

## *L'apprentissage sur corpus dans l'enseignement de l'écrit scientifique chez les locuteurs non-natifs en classe de FLE*

*DEMMANE Nadhira*  
*Université om-elbouaghi*

### Résumé

Face à l'acculturation des étudiants locuteurs non-natifs à l'écrit scientifique, notre travail porte sur le genre de l'article de recherche français dans le champ de la linguistique. Ce travail a double objectif, vise à aider les apprenants en classe de langue à mieux identifier les caractéristiques rhétoriques du discours scientifique au sein de leur discipline et à connaître ce qu'est la démarche d'apprentissage sur corpus (ASC). Ce travail se propose de mettre l'accent sur les apports et les limites de l'(ASC) en enseignement des écrits scientifiques. Pour ce faire, nous tentons tout d'abord d'identifier le concept de l'(ASC), ses apports dans l'enseignement /apprentissage des écrits scientifiques en classe de français langue étrangère. Puis, nous examinons les apports et limites de l'usage de l'(ASC), en enseignement des articles de recherche à partir d'une revue de littérature. Enfin, nous proposons quelques activités pédagogiques exploitables en classe de langue à partir du corpus Scientext disponible à l'adresse suivante : [scientext.msh-alpes.fr](http://scientext.msh-alpes.fr).

Mots clés : Écrit scientifique, activités Pédagogiques, plateforme, technologies de l'information et de la communication pour l'enseignement (tice), enseignement /apprentissage.

ملخص

لمواجهة الصعوبات التي يواجهها الطالب في الكتابة العلمية باللغة الفرنسية كلغة ثانية، يركز عملنا هذا على مساعدة المتعلمين للتعرف على الخصائص الأسلوبية للخطاب العلمي من خلال المدونات اللغوية ويهدف هذا العمل إلى التركيز على مجموعة من الأنشطة لتدريس النص العلمي. ولتحقيق ذلك، نحاول أولاً تحديد مفهوم، المدونة التعليمية في تدريس وتعلم الخطاب العلمي. ثم نعرض بعض الدراسات في تدريس النص العلمي بالمدونة وأخيراً، نقترح بعض الأنشطة البيداغوجية التي يمكن استغلالها من النصوص المتاحة على شبكة الانترنت على الرابط : [scientext.msh-alpes.fr](http://scientext.msh-alpes.fr)

الكلمات الدالة: الكتابة العلمية، الأنشطة البيداغوجية، برمجيات، تكنولوجيا المعلومات والاتصالات للتعليم، التدريس / التعلم.

### 1. Introduction

Ces dernières années, l'utilisation des corpus dans l'apprentissage des langues dans divers contextes est de plus en plus recommandée par les spécialistes. Le travail pionnier de (Johns,1991a ;1991b) a été suivi par une explosion d'études consacrées dans différents domaines spécialement en didactique des langues étrangères préconisant l'utilisation d'outils et de techniques de linguistique de corpus à des fins pédagogiques. Les corpus numériques font actuellement leur entrée dans le monde de l'enseignement/apprentissage des langues. Deux numéros spéciaux de revues spécialisés en témoignent : en 2014, ReCALL (Journal of the European association of computer-assisted language learning) a publié un volume appelé Special Issue on Researching New Uses of Corpora for Language Teaching and Learning, et en 2015, LLT (Language Learning & Technology), a lancé un appel concernant le thème Special Issue on : Corpora in Language Learning and Teaching, à paraître en 2017 cité par (Schaeffer-Lacroix E., 2015 :P47).

Cette démarche est venue s'imposer en tant que véritable nouvelle approche et en tant que paradigme de recherche en didactique des langues surtout dans les sphères anglo-saxonnes.

(Johns, 1986). En effet, des spécialistes tels que (Alex Boulton et Henry Tyne, 2014), reconnus de l'utilisation de corpus pour la didactique des langues recommandent aux enseignants et aux apprenants d'exploiter les corpus tout en les sensibilisant à la démarche d'apprentissage sur corpus (ASC) qui leur permettra d'évoluer dans l'enseignement /apprentissage des langues étrangères.

Beaucoup d'Enseignants Algériens refusent de rester enfermé dans le paradigme traditionnel surtout avec l'évolution continue des TICE (technologies de l'information et de la communication pour l'enseignement) et adhèrent à cette démarche innovante en vue de son intégration dans leurs pratiques pédagogiques.

Dans le cadre d'un article assez court, il ne peut être question d'une analyse détaillée, à partir d'un corpus complet. Nous nous limiterons à la présentation de quelques activités pédagogiques par l'ASC afin d'illustrer certains des points les plus évidents et qui nous semblent déterminants dans l'enseignement /apprentissage de l'écrit scientifique en classe de langue aux étudiants de Master.

La question de l'enseignement /apprentissage de l'écrit scientifique par emploi de l'ASC en classe de langue est une thématique qui intéresse actuellement beaucoup de chercheurs et concepteurs de programmes (Chambers A., 2005, 2010/2, Chambers A. et O'Sullivan I., 2004, Rézeau J., 2007, Schaeffer-Lacroix E., 2008 cité Cristelle Cavalla, Mathieu Loiseau, 2013 :p2). Une première analyse de la revue de littérature dévoile la diversité des méthodes et instruments utilisés pour juger l'ASC.

L'ensemble des spécialistes confirment que les apports de cette nouvelle approche dans l'enseignement sont multiples. Dans la revue de littérature sur les apports des corpus de la plateforme Scientext, (Cavalla et Loiseau, 2013 :p3) proposent quelques pistes d'analyse de ce type de corpus. Selon eux, divers aspects principaux doivent être observés, à savoir :

- L'enseignement /apprentissage de la phraséologie transdisciplinaire;
- La découverte du lexique universitaire et scientifique;
- La découverte des différents écrits exigés par le monde universitaire ;
- L'extraction de phrases -exemples pour l'élaboration de séquences ;
- L'exploitation des grammaires sémantiques.

En se situant dans ce cadre, on peut se demander comment aider les apprenants en classe de Français Langue Étrangère(FLE) à mieux identifier les caractéristiques rhétoriques du discours scientifique par emploi de l'apprentissage sur corpus (ASC) ?

#### 1. Qu'est ce que l'apprentissage sur corpus(ASC) ?

La présente étude est basée sur les présupposés théoriques du 'Data-driven learning'dont la forme abrégée en anglais est 'DDL' ou l'apprentissage « conduit ou porté par les données » qui s'est développée au cours des vingt dernières années. Les travaux pionnier de Johns (1986) et Tribble et Jones (1990) ont été suivis par une explosion d'études consacrées à différents aspects de l'utilisation des corpus dans l'apprentissage des langues dans divers contextes. Cette démarche est de plus en plus appliquée en tant que véritable nouvelle approche et en tant que paradigme de recherche en didactique des langues surtout dans les sphères anglo-saxonnes.Le 'Data-driven learning'« DDL » est traduit en français par Henry Tyne en « apprentissage sur corpus » désormais « ASC ». Boulton Alex et Henry Tyne,(2014 :p15) définissent l'acception de l'ASC comme : « Toutes données tirées de l'utilisation explicite d'un corpus à des fins d'apprentissage ou d'aide à l'expression écrite ».En outre l'ASC est conçue dans une perspective socioconstructiviste dont témoignent les exigences définies par le Cadre européen commun de références pour les langues (CECRL), formulées en termes de réalisation de tâches en langue cible .L'utilisateur de cette approche se retrouve devant un apprentissage à partir de documents

authentiques variés . La démarche proposée par l'ASC s'appuie sur cette évolution pédagogique et se fonde sur l'apprenant et sur son autonomisation.

## 2. Définition du terme corpus

Le terme de « corpus » que nous adoptons ici renvoie évidemment à la définition que donne Jean Dubois (1969 :p120), selon l'auteur, un corpus dans sa définition classique, désigne « un ensemble déterminé de textes sur lesquels on applique une méthode définie » Chez les linguistes, la construction du corpus, porte sur les ressources interprétatives des discours en les interprétant pas la médiation des formes linguistiques, et plus largement des formes de l'écriture.

Un corpus dans ce sens est une grande collection raisonnée de textes représentative d'une langue ou d'une variété et consultable sur ordinateur. Quant à Boulton Alex et Henry Tyne, (2014), les corpus représentent une ressource avec de multiples possibilités pour l'apprentissage des langues .Selon les deux linguistes, l'accès direct à des textes écrits et oraux permet de proposer à l'apprenant un « bouquet » choisi de données brutes sur lesquelles il fondera ses observations et la vérification d'hypothèses élaborées dans un contexte de communication. Le corpus est présenté comme un ensemble de textes collectés à des fins précises et interrogeables de manière automatisée. Selon (Fligelstone, 1993) cité par Boulton, 2008 : 38) les corpus ont trois usages :

« teaching about » : enseigner la linguistique de corpus

« exploiting to teach » : exploiter les corpus pour enseigner les langues

« teaching to exploit » : former les apprenants à exploiter les corpus pour apprendre une langue.

Dans notre recherche c'est le dernier usage qui nous intéresse puisque notre objectif principal est bien la formation des apprenants locuteurs non-natifs à l'apprentissage de l'écrit scientifique d'une langue cible.

## 3. Les apports de l'ASC dans l'enseignement de l'écrit scientifique

Depuis plusieurs années, la recherche en didactique des langues étrangères s'intéresse à la question suivante : en quoi l'ASC serait utile pour favoriser la maîtrise du discours scientifique par des non-natifs ? L'accent est mis sur ses avantages et ses apports en enseignement des écrits scientifiques. Nous nous sommes basées sur une revue de littérature des différentes recherches réalisées en ce domaine pour dégager quelques éléments principaux concernant les apports de l'usage de l'ASC en enseignement /apprentissage de l'article de recherche.

### 3.1 Utilisation de l'ASC dans l'enseignement des langues cibles.

L'ASC en éducation a montré un grand potentiel d'améliorer l'apprentissage des élèves s'il est intégré au bon moment et pour la bonne activité, en tant qu'un élément d'une approche logique d'enseignement. Plusieurs recherches importantes ont étudié l'utilité pédagogique des corpus extraits des plates formes (voir, par exemple, Bernardini, 2002; Chambers & O'Sullivan, 2004; Kennedy & Miceli, 2001, 2002 cité par (Cristelle Cavalla, Mathieu Loiseau, 2013). Nous pensons, comme (Henry Tyne, 2012) que les corpus nous renseignent aussi sur le fonctionnement de la langue comme système. De même que (Aston, 1997), confirme que l'analyse de petits corpus à des fins d'apprentissage de langues peut servir de point de départ utile pour les étudiants qui peuvent souhaiter plus tard passer à l'analyse de grands corpus dans un contexte de recherche. Rappelons que les apprenants sont amenés à découvrir et à cerner la langue, ainsi que les différents usages, à travers le contact avec les données généralement réunies sous forme de corpus( Chambers, 2010) .De sa part, (O'Sullivan, 2007) rapporte que le propre du 'data-driven learning', dans son état le plus « pur », est de faire de l'apprenant un chercheur,

capable de s'informer sur le fonctionnement de la langue cible à partir des données brutes. Tout en allant dans le sens des arguments cognitivo-constructivistes.

### 3.2. Utilisation de l'ASC dans l'enseignement de l'écrit scientifique

Dans le panorama scientifique français, des chercheurs ont mis l'analyse du discours scientifique au service des apprentis chercheurs : c'est le cas de Boch, F., Grossmann, F., et Rinck F. (2009), Poudat C. et Rinck F. (2006). Leurs travaux concernant l'écrit scientifique peuvent constituer l'arrière-plan de notre recherche. Dans ce contexte, nous avons développé une recherche pragmatique et inductive portant sur l'exploitation d'une plateforme numérique « Scientext » mise gratuitement en ligne par le laboratoire LIDILEM à la disposition des chercheurs et des étudiants.

### 3.3 L'outil Scientext

Scientext est une plateforme accessible à cette adresse : [scientext.msh-alpes.fr](http://scientext.msh-alpes.fr). Le projet met à la disposition des chercheurs et étudiants un large corpus d'écrits scientifiques de manière à permettre l'étude de leurs caractéristiques linguistiques. Cette plateforme permet de consulter le corpus selon différents critères : requêtes sémantiques, requêtes libres, requêtes avancées. A l'intérieur de ses requêtes a été principalement traitée l'étude linguistique du positionnement et du raisonnement à travers la phraséologie, les marques énonciatives et les marques syntaxiques liées à la causal.

À cette fin, nous proposons de mettre à la disposition de l'enseignant et de l'apprenant un ensemble d'aides susceptibles de faciliter l'accès à la plateforme et ses différents modes d'emplois. Le décodage linguistique peut être renforcé par l'utilisation<sup>1</sup> de la plateforme, qui assure ainsi une reconnaissance automatique des différentes sections d'un article. La structure étant annotée permet d'élargir le champ d'exploitation de l'article, une meilleure interaction entre les sections peut avoir lieu et conduire à une compréhension fine du texte lu. L'apprenant peut ainsi construire une représentation globale de l'article de recherche, étape importante de la compréhension. La plateforme devient alors un outil cognitif et aide à la compréhension dès que l'apprenant la consulte lors de sa lecture. De sa part (Carrell, 1990 : 19) rapporte qu'un texte relevant de la spécialité de l'apprenant est plus facile à lire et à comprendre qu'un texte linguistiquement comparable dans un domaine moins familier.

## 4. Constitution et interrogation du corpus

La constitution du corpus et la sélection des textes ont été menées en fonction des objectifs de cette étude qui sont:

- l'observation des phénomènes linguistiques au niveau d'un genre et d'un type de discours, en l'occurrence à l'intérieur d'un corpus d'articles de revues ;
- l'observation de la distribution de certains faits langagiers à l'intérieur des textes d'une même discipline.

Les objectifs poursuivis dans notre recherche, c'est-à-dire la quête des effets sur les apprenants de FLE d'un corpus d'articles de recherche, extrait de la plateforme Scientext ,adapté selon les nécessités de la spécialité des étudiants, constituerait un moyen pour la diminution des difficultés rencontrées par les apprenants lors de la lecture et l'écriture d'un écrit scientifique dans leur discipline comme l'affirment Christelle Annick Ferraris, Clara Cecilia Uribe Hernández : « nous voulons que nos étudiants commencent à écrire des articles académiques, il est primordial qu'ils en lisent régulièrement pendant leurs études de comprendre et de repérer les choix rhétoriques enracinés dans ces articles ». Comme le confirment certains spécialistes, les faits langagiers spécifiques à un genre et à un type de discours peuvent être observés de manière efficace dans un corpus comportant des documents du même genre portant sur le même domaine (Rinck ,2012).

#### 4.1 Choix du corpus

Nous avons choisi un échantillon pour l'observation de la langue. Cet échantillon est constitué d'articles de linguistique de revues scientifiques en tant que genre, nous proposons une analyse interne du corpus afin de dégager les représentations langagières liées à la linguistique comme discipline de spécialité de nos apprenants. En nous servant de ces critères d'analyse, nous avons pu sélectionner un corpus intéressant et riche en informations et en ressources dont l'exploitation vise à développer une attitude active de la part de l'apprenant à travers l'interactivité.

Le corpus utilisé rassemble douze articles appartenant à la linguistique. Ce corpus est extrait du corpus libre de droits mis à la disposition de la communauté scientifique par le projet ANR SCIENTEXT sous licence Creative Commons intitulé « Corpus et outils pour étudier le positionnement et le raisonnement de l'auteur dans les écrits scientifiques ». Le corpus utilisé, au format XMLTEI, comporte 61 942 occurrences. Les articles sont représentés comme suit : quatre articles de la revue Cahiers de Grammaire, n° 25, quatre articles de la revue Autour des langues et du langage: perspective pluridisciplinaire et quatre articles de la revue Lidil, n° 30. Les corpus à analyser proposent déjà, par leur constitution même, différentes interrogations. En effet, ils sont diversifiés selon plusieurs perspectives.

#### 4.2 Déroulement de l'intervention

La première étape consiste à sélectionner le corpus, selon un ensemble de paramètres, comme indiqué sur la copie d'écran (voir fig1). Nous avons choisi un corpus d'articles de revues appartenant à la linguistique comme discipline de spécialité.

#### 4.3. Exemples d'activités sur le lexique disciplinaire.

Ces activités visent à approfondir, à chercher et à évaluer des unités terminologiques du domaine de la linguistique par le biais du lexique disciplinaire comme le préconisent certains chercheurs « Enseigner le vocabulaire au moment opportun facilite de plus l'appropriation des capacités Langagières » (Meyerson & al., 1991 ; Wilson, 1998) cité par (Laurence Catel, 2001). Les apprenants procèdent à une extraction automatique des termes dans les articles choisis. Les unités terminologiques spécifiques à un genre et à un type de discours peuvent être observées de manière efficace dans un corpus comportant des documents du même genre : l'unité terminologique spécifique à un genre aura tendance à se répéter dans un corpus et par conséquent à se manifester à travers le corpus entier. Le tableau ci-dessous confirme ce que nous venons de dire. Ce genre d'activité vise deux objectifs majeurs :

- Cibler « les éléments récurrents » Alex, 2014 dans ce type de corpus. D'après Firth « Le voisinage » d'un mot est très important, il y'a des préférences, des contraintes qui ne peuvent pas toujours s'expliquer de façon « formelle » ce qu'il appelle le principe « d'idiomaticité ».
- Attirer l'attention des apprenants sur l'utilisation de la phraséologie appropriée lors de la rédaction scientifique. Voici des exemples d'exercices proposés dans cette partie.

##### 4.3.1 Formulation d'une hypothèse

L'enseignant peut proposer à titre d'exemple des exercices portant sur la structure de l'hypothèse ainsi que sur les verbes utilisés pour sa formulation. Une fois la requête lancée, nous pouvons obtenir la réponse suivante (voir figure 2)

##### 4.3.2 Regroupement lexical

Notre objectif à travers cette activité est de regrouper les mots du corpus et des textes en classes. Ces classes réduisent le nombre de mots inconnus et donc le nombre de phénomènes syntaxiques rares ou inconnus, liés au lexique, des textes à analyser. Notre objectif est donc de

proposer des regroupements lexicaux à partir d'informations tirées du corpus et sélectionnés selon les besoins de l'apprenant car selon (Boulton, 2016), une approche sur corpus privilégie l'étude des structures de la langue telles qu'on les conçoit selon la lexico-grammaire (la langue connaît des « préférences de voisinage »;celles-ci peuvent varier selon le genre ou type de production, ces « préférences »sont aussi importantes que les règles formelles de grammaire. A titre d'exemple la plateforme Scientext propose des regroupements lexicaux sur les expressions relatives aux auteurs cités, sur la formulation et validation des hypothèses, sur les verbes de dénomination (voir figure 3)

#### 4.3.3 Identification de la structure globale de l'article

Notre objectif didactique à travers cette activité est d'aider nos étudiants à devenir des lecteurs efficaces, à savoir bien traiter l'information textuelle, lors d'une situation d'apprentissage authentique où la lecture de tout l'article sera le moyen d'arriver au but visé : comprendre les articles de revue scientifiques qui leur sont donnés à lire. Sur le plan didactique (Boulton ,2016) met l'accent sur la question de la découverte de l'apprenant de « règles » non écrites, même si on peut aussi concevoir des exercices « classiques »on peut se servir de corpus pour relever des exemples authentiques.

À cette fin, nous proposons de mettre à la disposition de l'apprenant un ensemble d'aides susceptibles de faciliter la compréhension. Le décodage linguistique peut être renforcé par l'utilisation des articles sélectionnés sur la plate forme Scientext, qui assure ainsi une reconnaissance automatique des différentes sections d'un article. La structure étant rendu compréhensible, une meilleure interaction entre les sections peut avoir lieu et conduire à une compréhension fine du texte lu. L'apprenant peut ainsi construire une représentation globale de l'article, étape importante de la compréhension. La plateforme devient alors un outil cognitif et aide à la compréhension .Les étapes à suivre sont :

- Lire le résumé et relever le plan du texte.
- Relever l'hypothèse.
- Définir le but de l'étude
- Quelle méthodologie est utilisée ?
- Quels sont les principaux résultats (bruts) ?
- Quelles sont les principales conclusions ?
- Quelles sont les ouvertures proposées pour les chercheurs ?

#### 4.3.4 Rubriques et sous-rubriques de l'article

L'enseignant demande aux apprenants de lire rapidement pour trouver les titres des sections du premier et du deuxième article sélectionnés sur la plateforme. Une fois la lecture terminée, il pourra poser les questions suivantes :

- Comment chaque article est-il organisé?
- Quelles sont les principales rubriques et sous-rubriques? Faire de brèves notes.
- Vérifier les réponses dans les articles sélectionnés.
- Comparer les rubriques de l'article sélectionné par l'apprenant et l'article sélectionné d'un collègue. Noter les similitudes et les différences trouvées.

Certaines sections de l'article sont plus faciles à décrire comme les «méthodes» et les «résultats».Cependant les apprenants peuvent rencontrer plus de difficulté avec l'«introduction» et la «discussion» vu que leur vocabulaire est limité. Afin d'améliorer les compétences de compréhension des apprenants, un ensemble d'outils doit être développé Mais pour s'assurer que ces activités s'accompagneront d'un gain en termes d'apprentissage, il faut déployer des moyens

pédagogiques, des réflexions théoriques et des outils d'analyse et de gestion deviennent indispensables.

#### 4.3.5 Analyse lexico-sémantique du corpus

Dans cette activité, nous explorons les caractéristiques de l'article de revue en procédant à une analyse lexico-sémantique d'un corpus constitué de 38 896 mots à partir de huit articles de revues spécialisées en linguistique. Dans l'exercice en question, on attire l'attention des apprenants sur quelques collocations récurrentes dans les articles choisis tout en cherchant s'il ya des phénomènes d'organisation textuelle associés pour la compréhension de ce genre de texte. L'enjeu sera d'élaborer une approche qui permette de lier les questions du genre et le repérage du lexique terminologique, la structure discursive, la mise en relation entre la matérialité linguistique et sémiotique de ce genre de discours. L'apprenant joue le rôle d'expert dans la mesure où c'est lui qui doit tester la plateforme en consultant par exemple la rubrique « Recherche sémantique » (voir figure 4) et repérer des connaissances spécialisées concernant le positionnement de l'auteur dans un article de revue scientifique. La recherche aboutit aux données suivantes :

#### 4.3.6 Les citations des auteurs dans un article de linguistique

Il existe autant de « normes » que d'interlocuteurs. En effet, il n'y a pas de véritable consensus sur les normes de citations en sciences humaines et sociales : elles sont différentes en fonction des disciplines, des publications. Les résultats de la recherche dans les articles en linguistique ont données la forme suivante : Le nom de l'auteur ainsi que l'année indiqués avant la citation et séparés par un espace figure 5. Ni le prénom, ni les initiales ne sont indiqués. Les références des citations dans le texte peuvent donc être différentes de ce qui pourrait être demandé au sein d'autres disciplines et d'autres institutions (voir figure 5).

#### 4.3.7 Lecture compréhension de l'article de revue de recherche

L'objectif visé par cette activité comme le montre Kerr, 2013 cité par Boulton 2016 : « Le travail sur corpus permet de travailler sur un certain nombre de points précis en ce qui concerne les « structures » ». Dans cette activité l'enseignant demande aux apprenants de lire l'ensemble de l'article en suivant la démarche suivante :

- 1-Lire avec un objectif clair à l'esprit. La première lecture devra être faite pour la connaissance de base, et non pas appliquer à une question spécifique?
- 2-Attirer l'attention des apprenants sur les structures du titre, du résumé, et les différentes parties de l'article.
3. Faire repérer le sujet, les principales conclusions, la méthode et la conclusion.
4. Demander aux apprenants de Penser à leur propre orientation et quelle est leur biais vers le sujet, la méthode, la source de publication qui peuvent enrichir leur lecture.
- 5- Demander aux apprenants s'ils connaissent déjà le sujet et les méthodes utilisées et si la source de publication est crédible.
- 6-Faire évaluer la lecture de l'article. S'il existe des erreurs par rapport aux résultats et aux données. si l'article paraît cohérent avec les hypothèses de l'approche qu'il faut.
- 7- Demander aux apprenants de résumer tout l'article sous forme d'un petit texte avec le sujet, les méthodes utilisées et les résultats.
- 8-Évaluer l'exactitude factuelle des résultats cités dans l'article et quelles autres questions peuvent –ils poser sur l'article exploité.

Ce genre d'activité devra se faire à un niveau bien avancé de l'apprentissage.

Dans le cadre des études supérieures, la lecture d'articles scientifiques prend trop de temps et ne permet pas à l'enseignant de se focaliser sur la structure rhétorique liée à ce genre de texte ; en

revanche, l'ASC permet de mieux analyser le contenu et changer le regard des étudiants sur l'exploitation de ces articles dans leur travaux de recherche. On voit que l'ASC, s'il est pris en compte par l'enseignant, peut rendre l'enseignement formel de l'écrit scientifique d'autant plus pertinent.

Conclusion et orientations de recherche.

Nous avons essayé de démontrer comment l'ASC peut être utilisé pour améliorer l'enseignement de l'écrit scientifique des locuteurs non natifs en langue française. Nous avons proposé quelques activités pédagogiques qui peuvent être exploitées par les enseignants. L'objectif principal de cette recherche a été le développement d'une pratique numérique émergente, nous pouvons affirmer que les diverses activités pédagogiques proposent, chacune sous un angle d'approche différent, une contribution tournée vers l'apport de l'ASC en tant qu'outil pour stimuler et renforcer l'Enseignement / apprentissage de l'écrit scientifique en classe de langue, pour offrir de nouvelles perspectives et, en même temps, pour aider les différents acteurs (enseignants et étudiants) à repérer les difficultés relatives au genre scientifique et à les surmonter par l'exploitation adéquate des activités pédagogiques proposées.

Ce travail est encore à un stade embryonnaire et a un long chemin à parcourir. Les corpus, La concordance et les outils ont été mis à la disposition des étudiants et chercheurs du département de français de l'Université de Khenchela. Nous prévoyons d'exploiter d'autres bases de données de corpus plus sophistiquées pour offrir aux apprenants de nouveaux supports pédagogiques qui rendent l'apprentissage de l'écrit scientifique plus attrayant. Nous recommandons également que les enseignants et les apprenants en classe de langue utilisent l'ASC et bénéficier de ce qu'il a à offrir.

Nous suggérons également que les outils de construction des corpus soient raffinés et optimisés pour que les enseignants peuvent consolider leurs compétences numériques et, surtout, pour les apprenants qui doivent développer une bonne capacité d'autonomie pour s'approcher des pratiques numériques comme des outils complémentaires à l'acte pédagogique.

#### Bibliographie

ASTON G., "Small and large corpora in language learning", LEWANDOWSKA-TOMASZCZYK B. et MELIA P. J. (dir), PALC 97: Practical applications in language corpora, Łódź, Pologne, Łódź University Press, 1997, p. 51-62.

Boch, F., Grossmann, F. & Rinck, F. (2007). « Conformément à nos attentes... », ou l'étude des marqueurs de convergence/divergence dans l'article, *Revue Française de Linguistique Appliquée*, XII-2, 109-122  
Grossmann, F., & Boch, F. (2003). Production de textes et apprentissage lexical : l'exemple du lexique de l'émotion et des sentiments, *Repères*, 28, 117-135.

Boulton, Alex et Henry Tyne. 2014. Des documents authentiques aux corpus : démarches pour l'apprentissage des langues. Paris : Didier Coll. Langues et didactique, 309 pages.

BOULTON A., "Esprit de corpus : promouvoir l'exploitation de corpus en apprentissage des langues", *Texte et Corpus*, n°3, 2008, p. 37-46.

CHAMBERS A., "Integrating Corpus consultation in Language studies", *Language Learning & Technology*, n°9 - n°2, 2005, p. 111-125.

CHAMBERS A. et O'SULLIVAN I., "Corpus consultation and advanced learners' writing skills in French", *ReCALL*, n°16-1, 2004, p. 158-172.

Cristelle Cavalla, Mathieu Loiseau. SCIENTEXT COMME CORPUS POUR L'ENSEIGNEMENT. Agnès Tutin; Francis Grossmann. L'écrit scientifique : du lexique au discours. Autour de Scientext, Presses Universitaires de Rennes, pp.163-182, 2013.



- Cristelle Cavalla, Agnès Tutin. Étude des collocations évaluatives dans les écrits scientifiques. Europhras2008, Aug 2008, Helsinki, Finlande. Dubois Jean (1969). « Lexicologie et analyse d'énoncé ». Cahiers de lexicologie 2: 115-126.
- C. Tribble, G. Jones Concordances in the classroom: A resource book for teachers Longman, London (1990)
- Eva Schaeffer-Lacroix. Exploitation de corpus lors de la production d'écrits électroniques en L2". In Texte et corpus, n°3, août 2008. Actes des Journées de la linguistique de Corpus 2007. Lorient : Université de Bretagne-Sud. pp. 47-56. Sommaire des actes.
- Eva Schaeffer-Lacroix, « Analyse de trois systèmes de gestion de corpus pour l'enseignement-apprentissage des langues étrangères », Alsic [En ligne], Vol. 18, n° 1 | 2015, mis en ligne le 20 décembre 2015, Consulté le 26 août 2016. URL: <http://alsic.revues.org/2852>
- FLIGELSTONE S., "Some reflections on the question of teaching, from a corpus linguistics perspective", ICAME journal, n°17, 1993, p. 97-109.
- Henry Tyne. Acquisition d'une langue seconde en milieu naturel : contextes, contacts, enjeux. Les migrants face aux langues des pays d'accueil: acquisition en milieu naturel et formation, Septentrion, pp.288, 2012
- JOHNS T., "Micro-concord: a language learners research tool", System, n°14(2), 1986, p. 151-162.
- Johns, T. (1991a). Should you be persuaded: Two examples of data driven. ELR Journal 4, 1-16, University of Birmingham.
- Johns, T. (1991b). From printout to handout: Grammar and vocabulary learning in the context of data-driven learning. ELR Journal 4, 27-45.
- Joseph Rézeau, «L'apport du concordancier à l'analyse et à la remédiation des erreurs des apprenants dans les forums de discussion en ligne», Alsic [En ligne], Vol. 10, n° 2.2007, document alsic\_v10\_04-pra1, mis en ligne le 30 mars 2007, Consulté le 18 novembre 2017. URL : <http://alsic.revues.org/561>
- Laurence CATEL. Écrire pour apprendre ? Écrire pour comprendre ? État de la question .ASTER\_2001\_33\_17.pdf URL : <http://hdl.handle.net/2042/8775>
- Poudat, C. (2006). Étude contrastive de l'article scientifique de revue linguistique. Thèse de doctorat, Université d'Orléans.
- Rinck, F. (2006). L'article de recherche en Sciences du Langage et en Lettres, Figure de l'auteur et approche disciplinaire du genre. Thèse de doctorat, Université de Grenoble.
- TUTIN A., "Collocations du lexique transdisciplinaire des écrits scientifiques : annotation et extraction des propriétés linguistiques dans la perspective d'une application didactique", Cahiers de l'Institut de linguistique de Louvain, n°Vol. 31, n° 2-4, 2007, p. 247-262.
- TUTIN A. et GROSSMANN F., "Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif", Revue Française de Linguistique Appliquée, n°VII, 2002, p. 7-25.
- VERIN, A. (1993). Didactique des sciences et apprentissages méthodologiques : une recherche conduite à l'Institut National de Recherche Pédagogique. Cahiers de Baulieu, 16, 25-38.

Figures

Figure 1 : Page d'accueil du site Scientext

Figure2 :Formulation d'une hypothèse

no	résultat	hypothèse	Conclusion
2	am going remplace went Afin d'	étayer cette hypothèse	#193 - Article - Linguistique - Développement
3	régularité et la fréquence du verbe anglais. Ces résultats	valider l' hypothèse	#193 - Article - Linguistique - Développement
4	En ce qui concerne le modèle de représentation du texte, les résultats	confirmer nos hypothèses	#221 - Article - TAL - Développement
5	ou d' autres combinaisons appliquées à des corpus différents afin de	confirmer des hypothèses	#225 - Article - TAL - Conclusion
6	notes. Avant de calculer les coefficients Kappa, nous avons	vérifié l' hypothèse	#419 - Article - TAL - Développement
7	pu constater pour la distance et le temps. Ceci	renforce l' hypothèse	#420 - Article - TAL - Développement
8	ou encore d' un type de relation. Il permet de	valider les hypothèses	#421 - Article - TAL - Développement
9	dimension lexicale. Des études restent certes à mener pour	valider l' hypothèse	#423 - Article - Linguistique - Développement
10	un fonctionnement cognitif plus général. Ils tendent en effet à	valider l' hypothèse	#423 - Article - Linguistique - Développement
11	Concernant les aspects pragmatiques du récit, les résultats obtenus	confirment partiellement notre hypothèse	#523 - Article - Linguistique - Développement

Figure 3 : Les verbes d'opinion se rapportant à l'auteur

no	résultat	hypothèse	Conclusion
29	Dans cet article, et Moore, 2000 ). Dans cet article,	proposons nous proposons	#226 - Article - TAL - Développement
30	l' identification des changements de thème. La méthode que	nous proposons	#226 - Article - TAL - Développement
31	fondé sur la récurrence lexicale. L' approche	nous proposons	#226 - Article - TAL - Développement
32	Zock, 2006 ) est une possibilité allant dans ce sens que	nous souhaitons	#226 - Article - TAL - Conclusion
33	il est plutôt " opportuniste " dans le sens où	nous souhaitons	#418 - Article - TAL - Développement
34	période de temps sur laquelle s' échelonnent les dépenses dont	on veut	#418 - Article - TAL - Développement
35	se distinguent, est de 100 m. Ce que	nous voulons	#418 - Article - TAL - Conclusion
36	de trois ordres: Du point de vue pratique,	nous proposons	#418 - Article - TAL - Conclusion
37	", et de certaines récentes campagnes d' évaluation.	Nous proposons	#419 - Article - TAL - Résumé
38	par les conditions expérimentales. Dans la dernière	nous proposons	#419 - Article - TAL - Introduction
39	utilisées dans les campagnes actuelles Dans les	on distingue	#419 - Article - TAL - Développement
40	mesurer l' accord des juges ( Table 2 ),	nous distinguons	#419 - Article - TAL - Développement
41	externes de systèmes de TA et diminuer leur coût ?	Nous proposons	#419 - Article - TAL - Développement
42	liaison avec une tâche réaliste. Dans ce cadre,	nous proposons	#419 - Article - TAL - Développement
43	thème et la version de haute qualité. Par exemple,	on veut	#419 - Article - TAL - Développement
44	européen, en trente ou quarante langues. Ou encore,	on veut	#419 - Article - TAL - Développement
45	. Mesures externes objectives pour la TA de l' écrit	Nous proposons	#419 - Article - TAL - Développement
46	système opérationnel, dans un contexte réel, alors	on veut	#419 - Article - TAL - Développement
47	, cela peut être très utile pour le développement.	Nous proposons	#419 - Article - TAL - Développement
48	les déplacements à effectuer avec le dispositif de	nous proposons	#420 - Article - TAL - Introduction

Figure 4 : Recherche lexico- sémantique

01/11/2017		ScienQuest		Historique Connexion À propos Aide	
Corpus		Recherche Résultats		Sémantique Libre Avancée	
<input checked="" type="checkbox"/>	7	nous avons conduites sont assez hétérogènes, ce qui est	pour nous	une preuve de la souplesse de la plate-forme logicielle Navitexte	#221 - Article - TAL - Conclusion
<input checked="" type="checkbox"/>	8	est la raison pour laquelle nous avons adapté ces principes	pour nous	rapprocher d' une forme des documents plus réaliste tout en conservant	#226 - Article - TAL - Développement
<input checked="" type="checkbox"/>	9	fournis dans les tableaux 2 et 3 , a ,	selon nous	quatre conséquences majeures qui sont généralement peu prises en compte dans	#418 - Article - TAL - Développement
<input checked="" type="checkbox"/>	10	" . 7 6 Nous remercions Josette Rebeyrolle et Ludovic Tanguy	pour nous	avoir prêté les corpus sur lesquels ils ont testé leur	#418 - Article - TAL - Notes de bas de page
<input checked="" type="checkbox"/>	11	lesquels ils ont testé leur repérage automatique de définitions et	pour nous	avoir aussi donné les différents patrons produits lors de cette étude	#418 - Article - TAL - Notes de bas de page
<input checked="" type="checkbox"/>	12	les énoncés définitoires qu' il contient . L' essentiel ,	pour nous	, est donc d' évaluer la possibilité du repérage automatique	#422 - Article - Linguistique - Développement
<input checked="" type="checkbox"/>	13	que nous avons utilisées jusqu'ici . Elle se situe ,	selon nous	, à un second niveau , et constitue un indice	#422 - Article - Linguistique - Développement
<input checked="" type="checkbox"/>	14	membres du groupe Terminologie et Intelligence Artificielle ( TIA )	pour nous	avoir permis d' utiliser un recueil d' articles constitué dans le cadre d'	#422 - Article - Linguistique - Notes de bas de page
<input checked="" type="checkbox"/>	15	, Fiherty-Craig et Benton , 2004 ) . En effet ,	selon nous	, des déficits dans la régulation des émotions ( désinhibition	#523 - Article - Linguistique - Développement
<input checked="" type="checkbox"/>	16	susceptible de permettre la différenciation des outils épistémologiques que,	selon nous	, la pc recouvre indistinctement . 2 phrase canonique dans	#526 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	17	intersubjectives différentes . Ce phénomène énonciatif et discursif complexe mérite	à nos yeux	une attention toute particulière , tant sur le plan théorique	#222 - Article - TAL - Développement
<input checked="" type="checkbox"/>	18	situations réelles de communication . Par ailleurs , il est important	à nos yeux	de pouvoir travailler sur des séquences citationnelles " complètes "	#222 - Article - TAL - Développement
<input checked="" type="checkbox"/>	19	à la tâche , HWER . Nous expliquerons plus loin pourquoi,	à notre avis	, elle est biaisée ( pour qu' on puisse essayer de la	#419 - Article - TAL - Développement
<input checked="" type="checkbox"/>	20	cause d' incompréhension . La seule évaluation objective possible est	à notre avis	une mesure par QCM de compréhension , qu' on ne	#419 - Article - TAL - Développement
<input checked="" type="checkbox"/>	21	tropes est parfois dans les faits difficile à tracer , nous choisissons	pour notre part	de considérer la synecdoque comme une catégorie à part entière	#421 - Article - TAL - Développement

Figure 5 : Citations des auteurs

Corpus		Recherche Résultats		Historique Connexion Aide	
Sémantique Libre Avancée					
#	Contexte gauche	Occurrences	Contexte droit	Texte	
<input checked="" type="checkbox"/>	1	rejoignons les résultats précurseurs et les positions défendues par Louise	Dabène ( 1983 )	dans un cadre didactique : un travail métalinguistique sur la	#193 - Article - Linguistique - Conclusion
<input checked="" type="checkbox"/>	2	, nous avons utilisé les valeurs utilisées par Rumelhart et	McClelland ( 1987 )	élaborées à partir du corpus de Brown , établi par Kucera et	#193 - Article - Linguistique - Notes de bas de page
<input checked="" type="checkbox"/>	3	élaborées à partir du corpus de Brown , établi par Kucera et	Francis ( 1964 )	, et consistant en un million de mots provenant de 15 catégories	#193 - Article - Linguistique - Notes de bas de page
<input checked="" type="checkbox"/>	4	riches en informations sémantiques , on peut citer ceux de	Borillo ( 1996 )	, Cartier ( 1998 ) , Condamines & amp ;	#422 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	5	à paraître ) , Meyer ( à paraître ) ,	Pearson ( 1998 )	, Pascual & amp ; Péry Woodley ( 1995 )	#422 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	6	Pearson ( 1998 ) , Pascual & amp ; Péry	Woodley ( 1995 )	, Rebeyrolle & amp ; Péry Woodley ( 1998 )	#422 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	7	Woodley ( 1995 ) , Rebeyrolle & amp ; Péry	Woodley ( 1998 )	. D' autres travaux abordent ce problème du point de	#422 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	8	bases statistiques . C' est le cas des travaux de	Hearst ( 1998 )	, Morin ( 1999 ) et Séguéla ( 1999 )	#422 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	9	le cas des travaux de Hearst ( 1998 ) ,	Morin ( 1999 )	et Séguéla ( 1999 ) . Les premiers travaux (	#422 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	10	Hearst ( 1998 ) , Morin ( 1999 ) et	Séguéla ( 1999 )	. Les premiers travaux ( linguistiques ) s' en tiennent	#422 - Article - Linguistique - Introduction
<input checked="" type="checkbox"/>	11	verbes lors du repérage automatique constitue l' une des principales difficultés . M.	Derruau ( 1988 )	Précis de géomorphologie . Masson , 7 ème Edition .	#422 - Article - Linguistique - Notes de bas de page