPS) au cours du premier semestre de l'année universitaire 2016-2017. L'ENSPS est une grande école sélective accueillant des étudiants de divers domaines, à partir de la L3, sur concours ou dossier. Les deux classes ayant participé à l'expérience ont eu une formation de vingt heures en anglais scientifique. Dans le cadre du nouveau diplôme de l'ENSPS mis en place en septembre 2016, chaque étudiant, quel que soit son parcours, doit passer une certification d'anglais scientifique conçue par le département d'anglais : le SWAP (*Scientific Writing Assessment Program*). Cette certification, qui se veut professionnalisante, est en partie centrée sur la connaissance du genre du résumé et certaines des six sections font appel à une connaissance de la phraséologie employée ou encore des différentes étapes du résumé. La préparation au SWAP mène à une introduction au genre du résumé, sa phraséologie et son style. Comme tous nos étudiants sont formés au SWAP (ou reçoivent une formation similaire), ils abordent l'exercice de rédaction d'un résumé, que nous leur proposons dans le cadre de cette étude, avec un bagage de connaissances non négligeable, qu'il faut prendre en compte dans l'analyse des résultats.
- 20 Le public est constitué de vingt-quatre étudiants inscrits en master et dans deux parcours distincts à l'ENSPS. Le premier groupe est composé de dix étudiants en M1 Sciences sociales, la majorité des élèves ayant un niveau « avancé », selon les coordinatrices des cours LANSAD de l'ENSPS. Le deuxième groupe est composé de quatorze étudiants en M2 Génie mécanique, dont le niveau d'anglais est plus faible ; certains d'entre eux ont une langue maternelle différente du français, ce qui peut avoir une incidence sur leurs productions en anglais. Les métadonnées concernant les participants sont consignées dans un tableau (Annexe 1) L'ensemble des étudiants a reçu la même formation à la rédaction du résumé et à l'utilisation du corpus. Tous les étudiants ont été informés lors du premier cours de l'étude et ont accepté d'y participer tout en étant assurés que leurs travaux seraient anonymisés. La majorité des étudiants ont vocation à faire de la recherche dans le cadre de leur profession ; ils se sont montrés pour la plupart réceptifs à un cours d'anglais de la recherche et à l'utilisation d'un nouvel outil.

2.2. Ressources utilisées

- 21 Le choix des deux ressources a pour but de permettre une évaluation de l'apport d'un corpus par rapport à un outil plus classique. Nous avons initié nos étudiants à l'utilisation

d'un corpus spécialisé (de résumés), premièrement car ils devaient pouvoir s'en servir pour l'expérience, mais également car le but du cours est de les sensibiliser aux outils qu'ils peuvent utiliser pour parfaire leur anglais scientifique.

- 22 Le corpus spécialisé utilisé avec les étudiants est constitué d'une partie d'un corpus de résumés d'articles scientifiques compilé par Anthony Saber, directeur du département de langues à l'ENSPS, en utilisant les outils avancés de recherche bibliographique et d'exportation des références de ScienceDirect. Ce corpus de résumés est divisé en dix sous-corpus par domaine (biologie, chimie, linguistique, médecine, etc.). Nous avons choisi de travailler avec deux sous-corpus dans les domaines de nos étudiants : Sciences sociales et Génie mécanique. Nous justifions ce choix d'un corpus relativement restreint (382 942 mots) en nous appuyant sur les travaux de plusieurs chercheurs (Tribble 2001 ; Flowerdew 2004 ; Charles 2007) qui considèrent qu'un corpus restreint est plus adapté dans le cadre d'enseignements en anglais de spécialité. De plus, Aston (1997) considère que l'utilisation d'un corpus restreint est un bon premier pas vers des corpus plus larges et permet de comprendre les enjeux de l'ASC. Le corpus de résumés utilisé contient 1 995 résumés (1 000 en Génie mécanique et 995 en Sciences sociales) pour un total de 382 942 mots. Nous avons utilisé le concordancier en ligne IMS Corpus Workbench (Evert & Hardie 2011)¹ plutôt qu'un concordancier à télécharger tel que Antconc pour des raisons pratiques. En effet, nous avons encouragé les étudiants à utiliser le corpus chez eux et nous pensons que le fait de devoir télécharger une application sur leur ordinateur serait un frein. Le corpus a été mis en ligne sur IMS (plateforme par ailleurs utilisée pour les cours de linguistique de corpus de l'UFR EILA de Paris 7) et protégé par un mot de passe. Le concordancier offre plusieurs outils, mais nous n'en avons présenté que deux en détail aux étudiants : *standard query*, qui permet de chercher une suite de mots dans le corpus et d'avoir accès à toutes les occurrences de celle-ci, et l'outil *collocation*, qui permet de chercher les collocats les plus fréquents d'un mot ou d'une suite de mots. Les étudiants étaient vivement encouragés à se servir du corpus chez eux après la première approche que nous allons détailler dans la section 2.3.2. C'est à la suite de cette très courte introduction à l'outil corpus que nous avons demandé aux étudiants de réaliser l'exercice qui sert pour l'étude.
- 23 Afin de juger l'apport qu'un outil tel que le corpus peut avoir sur la rédaction d'un résumé, il est nécessaire d'avoir un « témoin » auquel comparer les écrits avec corpus. Le dictionnaire en ligne WordReference (WR) a été retenu pour plusieurs raisons. Les étudiants connaissent tous cet outil et beaucoup (79,2 %) le citent comme une ressource régulièrement utilisée, une initiation en classe n'était ainsi pas nécessaire. Le fonctionnement de WR est éloigné de la mise en contexte d'un corpus spécialisé. Enfin, ce dictionnaire, bien que très complet, concerne l'anglais général alors que le corpus ne contient que des éléments d'anglais scientifique et du genre du résumé.

2.3. Mise en place de l'expérience

- 24 L'exercice de rédaction mis en place essaye de répondre à une volonté de rigueur méthodologique tout en se pliant au cadre institutionnel (notation des étudiants, disponibilité de la salle informatique, nombre d'heures à dédier à l'expérience...). Après avoir présenté plus précisément les questionnaires soumis aux étudiants, nous expliquons en premier lieu comment nous avons initié les étudiants à l'outil corpus, puis nous présentons l'exercice en détail.

2.3.1. Questionnaires de début et de fin de semestre

- 25 Les étudiants ont rempli deux questionnaires en début et fin de semestre. Le premier visait principalement à recueillir des informations à leur sujet (profil, connaissance de l'anglais, connaissance des corpus) et a permis l'élaboration du tableau 5 (Annexe 1) présentant le profil des étudiants. Le deuxième questionnaire, plus long, demandait aux étudiants de donner leurs sentiments sur l'expérience. Ils devaient premièrement répondre à des questions avec une échelle de Likert en quatre points à propos de la prise en main du corpus, de son utilisation, de la rédaction avec corpus et puis de l'exercice de rédaction de résumés à proprement parler. Les étudiants pouvaient ensuite écrire un commentaire libre au sujet de l'outil corpus tel qu'il avait été présenté en cours.
- 26 Il n'est pas possible de présenter toutes les données dans cette étude et nous avons décidé de ne mentionner les résultats de ces différents questionnaires que pour illustrer notre propos, le cas échéant.

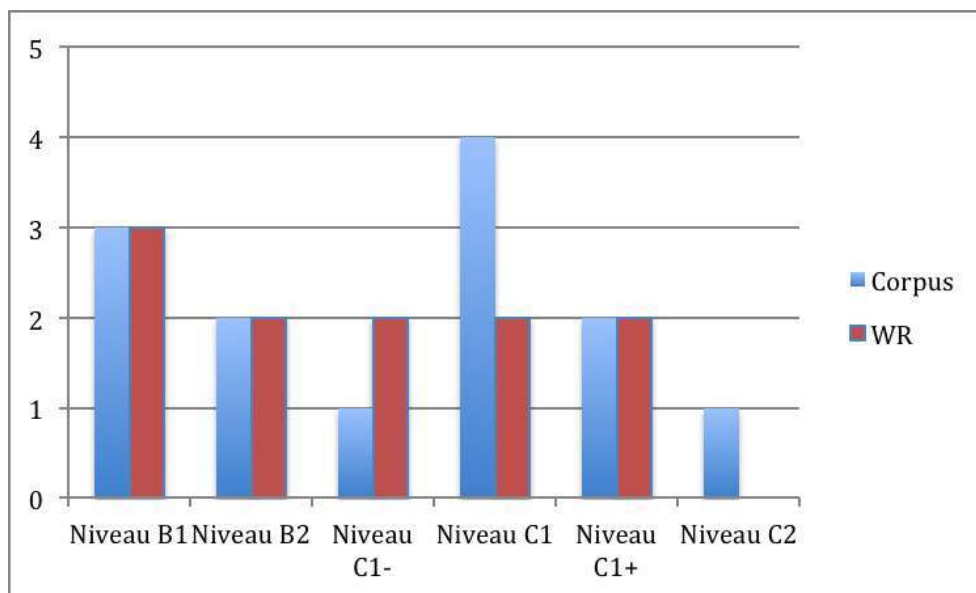
2.3.2. Première approche du corpus : exercice guidé et pratique autonome

- 27 Une majorité (70,8 %) des étudiants ignorait ce qu'était un corpus au début du semestre et tous étaient de nouveaux utilisateurs. La question de la première approche se posait et nous devions notamment faire un choix entre approche libre ou guidée. Dans son étude de 2007, Charles parle également de ce dilemme en opposant les bénéfices de l'approche guidée (résultats plus exploitables et interprétations guidées) à ceux de l'approche libre (rôle de chercheur [Johns 1991] et d'explorateur [Bernardini 2001] pour l'étudiant).
- 28 Tout comme Charles, nous avons choisi une première approche guidée afin de sensibiliser les étudiants le plus tôt possible à un outil qui ne leur est pas familier :
- Providing a controlled introduction to concordancing with a limited number of concordance lines and a clearly defined set of questions can assist students, not only to gain specific answers to the questions posed, but also to develop their searching and interpreting techniques. (2007 : 297)
- 29 De plus, le temps constituait un facteur crucial. En effet, les contraintes d'organisation du cours et d'accès à la salle informatique nous permettaient seulement deux heures de sensibilisation au corpus et il fallait que celles-ci soient les plus efficaces possible, étant donné le défi que représente la consultation d'un corpus pour des étudiants non initiés (Pascual Pérez-Paredes et al. 2012). Nous voulions surtout sensibiliser les étudiants à l'utilisation de l'outil, tout en continuant de travailler sur le genre du résumé. Tout comme Sun (2003), nous avons choisi d'allier une introduction théorique à une mise en pratique. Après une brève explication de deux des principaux outils du concordancier (voir 2.2), nous avons choisi de donner aux étudiants une liste d'expressions souvent relevées dans les résumés à compléter par des collocations trouvées grâce au concordancier. Par exemple, « *this study...* » devait être complété par des verbes adéquats. Cet exercice permet aux étudiants de comparer l'intérêt des deux fonctionnalités et de se familiariser avec l'interface du concordancier. Les étudiants devaient ensuite se servir du corpus pour essayer de corriger un résumé rédigé à la maison. Ils étaient guidés pendant la séance s'ils avaient des questions ou des problèmes de prise en main.

2.3.3. L'exercice de rédaction - comparaison entre l'apport d'un corpus et d'un dictionnaire de traduction en ligne (WordReference)

- 30 L'exercice de rédaction est réalisé par les deux groupes séparément et les groupes ne communiquent pas entre eux. Chaque étudiant dispose d'un même article de presse du journal britannique *The Guardian*² détaillant une étude sur l'origine de la sociabilité des chiens et donnant toutes les informations nécessaires à l'écriture d'un résumé correspondant. Plusieurs facteurs ont déterminé ce choix. Premièrement, les domaines de la biologie zoologique et de la génétique ne sont pas maîtrisés par nos étudiants, ce qui signifie qu'aucun d'entre eux n'a un avantage de connaissances phraséologiques ou terminologiques lors de la rédaction. Ensuite, le style journalistique dans lequel est rédigé l'article n'est pas ce qui est attendu dans un résumé d'article scientifique, ce qui les oblige à reformuler l'information afin de respecter le genre attendu. Enfin, l'article comprend 550 mots et la langue est accessible, ce qui est primordial pour éviter autant que possible que l'exercice soit moins bien réalisé par certains en raison de problèmes de compréhension. Les étudiants disposent d'une heure pour rédiger un résumé correspondant à l'article sur ordinateur.
- 31 Chaque groupe est divisé en deux, un demi-groupe a accès au corpus de résumés et l'autre à WR. À la suite de désistements d'étudiants, les groupes sont légèrement déséquilibrés : sur 24 étudiants, 13 travaillent avec le corpus et 11 avec WR. Nous veillons à ce que le niveau des étudiants dans chaque sous-groupe soit le plus équilibré possible afin que ce facteur ne soit pas trop important dans les résultats statistiques. La figure 1 indique la répartition des étudiants selon la ressource et le niveau.

Figure 1. Répartition des étudiants selon la ressource et le niveau d'anglais



- 32 Outre le résumé, il est demandé à chaque étudiant de rendre un deuxième document que nous appelons un « journal de bord »³, dans lequel sont consignées toutes les recherches faites avec la ressource à disposition. De plus, ils doivent mettre en gras tous les passages du résumé écrits avec l'aide de la ressource.
- 33 Les consignes sont les suivantes :

Write an academic abstract for the following article. You have to write the abstract as if you were one of the scientists on the research team. During your work, you will have access to [resource]. Please put in bold any part of your abstract that you will have written with the help of this resource.

You also have to create a separate document in which you will write down anything you looked up with the resource.

- 34 Cet exercice nous permet d'avoir vingt-quatre résumés contenant les mêmes informations ainsi que des explications sur l'utilisation de la ressource pour vingt-trois résumés, un étudiant n'ayant pas jugé nécessaire d'utiliser le corpus.

2.4. Traitement des données

2.4.1. Compilation du corpus d'apprenants

- 35 Comme nous l'avons mentionné précédemment, l'exercice mis en place permet de constituer un corpus de résumés d'apprenants (désormais CA) de 7 440 mots. Afin de faciliter la lecture des résultats, chaque résumé est nommé comme suit : Numéro d'anonymat_Ressource_Domaine_Version. Par exemple, le résumé de l'étudiant 1 qui est en Sciences sociales et a eu le corpus comme ressource est nommé : 1_C_SS. Notons par ailleurs que la taille restreinte du corpus nous permet de faire nos analyses sans outil, mais simplement en passant en revue chaque résumé.

2.4.2. Analyses quantitatives et qualitatives avec plusieurs sources

- 36 Ces dernières années, le développement des technologies et de la recherche en ASC a mené à une multitude d'approches afin de comprendre comment les étudiants utilisent le corpus. Ces données sont essentielles pour compléter les résultats d'une étude comme celle-ci (Pérez-Paredes et al. 2011 : 249). L'enregistrement en temps réel des pratiques de recherche des étudiants (Park & Kinginger 2010) et l'utilisation des historiques de consultations couplés à des questionnaires et entretiens (Landure 2014) font partie des méthodes de collectes de données sur l'utilisation des corpus par les apprenants.
- 37 Suivant une méthodologie proche de celle de Kennedy et Miceli (2017), nous avons choisi de demander aux étudiants de rendre compte de leurs recherches au fil des travaux demandés. Les étudiants ayant tous eu accès à une formation en anglais scientifique et à la rédaction du résumé, il nous semble que l'utilisation de la ressource est la variable la moins problématique à étudier. En effet, tout passage contenant des erreurs (ou, au contraire, des choix particulièrement judicieux) et qui n'est pas attribué à l'utilisation de la ressource ne peut être analysé dans le cadre d'une étude de l'apport du corpus à la rédaction. Ce n'est qu'en s'intéressant à ce que produisent les étudiants en utilisant leur ressource que nous pouvons tirer des conclusions véritables sur l'apport de ces ressources. C'est pourquoi nous analysons les productions des étudiants dans le cadre de l'utilisation de la ressource (autrement dit, les passages qu'ils ont mis en gras).
- 38 L'ensemble de ces données nous permet de chercher des réponses aux questions suivantes : Comment la ressource (corpus ou WR) est-elle utilisée ? Cette utilisation entraîne-t-elle une amélioration ou est-elle source d'erreur ?

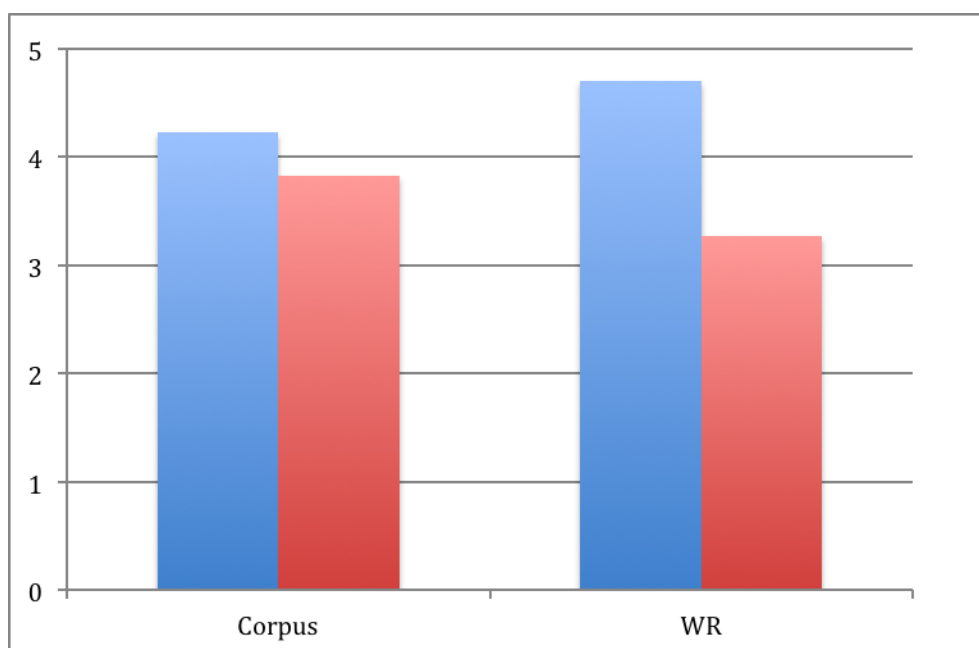
3. Utilisation des ressources par les étudiants

- 39 Nous proposons d'étudier la manière dont les étudiants se servent de la ressource à laquelle ils ont accès, ainsi que l'efficacité de cette utilisation. Les résultats proviennent des journaux de bord tenus par les étudiants pendant l'exercice. Les statistiques d'utilisation de la ressource sont d'abord présentées, puis le type de recherches mené pour chaque ressource et enfin l'efficacité des ressources.

3.1. Résultats statistiques

- 40 La figure 2 indique la répartition du nombre moyen de recherches et de leur utilisation par ressource. L'utilisation de la ressource se calcule en observant le nombre de passages mis en gras dans le résumé par l'étudiant.

Figure 2. Utilisation des recherches par ressource pour la rédaction



- 41 On peut voir que si les étudiants utilisant WR font légèrement plus de recherches en moyenne, la tendance s'inverse pour les recherches utilisées. Nous verrons si ce phénomène peut s'expliquer par la manière dont la ressource est utilisée.

3.2. Modes d'utilisation de la ressource

- 42 Il convient maintenant de se pencher plus en détail sur la manière dont les étudiants font leurs recherches, c'est-à-dire la manière dont ils se servent de la ressource pour faire des recherches. Pour cela, nous nous servons des journaux de bord des étudiants. Ces journaux ne sont pas tous écrits de la même manière, et certains étudiants n'utilisent pas de métalangage pour expliquer leurs recherches.

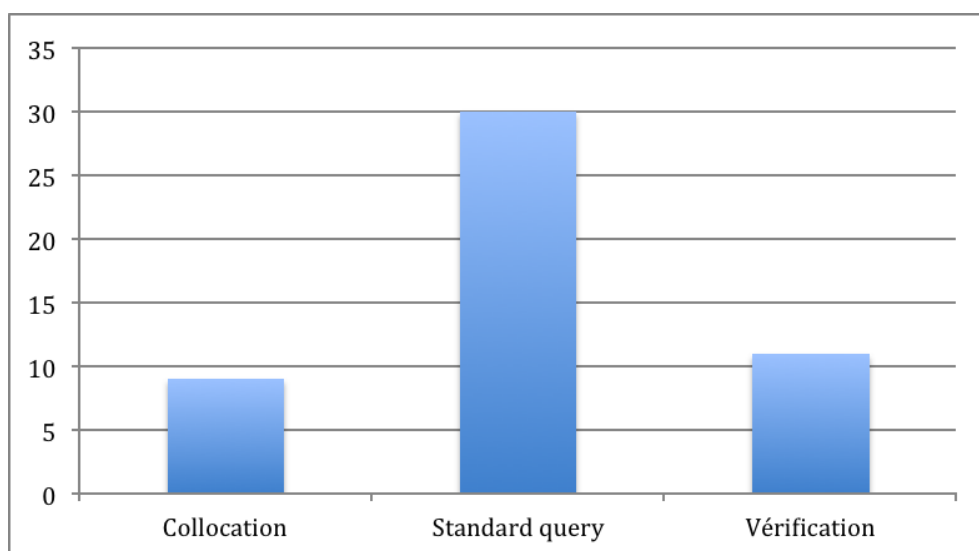
- 43 Nous avons néanmoins identifié trois types de recherches pour la ressource corpus (*collocation*, *standard query* et *vérification*) et quatre types de recherches pour la ressource WR (*synonyme*, *anglais vers français*, *français vers anglais* et *orthographe*).

Tableau 1. Types de recherches effectuées avec les ressources

Corpus	
Collocation	L'étudiant se sert de la fonction collocation sur le corpus afin de voir les collocats d'un mot ou d'une expression.
Standard Query	L'étudiant cherche toutes les occurrences du mot ou expression avec l'outil « standard query » et tire des conclusions des résultats en les voyant tous.
Vérification	Ceci est une sous-catégorie de recherche que nous avons cru bon d'identifier. L'étudiant dit avoir utilisé le corpus expressément pour vérifier une intuition, il cherche à confirmer que sa formulation est correcte.
WordReference	
Synonyme	L'étudiant cherche les synonymes d'un mot.
Anglais vers français	L'étudiant se sert de WR pour traduire de l'anglais vers le français, principalement pour comprendre l'article du <i>Guardian</i> .
Français vers anglais	L'étudiant se sert de WR pour traduire du français vers l'anglais.
Orthographe	L'étudiant se sert de WR pour vérifier l'orthographe d'un mot.

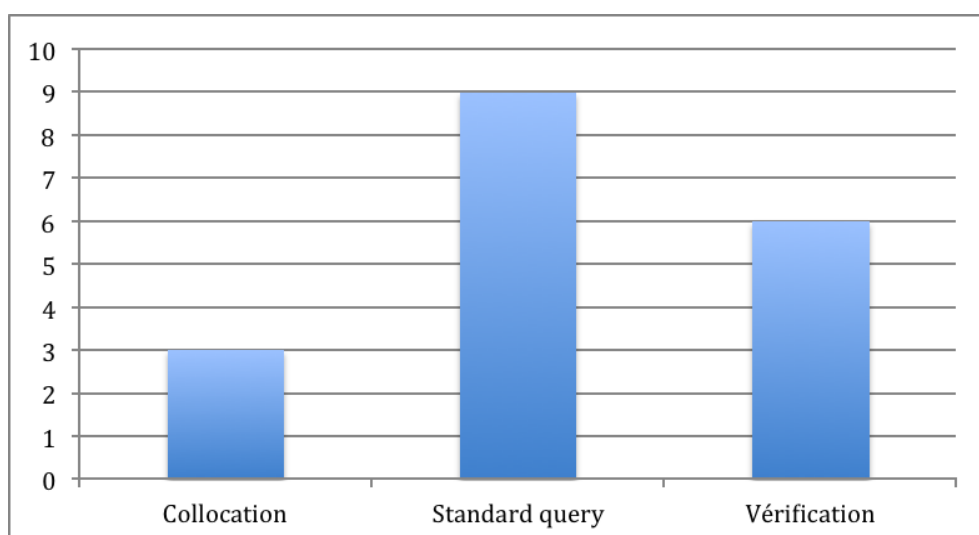
- 44 La figure 3 illustre la répartition du type de recherches menées avec le corpus. On peut voir que les recherches de type « *standard query* » sont les plus courantes (30 recherches). Ceci s'explique probablement par le fait que c'est l'outil le plus immédiat sur le concordancier et les étudiants ne sont pas forcément à l'aise avec l'outil collocation.

Figure 3. Répartition du type de recherches faites sur corpus



- 45 Nous étions avec les étudiants lors de l'exercice et avons proposé à ceux utilisant le concordancier de les aider s'ils avaient du mal à manier l'outil, sans pour autant les aider dans leur choix de recherches ou l'interprétation des résultats. Le plus faible nombre de recherches de collocation n'est normalement pas dû au fait que les étudiants avaient oublié la technique permettant de chercher des collocats.

Figure 4. Répartition du type de recherches sur corpus par nombre d'étudiants



- 46 Nous ne nous attendions pas à trouver autant d'utilisations du concordancier pour « vérification » et avons surtout présenté l'outil comme un moyen de trouver des expressions ou termes. Comme l'indique la figure 4, six étudiants sur les treize utilisant le concordancier ont utilisé cette fonction, ce qui indiquerait que même après une courte prise en main, ils se sentaient assez à l'aise avec l'outil pour s'en emparer eux-mêmes et l'adapter à leurs besoins.

- 47 Maintenant que nous avons vu les fréquences d'utilisation et la manière dont les ressources sont utilisées, la dernière étape de notre réflexion consiste à essayer de mesurer l'efficacité de la ressource.

3.3. Efficacité de la ressource pour la rédaction

- 48 L'efficacité de la ressource est calculée en utilisant le nombre de recherches utilisées et le nombre de recherches efficaces. Nous considérons qu'une recherche est efficace si elle ne se trouve pas dans un passage avec une erreur qui lui est liée.
- 49 Dix étudiants ont fait des erreurs en lien avec la ressource utilisée (5 des 12⁴ étudiants sur corpus et 5 des 11 étudiants sur WR). Le pourcentage d'efficacité moyen avec le corpus est 78,68 % contre 72 % pour l'utilisation de WordReference. L'utilisation du corpus mène à une rédaction légèrement plus efficace que celle de WR, ce résultat semble indiquer que l'utilisation du corpus de résumés est plus efficace pour éviter les erreurs dans un résumé qu'un dictionnaire de langue générale, l'écart restant modeste. Comme le soulignent Mueller et Jacobsen, il faudrait idéalement pouvoir comparer l'utilisation de ces deux ressources pour chaque erreur afin de statuer précisément sur l'utilité du corpus face à celle du dictionnaire (2015 : 18).
- 50 Notons que les deux ressources ont un pourcentage d'efficacité supérieur à la moyenne : dans les deux cas, l'utilisation de la ressource n'est pas principalement source d'erreur. Il est difficile de dire si l'utilisation entraîne une amélioration de la rédaction en ne regardant que ces chiffres. Il faudrait savoir ce que l'étudiant aurait écrit avant l'utilisation de la ressource afin de faire une comparaison avec le résultat final ; nous n'avons pas de données de cette nature.
- 51 L'outil corpus étant l'objet principal de cette étude, nous allons d'abord commenter certains cas d'utilisation efficace du corpus. Puis nous présenterons les cas d'utilisation inefficace du corpus et un cas d'utilisation inefficace de WR qui aurait pu être amélioré avec un accès au corpus.

3.3.1. Utilisations efficaces du corpus

- 52 Comme nous venons de l'indiquer, le corpus est principalement utilisé de manière efficace (dans 78,68 % des cas). Nous n'allons pas développer tous les cas d'utilisation efficace (il y en a 37), mais nous nous concentrons sur certains exemples qui présentent un intérêt particulier et que nous recensons dans le tableau 2.

Tableau 2 . Utilisations efficaces du corpus

Étudiant	Recherche	Résultat
3_C_SS	standard query "hypothesis"	This paper investigates the hypothesis
	standard query "experiment"	An experiment was conducted to examine how dogs cooperate with humans
	standard query "interest"	This is a topic of great interest for the field

	standard query “encourage”	That’s why we encourage further investigation to start with other breeds of dogs
6_C_SS	Classified : j’ai cherché ce mot pour savoir s’il était bien employé dans ce sens/contexte	Their reactions were recorded and the different behaviors were classified
	Associated + préposition	Yet, four of these have been previously associated with an opposite behavior in humans.
	Could be + verbe : j’ai voulu vérifier le temps du verbe qui suivait <i>could be</i>	Our investigation could be extended to other breeds of dogs for more reliability
13_C_GM	Collocation database: Responsible	genes are responsible for canine interest in humans
	Collocation database: Experiment	Experiments were conducted

53 *Standard query*

L’étudiant 3 se sert de la fonction *standard query* afin de trouver des contextes d’utilisation de certains mots (*hypothesis, experiment, encourage* et *interest*). Dans chaque cas, la recherche lui permet une rédaction sans erreur et avec un langage adapté au genre visé.

- 54 Par exemple, pour le mot *interest*, une recherche permet de trouver des contextes d’utilisation tels que :

Figure 5. Extraits de concordancier pour *interest*

Thermally-enhanced grouting can be of significant	interest	in a high thermal conductivity ground (such as saturated sand)
Bone is a topic of great	interest	for researchers
many instructors are adding elements of high	interest	to students to their teaching materials and activities

- 55 L’étudiant s’est sûrement inspiré de ce type de cas d’utilisation pour en venir à sa propre mise en anglais : *This is a topic of great interest for the field.*

56 *Collocation*

L’étudiant 13 a utilisé l’outil de recherche de collocations. On ne sait malheureusement pas quelle partie du discours il recherchait, mais on peut se douter qu’il voulait trouver un verbe dans le cas du mot *experiment*. Le verbe *conducted* est en effet le verbe le plus fréquent dans le contexte de *experiment*. On ne sait ce que l’étudiant cherchait pour *responsible*, mais sa recherche lui a sûrement permis de trouver la bonne préposition (*for*) car c’est le mot qui entre en collocation le plus fréquemment avec cet adjectif.

57 *Vérification*


Comme nous l'avons stipulé, nous ne nous attendions pas à ce que les étudiants se servent autant du corpus pour effectuer des vérifications.

58 L'étudiant 6 s'est servi du corpus pour vérifier deux aspects différents de la langue. Pour le mot *classified*, il voulait vérifier que le contexte d'utilisation était le bon, il s'agit donc d'une recherche liée au lexique. Pour *could be*, en revanche, c'est le temps du verbe qui suit qui intéressait l'étudiant. Le corpus sert afin de vérifier de la grammaire, et cela fonctionne.

59 Ces différents exemples nous permettent d'avancer que des étudiants aux niveaux variés (C1+, C1 et B2 respectivement pour les étudiants 3, 6 et 13) peuvent utiliser le corpus de différentes manières en obtenant des résultats très satisfaisants. Nous avons vu des cas d'utilisation du corpus pour des phénomènes de langue (le temps d'un verbe), mais aussi de genre résumé (des collocations liées au genre scientifique). Notons que les vérifications présentées sont très simples et sont effectuées en une étape, on peut donc stipuler qu'elles sont a priori accessibles à tout étudiant, si tant est qu'il se pose de telles questions, comme le remarquent également Kennedy et Miceli (2017). Cette remarque nous amène à évoquer maintenant les utilisations inefficaces du corpus et à analyser certains exemples où, justement, le niveau d'anglais de l'étudiant semble être une barrière que le corpus ne peut aider à surmonter totalement.

3.3.2. Utilisations inefficaces du corpus

Tableau 3. Utilisations inefficaces du corpus

Étudiant	Recherche	Erreur
1_C_SS	Collocation : The paper  Paper describes	The paper describes the methodology we used.
14_C_GM	Verification : On future works => correct	On future work , more experiments will be performed to refine these results
23_C_GM	Standard query : Several...nouns	They are severals genes
	Standard query : Sliding...nouns	Three transparent sliding plastics covering a treat
	Standard query : Allow...preposition Identifying...nouns	The result observed allow identifying of five genes

60 Le tableau 3 présente des utilisations inefficaces du corpus. Nous allons commenter le cheminement de ces étudiants.

61 **Étudiant 1**

Cet étudiant a utilisé le corpus trois fois, dont une fois de manière inefficace. L'étudiant a utilisé la fonction collocation pour chercher des collocats de *paper*. Le verbe *describe* en est bien un ; on trouve treize occurrences de *paper describes* dans le corpus. Le choix de *paper describes* est donc a priori correct. Ce qui pose problème est le contexte plus large d'utilisation. L'étudiant écrit *the paper describes the methodology we used*. ce qui constitue une erreur de transfert de contenu car cette phrase laisse à penser que seule la méthode est décrite dans l'article, et non toute l'expérience.

Figure 6. Résultats dans le concordancier pour *paper describes*

is . First , the **paper** **describes** suitable methods for the determination of fracture-mechanical parameters . Second ,
 mmals . This **paper** **describes** a signal enhancement strategy to improve the sensitivity of an antibody-based
 oducts . This **paper** **describes** the optimization of hot rolling strategies based on the estimation of
 rete life . The **paper** **describes** mechanical properties , self-healing capacity and bonding behaviour of a sustainable
 . The current **paper** **describes** the results of an extensive experimental program which focused on determining
 aterial . This **paper** **describes** salient features of the test setup developed and used for conducting
 e tests . This **paper** **describes** the construction of PREMUX , its rationale and the experimental campaign
 oposed in this **paper** , **describes** well the deformation behavior of frozen loess under different stress
 energy . This **paper** **describes** a completely passive aortic counterpulsation solution , with an intra-aortic balloon
 signals . This **paper** **describes** a new converter based on an angle-tracking technique that employs synchronized
 ncoder . This **paper** **describes** the proposed method , its simulation and experimental results . Optical
 istance . This **paper** **describes** a trial production of SMC roof tiles in a composite manufacturing
 arison . This **paper** **describes** the process of identifying performance measures and associated indicators to serve

62 On voit dans les résultats du concordancier (figure 6) pour une *standard query* de *paper describes* que ce qui vient après *paper describes* est une description des contenus de l'article, plus ou moins détaillée. Le mot *method* est listé comme une première étape (*First, the paper describes suitable methods...*), ou fait partie d'une présentation du plan (*the paper describes the proposed method, its stimulation and experimental results.*) Si l'étudiant avait suivi sa première recherche de collocation par une recherche de *paper describes* en contexte, il aurait peut-être évité son erreur.

63 **Étudiant 14**

Cet étudiant a utilisé le corpus quatre fois, dont une fois de manière inefficace (il a donc un pourcentage d'efficacité de la ressource de 75%). On trouve un élément notable dans son utilisation menant à une erreur : l'expression **on future works*, qu'il dit avoir confirmée comme correcte avec le corpus, ne donne aucun résultat. En effet, une vérification de cette expression lui aurait normalement confirmé qu'elle est erronée et l'aurait empêché de commettre une erreur. On ne peut savoir ce qui s'est réellement passé quand l'étudiant a fait sa recherche. L'expression *future works* donne un résultat qui ne permet pas de confirmer que **on future works* est correct. Ici c'est donc une mauvaise utilisation de la ressource qui mène à une erreur.

Figure 7. Extrait du concordancier pour *future works*

Finally , this paper is concluded with possible	future works	.
---	--------------	---

64 **Étudiant 23**

Cet étudiant, de faible niveau, a utilisé le corpus cinq fois lors de la rédaction, dont trois fois de manière inefficace (il a donc un pourcentage d'efficacité de la ressource de 40 %). Toutes ses recherches sont des *standard query* et nous allons nous intéresser aux deux dernières (*allow* et *identifying*) qui ont mené à la phrase : **The result observed **allow identifying** of five genes.*

- 65 Cette phrase contient deux erreurs de grammaire : il y a un problème d'accord sujet-verbe, ainsi qu'un problème de forme grammaticale (on ne peut faire suivre *allow* d'une forme en -ING seule dans cette phrase). La première erreur est très courante en grammaire et l'étudiant connaît cette règle, elle n'est pas directement due à la ressource et n'est pas comptée comme telle. Le corpus comporte quarante occurrences de *allow* qui permettent de voir son contexte grammatical d'utilisation : *allow* + GN⁵ + TO + VERBE ou alors *allow* + GN. Il est possible que l'étudiant ait compris cela et ait essayé de faire une structure nominale en OF. Néanmoins, cette structure est fautive à cause de l'absence d'article défini et car on utiliserait plutôt *identification* que le gérondif *identifying* dans ce cas-là.
- 66 Si l'étudiant pense maîtriser cette structure alors il n'a pas de raison de faire des recherches à ce propos, comme le souligne également Frankenberg-Garcia (2014). De plus, il ne serait pas sans doute guère aisé pour un étudiant maîtrisant mal la langue de faire des recherches de ce type afin de confirmer des structures grammaticales. Le corpus n'aura pas suffi à combler les lacunes linguistiques de l'étudiant.

3.3.3. Utilisation erronée de WR rectifiable avec le corpus

Tableau 4. Cas d'utilisation inefficace de WR pour l'étudiant 17

Étudiant	Recherche	Erreur
17_WR_GM	Synonyme of Experiment => discovery - empiricism - experiment - research - testing - trial	Research was made in two steps.
		Dogs were tested in the first part of experiment

- 67 L'étudiant 17 avait accès à WR et s'en est servi deux fois, dont une fois de manière inefficace. Sa recherche de synonymes de *experimentation* mène à trois erreurs.
- 68 Pour **research was made in two steps*, le verbe *made* n'est pas une collocation habituelle de *research*. De plus, le mot *research* aura tendance à être défini dans ce contexte. Le corpus propose par exemple :

Figure 8. Extrait du concordancier pour *research was*

The experimental	research was	divided in two series
The	research was	conducted in two stages .

- 69 On voit la présence du défini et de verbes plus fréquents dans ce type de contexte : *divided* et *conducted*. L'étudiant n'avait pas accès au corpus et ne pouvait donc pas voir ces résultats. Il est possible que l'utilisation du corpus lui ait évité ces erreurs.
- 70 Pour **Dogs were tested in the first part of experiment*, là encore il y a un problème de détermination, *experiment* doit être défini. Une recherche de *the first part of* dans le corpus met ce phénomène en avant.

Figure 9. Résultats dans le concordancier pour *the first part of*

nechanical load for arbitrary milling operations . In [the first part of](#) the paper the modeling of the NC-simulation based approach is presented. The effect of relative density on the mechanical properties is investigated in this paper . In [the first part of](#) the work , the effect of relative density on the mechanical properties of binary cementing material in mortar and concrete . [The first part of](#) the investigation included testing the pozzolanic activity index , sodium chloride content and water absorption . In order to achieve this purpose , in [the first part of](#) the experimental study , conventional asphalt tests were carried out in the laboratory . [The first part of](#) the paper underlines the need for economic analysis of these services ;

- 71 Il faut tout de même rappeler que nos réflexes sur corpus ne sont pas forcément les mêmes que ceux des étudiants : nous ne pouvons être sûre que l'étudiant 17 aurait pensé à faire les recherches que nous venons de présenter. De plus, si l'étudiant ne pense pas se tromper sur la détermination de *the first part of experiment*, il n'a pas de raison de faire une telle recherche, comme nous l'avons souligné dans le cas de l'étudiant 23.
- 72 Les différentes analyses effectuées au cours de cette partie nous permettent d'avancer plusieurs raisons pour lesquelles le corpus peut être inefficace. Il est possible que l'étudiant fasse une erreur d'utilisation (étudiant 14) et dans ce cas nous n'avons pas assez d'informations pour comprendre cette erreur et proposer une amélioration de la méthode. Un étudiant peut également formuler ses conclusions trop rapidement (étudiant 1) et, dans ce cas-là, il est probable que des tutoriels plus complets sur l'exploitation des résultats permettraient une amélioration de l'utilisation de l'outil corpus, c'est également ce que pensent Landure (2014) et Kennedy et Miceli (2017). Enfin, il est possible que les difficultés langagières de l'étudiant (étudiant 23) ne lui permettent pas de résoudre certains problèmes rencontrés avec le corpus. Notons par ailleurs que si l'apprenant ne pense pas avoir commis une erreur, il ne fera pas de recherches. La connaissance langagière peut être un frein à l'utilisation du corpus. Kennedy et Miceli abondent dans ce sens :

We recognize that during corpus investigations by language learners, there is considerable room for error due to lack of knowledge of the target language. (2001 : 88)

4. Discussion

4.1. L'étude de l'utilisation de la ressource : un angle prometteur

- 73 Notre étude ayant lieu dans le cadre d'une formation en anglais scientifique et avec des étudiants régulièrement confrontés à des écrits scientifiques en anglais, il était évident que chaque étudiant avait un bagage de connaissances différent en anglais scientifique. Une étude de leurs écrits en tant que tels n'aurait pas permis de discerner l'apport du corpus (ou de WR) de leur propre apport lors de la rédaction du résumé. La méthode que nous avons utilisée (étude des recherches avec la ressource, puis des recherches utilisées et enfin de l'efficacité de cette utilisation) nous semble être un moyen satisfaisant d'avoir des résultats permettant de statuer sur l'apport du corpus et la manière dont il est utilisé.

- 74 De plus, les données permettent une meilleure compréhension du rapport au corpus même si elles ne présentent qu'une vision incomplète de la réflexion des étudiants face au corpus. C'est ce que soulignent également Kennedy et Miceli :

There were no feasible mechanisms for capturing completely the use the students made of CWIC⁶ (including searches they deemed not useful) at any time and from any place, let alone their intentions and thought processes. (2017 : 96)

- 75 Les journaux de bord des étudiants permettent d'entrevoir leur processus de réflexion. Nous avons noté que ceux-ci n'ont pas tous le même niveau de détails. Néanmoins, ce sont des « coulisses » non négligeables dans l'étude de l'utilisation du corpus pour la rédaction.

4.2. Applications pédagogiques

- 76 Une étude de terrain avec des étudiants est source de nombreuses remises en question, mais également d'inspiration. Nous détaillons ici nos deux intuitions pédagogiques les plus importantes à la fin de cette étude. Nous voulons insister sur la multifonctionnalité du corpus, mais également présenter des pistes pour un cours d'introduction à l'ASC.

4.2.1. Outil de rédaction et d'apprentissage

- 77 Nous avons mentionné que le corpus peut aider à l'apprentissage, à la rédaction, mais également à la création de supports pédagogiques. Nous traitons les deux premiers aspects ici, puis reviendrons sur le dernier point dans un second temps.
- 78 Si le corpus a des bénéfices directs pour la rédaction chez nos étudiants, nous avançons également que celui-ci leur permet de développer leurs connaissances langagières. L'effet le plus évident du corpus est sûrement l'apprentissage de la lexicogrammaire (quand un étudiant cherche un verbe qui entre en collocation avec *results* par exemple). Ceci est soutenu par Boulton et Cobb (2017 : 385) : « *When it comes to the language objectives themselves, we found large or medium effects for vocabulary and lexicogrammar.* »
- 79 Cependant, nous pensons également que le corpus a permis aux étudiants de développer leurs capacités de réflexion linguistique : par exemple, en analysant des lignes de concordancier ou en faisant des vérifications avec le corpus. Ceci se voit dans la réflexion menée par certains dans le journal de bord (voir notamment les étudiants 6 et 15 en Annexe 2). Si ces résultats ne font que confirmer ce que nous savons déjà, il pourrait être intéressant de s'en servir afin de convaincre un public moins enthousiaste : les étudiants. En effet, dans le questionnaire de fin de semestre, la moitié des étudiants pense que le corpus n'aide pas à apprendre la langue.

4.2.2. Conception d'un cours d'introduction à l'ASC

- 80 Les contraintes de l'environnement dans lequel nous avons fait cette étude ne nous ont pas permis de faire une introduction au corpus aussi développée que nous l'aurions aimé. Les étudiants ont d'ailleurs été plusieurs à faire des remarques allant dans ce sens dans les questionnaires de fin de semestre.
- 81 Il ne faut pas oublier que les étudiants se servent déjà d'outils informatiques pour l'apprentissage des langues et beaucoup utilisent Google comme un concordancier :

Learners are already involved in using information and communication technology (ICT) to search for answers to their language questions, especially via the use of Google as a "concordancer" for the Web as "corpus" (Chinnery 2008; Kilgarriff &

Grefenstette, 2003). Properly conceived DDL activities can build on these existing behaviors, refining them and using them as a way in to corpus work (Boulton 2015). (Boulton & Cobb 2017 : 4)

- 82 Il semblerait donc judicieux d'utiliser ce phénomène comme point d'entrée dans le monde de l'ASC. Nous avons eu l'impression que certains étudiants ne voyaient le corpus que comme un outil présenté à des fins précises ; une introduction présentant Google comme un outil linguistique et le comparant à d'autres corpus permettrait peut-être de les convaincre de l'accessibilité de l'outil corpus et des spécificités de chaque type de corpus.
- 83 Nous pensons qu'une introduction plus efficace devrait également inclure des exemples de bonnes (3.3.1) et de mauvaises (3.3.2) pratiques d'étudiants comme ceux que nous avons développés dans les sections mentionnées. De manière générale, une initiation devrait permettre aux étudiants de se familiariser avec l'outil corpus, mais également d'en comprendre l'utilité et les enjeux. Au vu des résultats de l'étude, nous suggérons d'insister sur le A de ASC : un corpus est un outil de rédaction/correction, mais également (et principalement ?) un outil d'apprentissage. En effet, les 91 % des étudiants convaincus de l'apport d'un corpus pour la rédaction d'un résumé sont à contraster avec les 48 % d'étudiants déjà mentionnés qui ne croient pas que le corpus aide à l'apprentissage de la langue et ne perçoivent donc les bénéfices de l'outil que dans le contexte d'un exercice de rédaction.

Conclusion

- 84 Cette étude a été conçue pour mesurer l'apport d'un corpus lors d'un exercice de rédaction. Nos résultats nous permettent d'avancer que, dans notre échantillon, les étudiants, même de niveau plus faible, parviennent majoritairement à se servir efficacement de cette ressource et viennent confirmer l'apport supérieur d'un corpus face à un dictionnaire dans le cas d'un exercice de rédaction spécialisé. En revanche, l'étude des cas d'utilisation inefficace et le retour des étudiants nous indiquent qu'une formation plus poussée à l'utilisation d'un corpus serait bénéfique. C'est pourquoi nous pensons que la « mission corpus » de certains chercheurs essayant de convaincre les enseignants de familiariser leurs étudiants à cet outil est justifiée et pourrait permettre aux étudiants de profiter de ressources disponibles autrement inexplorées. Le manque d'enthousiasme de certains étudiants en fin de semestre est potentiellement dû à une frustration face à un outil pas assez bien maîtrisé. De plus, notre étude montre que dans le cadre d'un exercice de rédaction d'un genre spécialisé, un dictionnaire en ligne de langue générale (comme WR) ne permet pas forcément d'avoir de bons résultats et a des limites parfois liées à la spécificité de la langue. Notre étude indique également qu'un corpus peut avoir un apport pour différents aspects d'une langue pour un public hétérogène. Enfin, même si cette étude ne présente pas de données sur l'utilisation à long terme du corpus, on peut noter que 66,7 % des étudiants ont affirmé qu'ils aimeraient utiliser des corpus spécialisés à l'avenir dans le cadre de leurs études.
- 85 Comme nous l'avons déjà mentionné, nous pensons que le processus de réflexion de l'étudiant lors de l'utilisation d'un corpus est une piste très enrichissante pour créer du contenu pédagogique. Nous avons essayé d'étudier ce processus, mais nous devons reconnaître que, même avec les journaux de bord des étudiants, certaines de nos interrogations quant à leur cheminement intellectuel restent sans réponse. De plus, le

petit échantillon d'étudiants avec lequel nous avons travaillé ne nous permet que d'avancer des pistes de réflexion et non des certitudes. Enfin, les contraintes temporelles de l'étude nous empêchent de savoir si les étudiants continuent d'utiliser un corpus sur le long terme. Plus de données sur la manière dont différents étudiants de différents niveaux et domaines accomplissent différentes tâches avec un corpus pourraient aider à convaincre que l'outil corpus n'est pas destiné qu'à un public spécifique dans un contexte particulier.

BIBLIOGRAPHIE

- AIJMER, Karin. 2009. *Corpora and Language Teaching*. Amsterdam : John Benjamins, 225–234.
- ASTON, Guy. 1997. « Small and large corpora in language learning ». In LEWANDOWSKA- TOMASZCYK B. & P. J. MELIA (dir.), *Practical Applications in Language Corpora*. University Press : Lodz, Poland, 51–62.
- BERNARDINI, Silvia. 2000. « Systemizing serendipity: proposals for concordancing large corpora with learners ». In BURNARD L. & T. MCENERY (dir.), *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the Third International Conference on Teaching and Language Corpora*. Peter Lang : Frankfurt am Main.
- BERNARDINI, Silvia. 2001. « 'Spoilt for choice': A learner explores general language corpora ». In ASTON, G. (ed.), *Learning with Corpora*. CLUEB : Bologne, 220–49.
- BOULTON, Alex. 2010. « Learning outcomes from corpus consultation ». In MORENO JAÉN M., F. S. VALVERDE & C. PÉREZ (dir.), *Exploring New Paths in Language Pedagogy: Lexis and Corpus-Based Language Teaching*. Equinox : Londres, 129–44.
- BOULTON, Alex & Tom COBB. 2017. « Corpus use in language learning: A meta-analysis ». *Language Learning* 67/2, 348–93.
- BOULTON, Alex & Corinne LANDURE. 2016. « Using corpora in language teaching, Learning and use ». *Recherche et pratiques pédagogiques en langues de spécialité*. [En ligne], 35/2, mis en ligne le 30 juin 2016. DOI : 10.4000/apliut.5433
- BOULTON, Alex & Henry TYNE. 2013. « Corpus linguistics and data-driven learning: A critical overview ». *Bulletin Suisse de Linguistique Appliquée* 97, 97–118.
- CAZADE, Alain. 2001. « LIED2, Longman Interactive Electronic Dictionary ». *Recherche et pratiques pédagogiques en langues de spécialité*, 21/2, 79–88.
- CHAMBERS, Angela. 2005. « Integrating corpus consultation in language studies ». *Language Learning & Technology* 9/2, 111–125.
- CHAMBERS, Angela. 2010. « L'apprentissage de l'écriture en langue seconde à l'aide d'un corpus spécialisé ». *Revue Française de Linguistique Appliquée* 15/2, 9–20.
- CHAMBERS, Angela & Íde O'Sullivan. 2004. « Corpus consultation and advanced learners' writing skills in French ». *ReCALL* 16/1, 158–172.

- CHARLES, Maggie. 2007. « Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions ». *Journal of English for Academic Purposes* 6/4, 289–302.
- DE COCK, Sylvie & Henry TYNE. 2014. « Corpus d'apprenants et acquisition des langues ». *Recherches en didactique des langues et cultures: Les Cahiers de L'acedle* 11/1, 137–168.
- EVERT, Stefan & Andrew HARDIE. 2011. « Twenty-first century corpus workbench: Updating a query architecture for the new millennium ». In *Proceedings of the Corpus Linguistics 2011 Conference*. University of Birmingham, UK. <<http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/Paper-153.pdf>>.
- FLOWERDEW, Lynne. 2001. « The exploitation of small learner corpora in EAP material design ». In GHADESSY M., A. HENRY & R. L. ROSEBERRY (dir.), *Small Corpus Studies and ELT. John Benjamins : Amsterdam*, 363–80.
- FLOWERDEW, Lynne. 2004. « The argument for using English specialized corpora to understand academic and professional language ». In CONNOR U. & T. A. UPTON (dir.), *Discourse in the Professions: Perspectives from Corpus Linguistics*. Benjamins : Amsterdam, 11–33.
- FRANKENBERG-GARCIA, Ana. 2014. « The use of corpus examples for language comprehension and production ». *ReCALL* 26/2, 128–46.
- FRANKENBERG-GARCIA, Ana. 2015. « Dictionaries and Encoding Examples to Support Language Production ». *International Journal of Lexicography* 28/4, 490–512.
- HARWOOD, Nigel. 2005. « What do we want EAP teaching materials for? ». *Journal of English for Academic Purposes* 4/2, 149–61.
- JOHNS, Tim. 1991. « Should you be persuaded: Two examples of data-driven learning ». *English Language Research Journal* 4, 1–16.
- JOHNS, Tim. 1997. « Contexts: The background, development, and trialling of a concordance-based CALL program ». In WICHMANN A., S. FLIGELSTONE, T. MCENERY, G. KNOWLES & T. JOHNS (dir.), *Teaching and Language Corpora*. Longman : New York, 100–115.
- KENNEDY, Claire & Tiziana MICELI. 2001. « An evaluation of intermediate students' approaches to corpus investigation ». *Language Learning & Technology* 5/3, 77–90.
- KENNEDY, Claire & Tiziana MICELI. 2017. « Cultivating effective corpus use by language learners ». *Computer Assisted Language Learning* 30/1–2, 91–114.
- KOO, Kyosung. 2006. « Effects of Using Corpora and Online Reference Tools on Foreign Language Writing: A Study of Korean Learners of English as a Second Language ». PhD thesis, University of Iowa.
- LANDURE, Corinne. 2014. « Étude comparative de l'exploitation directe de corpus générique et spécifique L2 par des apprenants LANSAD ». Thèse de doctorat, Université Paris 7– Paris Diderot.
- LEE, David & John SWALES. 2006. « A corpus-based EAP course for NNS doctoral students: moving from available specialized corpora to self-compiled corpora ». *English for Specific Purposes* 25/1, 56–75.
- LEECH, Geoffrey. 1997. « Teaching and language corpora: A convergence ». In WICHMANN A., S. FLIGELSTONE, T. MCENERY, G. KNOWLES & T. JOHNS (dir.), *Teaching and Language Corpora*. Longman : New York, 1–23.

- MANGENOT, François. 2000. « Contexte et conditions pour une réelle production d'écrits en ALAO ». *ALSIC* 3/2, 187–206.
- MCCARTHY, Michael. 2008. « Accessing and interpreting corpus information in the teacher education context ». *Language Teaching* 41/4, 563–574.
- MCENERY, Tony & Andrew WILSON. 1997. « Teaching and language corpora ». *ReCALL* 9/1, 5–14.
- MILLOT, Philippe. 2017. « Inclusivity and exclusivity in English as a business lingua franca: The expression of a professional voice in email communication ». *English for Specific Purposes* 46, 59–71.
- MUELLER, Charles M. & Natalia D. JACOBSEN. 2016. « A comparison of the effectiveness of EFL students' use of dictionaries and an online corpus for the enhancement of revision skills ». *ReCALL* 28/1, 3–21.
- O'SULLIVAN, Íde. 2007. « Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy ». *ReCALL* 19/3, 269–286.
- O'SULLIVAN, Íde & Angela CHAMBERS. 2006. « Learners' writing skills in French: corpus consultation and learner evaluation ». *Journal of Second Language Writing* 15/1, 49–68.
- PARK, Kwanghyun, & Celeste KINGINGER. 2010. « Writing/thinking in real time: Digital video and corpus query analysis ». *Language Learning & Technology* 14/3, 31–50.
- PÉREZ-PAREDES, Pascual, María SÁNCHEZ TORNEL, Jose María ALCARAZ CALERO, & Pilar AGUADO JIMÉNEZ. 2012. « Learners' search patterns during corpus-based focus-on-form activities: A study on hands-on concordancing ». *International Journal of Corpus Linguistics* 17/4, 483–515.
- RICHARDS, Keith. 2009. « Trends in qualitative research in language teaching since 2000 ». *Language Teaching* 42/2, 147–80.
- RÖMER, Ute. 2006. « Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments ». *Zeitschrift Für Anglistik Und Amerikanistik* 54/2, 121–34.
- RÖMER, Ute. 2009. « Corpus research and practice: What help do teachers need and what can we offer? » In AIJMER, K. (ed.), *Studies in Corpus Linguistics*. John Benjamins : Amsterdam, 83–98.
- SUN, Yu-Chih. 2003. « Learning process, strategies and web-based concordancers: A case-study ». *British Journal of Educational Technology* 34/5, 601–13.
- SWALES, John. 1997. « English as tyrannosaurus rex ». *World Englishes* 16/3, 373–82.
- TRIBBLE, Christopher. 2001. « Small corpora and teaching writing ». In GHADESSY M., A. HENRY, & R. L. ROSEBERRY (dir.), *Small Corpus Studies and ELT*. John Benjamins : Amsterdam, 381–408.

ANNEXES

Annexe 1 ; Profil des étudiants

Tableau 5. Profil des étudiants

Étudiant	Ressource utilisée	Niveau d'anglais	Langue maternelle	Niveau d'études	Matière étudiée ⁷
1_C_SS	Corpus	B1	Français	M1	SS

2_WR_SS	WR	C1+	Français	M1	SS
3_C_SS	Corpus	C1+	Français	M1	SS
4_WR_SS	WR	C1	Français	M1	SS
5_C_SS	Corpus	C1+	Français	M1	SS
6_C_SS	Corpus	C1	Français	M1	SS
7_C_SS	Corpus	C1	Français	M1	SS
8_WR_SS	WR	C1	Français	M1	SS
9_WR_SS	WR	B1	Français	M1	SS
10_WR_SS	WR	C1+	Français	M1	SS
11_C_SS	Corpus	C2	Français	M1	SS
12_C_GM	Corpus	C1	Français	M2	GM
13_C_GM	Corpus	B2	Français	M2	GM
14_C_GM	Corpus	B2	Arabe	M2	GM
15_C_GM	Corpus	C1	Français	M2	GM
16_WR_GM	WR	B1	Arabe	M2	GM
17_WR_GM	WR	B2	Français	M2	GM
18_WR_GM	WR	C1-	Français	M2	GM
19_WR_GM	WR	B2	Français	M2	GM
20_C_GM	Corpus	C1-	Arabe	M2	GM
21_WR_GM	WR	C1-	Français	M2	GM
22_WR_GM	WR	B1	Arabe	M2	GM
23_C_GM	Corpus	B1	Français	M2	GM
24_C_GM	Corpus	B1	Arabe	M2	GM

Annexe 2 : Journaux de bord des étudiants

NOTES

1. La plateforme a été installée à l'Université Paris Diderot pour des besoins de recherche (CLILLAC-ARP) et d'enseignement (UFR EILA).
 2. < <https://www.theguardian.com/science/2016/sep/29/secret-of-connection-between-dogs-and-humans-could-be-genetic>> (consulté le 4 janvier 2017)
 3. Nous avons consigné l'intégralité des journaux de bord, sans modification, dans le tableau 6 en annexe 2.
 4. Il y a 13 étudiants sur corpus mais l'un d'entre eux n'utilise pas la ressource ; ainsi 12 s'en sont servis.
 5. GN : groupe nominal
 6. Contemporary Written Italian Corpus
 7. SS : Sciences Sociales / GM : Génie Mécanique
-

RÉSUMÉS

L'essor de l'apprentissage sur corpus dans le monde universitaire peine à se refléter dans les pratiques pédagogiques. Ce phénomène est notamment imputable au manque de lien entre recherche et pédagogie. La présente étude cherche à combler un vide informationnel quant à la manière dont les étudiants LANSAD se servent d'un corpus spécialisé. Nous proposons de comparer l'apport d'un corpus spécialisé à celui d'un dictionnaire en ligne dans le cadre de la rédaction de résumés scientifiques. Les données au sujet de l'utilisation des ressources par les étudiants permettent de démontrer que l'outil corpus est efficacement utilisable par tous, même en contexte spécialisé, et qu'il est une source d'apprentissage langagier. Néanmoins, une initiation en amont semble nécessaire afin que les apprenants puissent se servir de cette ressource le mieux possible. À cet effet, des activités d'introduction au corpus sont proposées. Des études similaires à celle-ci sont requises afin de mieux comprendre l'interaction de l'étudiant LANSAD avec le corpus.

The popularity of data-driven learning popularity among academics has yet to reach the classrooms. This phenomenon can be attributed to a gap between research and teaching. This study's aim is to fill an informational void about the way ESP students use a specialized corpus. We compare the advantages of using a specialized corpus when writing academic abstracts vs. using an online dictionary. Data about the students' use of the resources indicate that a corpus can be effectively used by all, even in a specialized context, and that it can develop the linguistic knowledge of its users. A proper initiation is nevertheless necessary to enable the learners to fully exploit this resource adequately; examples of introductory activities to corpora are given. Similar studies are required for a better understanding of the ESP student's relationship to corpora.

INDEX

Mots-clés : apprentissage sur corpus, didactique de l'anglais de spécialité, LANSAD, rédaction de résumés d'articles de recherche, TICE

Keywords : academic abstract writing, data-driven learning, ESP didactics, ICTE

AUTEUR

LAURA-MAY SIMARD

Laura-May Simard, agrégée d'anglais, a donné des cours à des spécialistes d'autres disciplines à l'École normale supérieure Paris-Saclay. Elle est en poste dans le secondaire au lycée Louis-le-Grand à Paris. Elle est membre d'ESPri (ENS Paris Saclay's English for Specific Purposes Research Initiative) pour l'année 2018. Ses recherches portent sur la didactique des langues et l'utilisation des TICE, plus particulièrement dans le secteur LANSAD. laura-may.simard@ac-paris.fr