



**HAL**  
open science

## Recension : Eva Schaeffer-Lacroix : Corpus et didactique de l'allemand. La langue à bras le corps.

Elena Berthemet

### ► To cite this version:

Elena Berthemet. Recension : Eva Schaeffer-Lacroix : Corpus et didactique de l'allemand. La langue à bras le corps.. ALSIC - Apprentissage des Langues et Systèmes d'Information et de Communication, 2019, Vol. 22, n° 2, 10.4000/alsic.4082 . hal-03019654

**HAL Id: hal-03019654**

**<https://hal.univ-brest.fr/hal-03019654>**

Submitted on 23 Feb 2023

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

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

# Alsic

Apprentissage des Langues et Systèmes d'Information et de Communication

Vol. 22, n° 2 | 2019

Varia

Analyse de livres

---

## **Analyse de Corpus et didactique de l'allemand. La langue à bras le corps**

ELENA BERTHEMET

<https://doi.org/10.4000/alsic.4082>

**Référence(s) :**

Limoges : Lambert-Lucas. 2019

ISBN 978-2-35935-276-4.

192 pages

ISBN/EAN 978-2-35935-276-4

Broché : 24,00 €

---

### ***Entrées d'index***

**Mots-clés :** analyse

**Keywords:** review

**Rubriques :** Analyse de livres

---

### ***Texte intégral***





Afficher l'image

## 1. Introduction

- 1 Le livre *Corpus et didactique de l'allemand. La langue à bras le corps* est écrit par Eva Schaeffer-Lacroix, maître de conférences en études germaniques à Sorbonne Université – Éspé de Paris. Ses recherches portent sur l'enseignement-apprentissage des langues à l'aide des technologies numériques, en particulier des corpus numériques. Comme l'indique son titre, l'ouvrage est dédié à l'enseignement-apprentissage de l'allemand langue étrangère. Il s'adresse à un public spécialiste d'allemand et aussi à un public de spécialistes d'autres disciplines. L'originalité de cette monographie réside dans son intérêt pour l'enseignement des compétences écrites en classe de langue à l'aide des corpus.
- 2 Le lecteur est guidé par une structure semblable des chapitres, qui se composent comme suit : définition des termes utilisés, aperçu de la littérature, avantages et inconvénients des outils numériques mentionnés, possibilités de leur exploitation dans l'enseignement secondaire et supérieur. Ce fil conducteur permet de retrouver rapidement les informations recherchées. La monographie propose des réponses aux questions suivantes : Quels sont le cadre théorique et les termes utilisés dans l'étude (chapitre 1) ? Quels sont les ressources, les outils et les méthodes de linguistique de corpus qui conviennent à l'enseignement-apprentissage des langues (chapitres 2-4) ? Quelles sont les tâches qui peuvent illustrer telle ou telle méthode ? Avec quels apprenants ? Dans quel contexte (chapitre 5) ? Quels scénarios est-il possible d'élaborer en utilisant les corpus (chapitre 6) ?
- 3 Le livre étant une analyse d'outils riche en recommandations, en critiques ainsi qu'en références, cet article est davantage une synthèse de la recherche menée par son auteure. En effet, dans un souci pédagogique, Eva Schaeffer-Lacroix construit sa présentation de la manière suivante : partant de réflexions théoriques, des points essentiels positifs et négatifs des outils sont rassemblés et des lectures complémentaires sont conseillées.

## 2. Public visé

- 4 Bien que cet ouvrage soit adressé en premier lieu aux enseignants d'allemand langue étrangère, sa portée ne se limite pas à ces derniers. Les analyses de l'auteure sont transférables aux autres langues étrangères ayant pour objectif l'enseignement des compétences écrites à l'aide des corpus. Deux approches sont considérées, (1) *data-driven learning DDL* (terme introduit dans Johns, 1992), autrement dit *corpus-driven* (Tognini-Bonelli, 2001), *directe* (Römer, 2011) ou, généralement parlant, *inductive*, où l'apprenant repère et analyse directement lui-même des phénomènes linguistiques afin de découvrir une règle, et (2) *corpus-based*, approche *indirecte* ou *déductive*, où c'est l'enseignant qui élabore des activités à partir d'une hypothèse ou d'une théorie en utilisant des corpus. Soucieuse de faciliter la lecture, l'auteure propose aussi bien des

références pour un public novice que pour les utilisateurs avertis désireux d'aller plus loin. Ainsi, chaque lecteur a à sa disposition des renvois pouvant lui être utiles, soit pour comprendre de quoi il s'agit, soit pour approfondir ses connaissances en la matière. Enfin, le lecteur trouvera des analyses et des activités se basant sur des corpus de natifs et d'apprenants.

## 3. Présentation générale de l'ouvrage

### 3.1. Introduction

5 Dans l'introduction, Eva Schaeffer-Lacroix commence par décrire deux avantages principaux des corpus : recherches sémasiologiques et recherches onomasiologiques. Ainsi, pour ce qui est du premier type, i.e. recherche par des mots, un apprenant de français langue étrangère peut, grâce aux statistiques<sup>1</sup>, constater que "merci" demande l'emploi de la préposition "pour" s'il est suivi d'un substantif et de la préposition "de" lorsqu'il est suivi d'un verbe. Dans une perspective didactique, la définition suivante de corpus est proposée : "base de données contenant des unités linguistiques attestées écrites, orales ou de type multimédia" (page 7). Parmi de nombreux avantages des corpus sont également citées la possibilité d'effectuer des recherches diachroniques de la fréquence d'un mot à une époque donnée ainsi que celle de faire un choix éclairé entre les formes quasi-synonymes d'un point de vue synchronique.

6 Après avoir nommé les travaux fondateurs ainsi que quelques travaux récents dans le domaine de l'enseignement-apprentissage des langues (pages 9-10), l'auteure constate que les corpus pouvant servir aux perspectives didactiques dans les langues autres que l'anglais sont manquants. En effet, la plupart des colloques et des ouvrages concernent la langue anglaise (Aijmer (ed.), 2009 ; Aston (ed.), 2001 ; Aston, Bernardini & Stewart, 2004 ; Burnard & McEnery, 2000 ; Beeby, Rodríguez Inés & Sánchez-Gijón, 2009 ; Campoy-Cubillo, Bellés-Fortuño & Geo-Valor, 2010 ; les conférences Eurocall ; Gavioli, 2005 ; McEnery, Xiao & Tono, 2006 ; les conférences "Teaching and Language Corpora Conference TaLC" ; Reppen, 2010 ; Zanettin, Stewart & Bernardini, 2003). Partant de ce constat, elle passe en revue les travaux publiés en France en didactique outillée de l'allemand. Après avoir référencé des recherches concernant les méthodes<sup>2</sup> et les outils<sup>3</sup>, elle arrive à la conclusion que l'usage des corpus dans l'enseignement de l'allemand, à l'exception de la langue de spécialité, n'est pas courant. Elle note que le *Cadre européen commun de référence pour les langues, CECR* (Conseil de l'Europe, 2001) ainsi que sa version de 2018 restent muets quant à l'intégration concrète des innovations en cours de langue. L'auteure consacre la fin de l'introduction à expliquer l'intérêt des corpus en didactique des langues (par exemple, la possibilité d'observer des énoncés-modèles) et à souligner les particularités des corpus d'apprentissage (telles que, par exemple, la quantité d'occurrences nécessaires à la compréhension d'un phénomène linguistique). Elle termine par une description des chapitres ainsi que deux remarques utiles au lecteur : un lien étroit entre la théorie et la pratique, et le souci d'une écriture inclusive tout en permettant d'alléger la lecture.

### 3.2. Chapitre 1 – "Repères scientifiques"

7 Le premier chapitre commence par le constat que l'intérêt de l'utilisation des corpus dans l'enseignement-apprentissage des langues étrangères fait débat. Le principal défaut cité par les chercheurs serait la non-rentabilité du temps investi dans les activités liées à l'usage des corpus par rapport aux résultats obtenus, les activités liées aux corpus étant chronophages. La question première est, ici, de savoir comment l'humain accède à la connaissance, ce qui permettrait ensuite de répondre à la question de l'intérêt de l'exploration des corpus dans l'apprentissage des langues. Avec de nombreuses

recherches à l'appui<sup>4</sup>, Eva Schaeffer-Lacroix distingue neuf chemins principaux de l'évolution de la pensée humaine. Après avoir argumenté chaque étape, l'auteure arrive à des conclusions récapitulées dans la deuxième colonne du tableau 1.

**Tableau 1 – Liens entre les corpus et les moyens d'accès à la connaissance.**

<b>Moyen d'accès à la connaissance</b>	<b>Facteurs favorisant l'apprentissage grâce aux corpus</b>
(1) franchissement des obstacles de l'apprentissage	statut herméneutique des informations obtenues par les corpus
(2) développement d'une culture scientifique	soutien du développement cognitif grâce à la comparaison des productions écrites des novices avec celles des experts et aide à la décentration
(3) recours à des experts	mise en situation des apprenants en tant qu'enquêteurs linguistes
(4) évolution grâce à la pensée intuitive	conditions favorables à la reconnaissance, à la combinaison et à l'extrapolation des éléments répétés
(5) réduction de la charge cognitive	mise à disposition de modèles reproductibles et possibilité de varier la taille des lignes
(6) conceptualisation et catégorisation	réflexion sur la langue grâce à l'observation
(7) développement d'une conscience métalinguistique	prise de distance par rapport à sa langue maternelle menant à la conscientisation du fonctionnement de la langue
(8) interaction verbale	soutien de l'apprentissage grâce aux interactions verbales
(9) mise de la pensée en mouvement	questionnements incessants des enquêteurs apprenants

8 Comme on peut le constater d'après la deuxième colonne, l'accent est mis sur l'utilité d'investiguer les corpus en didactique des langues. Les chapitres 2, 3 et 4 répondent à la question de la mise en pratique de l'utilisation des corpus. Ils sont destinés aux enseignants de langues étrangères, en particulier de l'allemand.

### 3.3. Chapitre 2 – "Ressources"

9 L'auteure consacre le deuxième chapitre à l'étude minutieuse des ressources pour les didacticiens germanistes. Elle commence par en mentionner deux spécifiques pour la langue allemande : le Digitales Wörterbuch des Deutschen Sprache DWDS et le Wortschatz-Portal de l'université de Leipzig et deux incluant d'autres langues : le Cosmas II et le Clarin Virtual Language Observatory. Après avoir esquissé les différences principales entre les corpus représentatifs et les corpus réflexifs, l'auteure conclut que les seconds, permettant de mener une réflexion sur le texte, sont plus utiles à des fins d'enseignement-apprentissage d'une langue.

10 Chaque partie commence par une définition des types de corpus (monolingue, parallèle, comparable, d'experts et d'apprenants), avec un ou plusieurs exemples suivis des fonctionnalités disponibles ainsi que de pistes d'exploitation possibles. L'objectif est de porter l'attention non seulement sur les avantages de chacune des ressources, mais aussi sur ses limites. Pour ce qui est des corpus d'apprenants, l'auteure donne des exemples des difficultés que les outils automatiques ont à analyser et à interpréter les erreurs des scripteurs, en particulier les erreurs syntaxiques et sémantiques. Elle souligne la nécessité d'une intervention humaine lors de la constitution des corpus d'apprenants mais exprime un espoir du perfectionnement des fonctionnalités des outils à l'avenir.

### 3.4. Chapitre 3 – "Outils"

- 11 Dans le chapitre 3, Eva Schaeffer-Lacroix propose des réponses aux questions suivantes : (1) Comment créer un corpus ? (2) Comment interroger les données ? (3) Comment visualiser les données ? Dans la lignée du chapitre 2, le lecteur trouve les noms des outils existants, une description de leurs fonctionnalités, ainsi que les avantages, les critiques et des conseils d'utilisation.
- 12 Le but du chapitre est, à partir d'un exemple concret – plus précisément des préverbes séparables en allemand – de montrer quels sont les outils existants qui facilitent le travail de l'enseignant. Le raisonnement se développe autour des questions auxquelles un praticien pourrait faire face. La première et la deuxième, concernant respectivement la création d'un corpus web et d'un corpus réflexif, font référence à plusieurs outils, en particulier à Sketch Engine, dont les avantages suivants sont cités : constitution de corpus web autour d'un thème, possibilité d'insérer des fichiers de formats différents, fonction Copier-Coller, étiquetage automatique, consultation en ligne, partage avec d'autres usagers. L'utilité des corpus annotés (approche déductive) et non annotés (approche inductive) est également mentionnée. La troisième question, celle de la préparation des fichiers, est traitée ensuite. Sont cités quelques outils automatiques permettant d'accélérer le travail de nettoyage des textes. Une fois que ces derniers sont au bon format, il est possible de les annoter. La quatrième partie est consacrée à la présentation des deux langages de modélisation, TreeTagger et XML, qui peuvent être utilisés pour un étiquetage automatique des données. Le tableau 2 recense leurs avantages et inconvénients pour une telle utilisation.

**Tableau 2 – Avantages et inconvénients des outils d'annotation TreeTagger et XML.**

Nom de l'outil d'annotation	Avantages	Inconvénients
<b>TreeTagger</b>	Gratuit, bonnes performances	Annotation parfois erronée
<b>XML</b>	Choix libre du contenu des étiquettes	Nécessité des respecter des conventions afin de pouvoir collaborer

- 13 Les deux points suivants sont la saisie des métadonnées et la création des corpus parallèles. Les sujets étant complexes, des références ainsi que quelques conseils pratiques sont donnés. Enfin, la dernière question posée, celle des droits d'auteur dans le domaine de l'enseignement-recherche est traitée de la même manière, avec de nombreux renvois. En effet, l'auteure fait une présentation que l'on pourrait qualifier de globale, car les limites d'un livre ne permettent pas de rentrer dans les détails de chaque sujet.
- 14 Une fois que le corpus est créé, se pose la question de l'interrogation des données. Parmi une multitude de choix qui se présentent aux utilisateurs, l'auteure sélectionne les outils pouvant convenir tout particulièrement à l'enseignement-apprentissage des langues. Consciente de la complexité de la prise en main de ces outils, surtout par des personnes non ou peu initiées à l'informatique, Eva Schaeffer-Lacroix mentionne également ceux qui sont accessibles à un novice. Elle propose ainsi un aperçu de cinq outils pouvant être utilisés dans une perspective pédagogique, dont le tableau 3 recense les caractéristiques principales.

**Tableau 3 – Avantages et inconvénients des outils d'interrogation.**

Nom de l'outil d'interrogation	Avantages	Inconvénients
<b>Compleat Lexical Tutor</b>	Possibilité de travailler sur des corpus déjà constitués ainsi que sur ses propres corpus Possibilité de lire les définitions des mots <sup>5</sup>	Système d'interrogation limité Absence d'accès aux sources

		Limite fixée à 60 000 mots par fichier téléchargé
<b>Corpuseye</b>	Disponibilité de corpus en treize langues Facilité d'utilisation pour les personnes non initiées à l'annotation des données Accès aux informations détaillées concernant les annotations Disponibilité d'annotations syntaxiques	La totalité des corpus n'est pas libre d'accès Impossibilité de retourner aux sources d'origine
<b>Sketch Engine</b>	Disponibilité de nombreuses options : <ul style="list-style-type: none"> <li>• Élargissement du contexte du mot</li> <li>• Accès aux sources</li> <li>• Enquêtes par formes et par lemmes</li> <li>• Recherches de combinaisons de mots qui ne se suivent pas</li> </ul>	Métatermes peu adaptés aux langues autres que l'anglais
<b>TXM</b>	Nombre de fonctionnalités encore plus important que Sketch Engine : <ul style="list-style-type: none"> <li>• Choix asymétrique du nombre de mots à gauche et à droite</li> <li>• Affichage des propriétés du mot-clé (étiquette, lemme, classe)</li> <li>• Création de tables lexicales</li> </ul> Possibilité de l'étiquetage à l'aide de TreeTagger	
<b>ANNIS</b>	Particulièrement adapté aux corpus d'apprenants Affichage et interrogation des annotations de manière isolée ou combinée Requêtes sur les relations complexes : <ul style="list-style-type: none"> <li>• Référentielles</li> <li>• Anaphoriques</li> <li>• Structure syntaxique d'une phrase</li> </ul> Interopérabilité possible avec des fichiers annotés par d'autres outils	Prise en main difficile

15 La dernière section du chapitre porte sur la visualisation des données. Comme précédemment, l'auteure esquisse les fonctionnalités principales de trois outils : Voyant Tools, Sketch Engine et TXM. Le tableau 4 reprend les points forts et les points faibles soulignés par Eva Schaeffer-Lacroix.

**Tableau 4 – Avantages et inconvénients des outils de visualisation des données.**

Nom de l'outil de visualisation des données	Avantages	Inconvénients
<b>Voyant Tools</b>	Rapidité de l'analyse Interface graphique agréable Convient aux débutants	Peu d'éléments paramétrables
<b>Sketch Engine</b>	Présentation de la fréquence du mot à un endroit précis du corpus Nuage de mots apparaissant dans le même contexte syntaxique	

TXM	Présentation de la fréquence du mot à un endroit précis du corpus (comme dans Sketch Engine) Possibilité : <ul style="list-style-type: none"> <li>• de représenter la distribution de plusieurs formes dans le même graphique</li> <li>• d'afficher des étiquettes contenant des informations métalinguistiques</li> </ul>	Nécessité d'une intervention humaine lors de l'interprétation des courbes
-----	---	---

16 L'accent est mis sur les points ayant un lien direct avec l'utilisation des corpus dans l'enseignement des langues étrangères, à partir de l'exemple allemand.

17 Le choix de faire un classement thématique des outils par les fonctionnalités, ainsi que le fait de naviguer entre les chapitres pour retrouver toutes les informations concernant tel ou tel outil pourraient être critiqués par un lecteur souhaitant avoir une vue globale sur tel ou tel outil. Cependant, l'objectif de l'auteure étant, avant tout, d'aider l'utilisateur à sélectionner les outils les mieux appropriés pour une tâche précise, cette option de classement thématique semble justifiée.

18 Autre réflexion que l'on pourrait faire : plusieurs parties de l'ouvrage, par exemple celles qui décrivent les fonctionnalités de certains outils et les perspectives pédagogiques, s'appuient sur une terminologie qui n'est pas nécessairement connue des enseignants de langues étrangères ("robot d'indexation", page 43 ; "corpus-minute", page 44 ; "fichier non structuré", page 45 ; "espace d'interrogation", page 45 ; "éditeur de texte léger", page 46 ; "interopérabilité", page 57 ou "résonances", page 7).

19 Les sous-chapitres, étant brefs, permettent difficilement d'avoir un aperçu global à un lecteur qui est à la recherche d'un outil pouvant l'aider. Cette brièveté, s'expliquant probablement par le manque d'espace qu'un livre peut offrir, mais également par la volonté de présenter le maximum d'informations pratiques, est cependant compensée par de nombreuses références permettant au lecteur d'explorer le sujet en détail.

### 3.5. Chapitre 4 – "Méthodes"

20 Dans le chapitre 4, des méthodes sont prises comme objet d'étude. Il s'agit plus précisément des calculs statistiques et de la prise en compte du contexte. Concernant les statistiques, il est conseillé de ne pas leur faire une confiance aveugle, mais de savoir les interpréter. Par exemple, l'auteure met en garde contre l'utilisation des valeurs moyennes écartant des élèves produisant une grande quantité d'erreurs.

21 S'appuyant sur les définitions des deux expressions "signe de langue" et "signe de discours" de Charaudeau (2015), l'auteure insiste sur le fait que "afin de donner toute sa place à l'analyse du sens (...), la prise en compte du contexte est nécessaire" (page 65). Elle reprend ensuite quatre types de contextes parmi les six définis par Charaudeau (2015) et montre dans quelle mesure chacun de ces types (linguistique, textuel, paratextuel et situationnel) peut être appliqué à l'exploration de corpus. Ces informations sont réunies dans un tableau où le lecteur trouvera également les fonctionnalités techniques associées à chaque contexte. Il est précisé qu'afin de mener une analyse profonde du sens, il est nécessaire que l'outil permette d'avoir accès au texte complet incluant également les métadonnées. Dans le cas contraire, lorsque seul le cotexte est présenté, l'analyse du signe de langue sera restreinte : malgré la possibilité d'observer ses caractéristiques morpho-syntaxiques et son profil lexical, l'analyse du sens sera limitée. Plusieurs corpus et outils sont brièvement présentés avant de clore le chapitre : ChildLex, Corpuseye, Cosmas II, TXM et Sketch Engine. Pour finir, Eva Schaeffer-Lacroix reprend une illustration de Charaudeau (2015) du lien entre le signe de discours et le signe de langue, et émet une proposition sur son utilisation, afin de "mettre en synergie la production et la réception langagières" (page 69).

### 3.6. Chapitre 5 – "Perspectives pédagogiques"

- 22 En accord avec les chapitres précédents, le chapitre 5 présente diverses activités focalisées sur la production écrite faisant appel aux corpus. Les ressources présentées concernent, comme l'ensemble de l'ouvrage, des spécialistes de langues étrangères ou d'autres disciplines mais, comme il est précisé à la page 19, conviennent plus spécialement aux étudiants se destinant aux métiers administratifs ou médicaux. De nombreuses activités y sont proposées comme, par exemple, l'identification des genres textuels grâce aux phénomènes de fréquence et de segmentation, ainsi qu'une réflexion métalinguistique grâce à l'observation de concordances, de listes de fréquence et de productions des apprenants.
- 23 S'appuyant sur des approches socioconstructivistes de l'apprentissage des langues (Bruner, 1960 ; Vygotski, 1934), qui encouragent "les apprenants à aborder la langue d'une façon flexible, intégrée et personnalisée" (page 73), l'auteure conseille d'alterner les approches "*bottom up*" (de bas en haut, par exemple, l'observation des lignes de concordance) et "*top down*" (de haut en bas, par exemple, le repérage des représentants d'une classe de mots). Ainsi, l'apprenant devient l'acteur de l'apprentissage et l'enseignant est l'accompagnateur ou, autrement dit, le facilitateur. En effet, les concordanciers permettent aux apprenants de "faire des trouvailles, découvrir les contours, passer au scanner, reproduire, comparer et regarder à la loupe" par eux-mêmes (pages 73-90). Chacun des avantages est détaillé dans le sous-chapitre correspondant, suivant le même principe : on débute par une explication du sous-titre (ainsi, pour le premier sous-titre, est expliqué ce qui est compris sous l'expression "faire des trouvailles"), viennent ensuite des activités possibles correspondantes (pouvant être prévues ou non par l'enseignant) ainsi qu'une illustration dans un contexte précis ; enfin, des conseils et des remarques sur ces activités sont donnés.

### 3.7. Chapitre 6 – "Scénarios"

- 24 Le dernier chapitre débute par un rappel de ce qu'est un scénario, la définition suivante est proposée : "Un scénario est un plan d'enseignement-apprentissage dont la construction est motivée par une préoccupation d'enseignement-apprentissage centrale" (page 91). L'auteure précise qu'il peut s'agir de productions de différents degrés de difficulté, allant de la rédaction d'un poème ou d'une nouvelle jusqu'à une discussion métalinguistique, en passant par des jeux de société.
- 25 La discussion se développe autour de cinq scénarios. Le lecteur y trouvera le type de public visé, le niveau de compétence en langue requis, les outils proposés, ainsi que la démarche pédagogique à suivre, allant d'activités préparatoires aux activités d'entraînement. Sont également proposées des analyses de productions écrites ainsi que des activités possibles sur le corpus. Cependant, un scénario d'enseignement-apprentissage étant "lié au contexte institutionnel et situationnel" (page 92), il est précisé qu'il est nécessaire de l'adapter en fonction du public. En ce sens, l'auteure rejoint les préconisations de Meunier (2011).

Si les méthodes de corpus doivent être intégrées dans des activités d'enseignement ordinaires, elles doivent être centrées sur l'apprenant, dépendre du contexte et être liées à la culture. Du temps, de l'attention et de l'écoute sont également essentiels pour promouvoir la littéracie du corpus comme un outil d'autonomisation pour les apprenants et les enseignants.

*If corpus methods are to be integrated in normal teaching activities they must be learner-centered, context-dependent and culture-bound. Time, care and attentiveness are also essential when promoting corpus literacy as empowerment tools for learners and teachers.*

## 4. Conclusion

26 En tant qu'enseignante de français langue étrangère, je voudrais remercier Eva Schaeffer-Lacroix pour ce travail riche en références recensant des recherches internationales et dédié à l'enseignement-apprentissage d'une langue étrangère autre que l'anglais. La langue est plus complexe que ce que les ressources pédagogiques traditionnelles laissent paraître. En ce sens, les corpus sont des matériaux d'apprentissage intéressants.

27 Partant du constat que les corpus ne sont pas suffisamment exploités en tant qu'outils pédagogiques, le livre a pour objectif de mettre en avant les recherches actuelles en matière de création et d'exploitation de corpus dans l'enseignement-apprentissage des langues. Il analyse différents outils numériques pouvant être exploités dans des buts pédagogiques. Tout en soulignant des côtés positifs de ces supports numériques, l'auteure fait état des points qui pourraient poser problème. Ainsi, elle souligne le fait que les outils numériques ne sont que des instruments, d'une part, perfectibles et, d'autre part, à utiliser avec prudence car, seul l'esprit humain est capable d'analyser les données. Malgré les difficultés liées à l'exploitation des corpus, leur potentiel est significatif. Riche en recommandations, le livre guide le lecteur et l'oriente vers les outils correspondant au mieux à ses attentes.

---

### Bibliographie

Les liens externes étaient valides à la date de publication.

Aijmer, K. (ed.) (2009). *Corpora and language teaching*. Amsterdam, Philadelphia : John Benjamins.

ANNIS (2016). A web browser-based search and visualization architecture for complex multilayer linguistic corpora with diverse types of annotation. <https://corpus-tools.org/annis/>

Aston, G. (ed.) (2001). *Learning with Corpora*. Houston : Athelstan.

Aston, G., Bernardini, S. & Stewart, D. (2004). *Corpora and language learners*. Amsterdam, Philadelphia : John Benjamins.

Audin, L. (2008). "Langues étrangères à l'école primaire : recherches INRP 1988-2003". In Chini, D. & Goutéraux, P. (éd.). *Psycholinguistique et didactique des langues étrangères – Travaux du Geped en hommage à Danielle Bailly*. Paris : Ophrys. pp. 143-154.

Bachelard, G. (1934). *La formation de l'esprit scientifique – Contribution à une psychanalyse de la connaissance objective*. Paris : Vrin, 1967.

Beeby, A., Rodríguez I. & Sánchez-Gijón, P. (2009). *Corpus use and translating. Corpus use for learning to translate and learning corpus use to translate*. Amsterdam, Philadelphia : John Benjamins.

Bereiter C. & Scardamalia, M. (1983). "Does learning to write have to be so difficult?". In Freedman, A., Pringle, I. & Yalden, J. (eds.). *Learning to write: First language, second language*. New York : Longman.

Bruner, J. S. (2002). *Comment les enfants apprennent à parler*. Paris : Retz.

Bernardini, S. (2004). "Corpora in the classroom. An overview and some reflections on future developments". In Sinclair, J. (ed.). *How to Use Corpora in Language Teaching*. Amsterdam : John Benjamins. pp. 15-36.

Burnard, L. & McEnery, T. (2000). *Rethinking language pedagogy from a corpus perspective. Papers from the third international conference on teaching and language corpora*. Frankfurt am Main, Berlin, Bern : Peter Lang.

Bigot, V. (2016). "L'analyse des interactions didactiques dans la formation initiale des enseignants de français langue seconde". *Communiquer*. pp. 5-27. <http://communiquer.revues.org/2006>  
DOI : 10.4000/communiquer.2006

Bronckart, J.-P. (1997). *Activité langagière, textes et discours. Pour un interactionisme socio discursif*. Neuchâtel : Delachaux et Niestlé. <https://archive-ouverte.unige.ch/unige:37758>

Bruner, J. (1960). *The process of Education. A Landmark in Educational Theory*. Cambridge, MA, London : Harvard University Press, 2002.  
DOI : 10.1119/1.1969598

- Campoy-Cubillo, M. C., Bellés-Fortuño, B. & Geo-Valor, M. L. (2010). *Corpus-based approaches to English language teaching*. London, New York : Continuum.
- Castellotti, V. (2015). "Diversité(s), histoire(s), compréhension... Vers des perspectives relationnelles et altdidactiques pour l'appropriation des langues". *Recherches en didactique des langues et des cultures (RDLC)*, vol. 12, n° 1. pp. 293-331. <https://journals.openedition.org/rdlc/420>
- Chanquoy, L., Tricot, A. & Sweller, J. (2007). *La charge cognitive*. Paris : Armand Colin.
- Charaudeau, P. (2015). "De la linguistique de la langue à la linguistique du discours, et retour". In Engwall, G. & Fant, L. (eds.). *Festival Romanistica. Contribuciones lingüísticas – Contributions linguistiques – Contributi linguistici – Contribuições linguísticas*. Stockholm Studies in Romance Languages. Stockholm : Stockholm University Press. pp. 3-12.
- ChildLex. A lexical database of German read by children. <https://www.ncbi.nlm.nih.gov/pubmed/25319039>
- Clarín Virtual Language Observatory. <https://vlo.clarin.eu/>
- Cobb, T. (2006). "Constructivism, applied linguistics, & language education" In Brown, K. (ed.). *Encyclopedia of language and linguistics. Foundations of linguistics, vol. 32*. Oxford : Elsevier. pp. 85-88.
- Collins Concise English Dictionary*. <https://www.collinsdictionary.com/dictionary/english>
- Compleat Lexical Tutor. For data-driven language learning on the Web. <https://www.lex Tutor.ca/>
- Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues – Apprendre, enseigner, évaluer*. Paris : Didier.
- Conseil de l'Europe (2018). *Cadre européen commun de référence pour les langues – Volume complémentaire avec de nouveaux descripteurs*. <https://rm.coe.int/cecr-volume-complementaire-avec-de-nouveaux-descripteurs/16807875d5>
- Cosmas II. Corpus Search, Management and Analysis System. <http://www.ids-mannheim.de/cosmas2/>
- Corpuseye. <https://corp.hum.sdu.dk/>
- Dictionary.com <https://www.dictionary.com/browse/cat>
- Digitales Wörterbuch des Deutschen Sprache DWDS. Der deutsche Wortschatz von 1600 bis heute. <https://www.dwds.de/>
- Ericsson, K. A. & Simon, H. A. (1987). "Verbal reports on thinking". In Faerch, C. & Kasper, G. (eds.). *Introspection in Second Language Research*. Philadelphia : Multilingual Matters, pp. 24-54.
- Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam : John Benjamins. DOI : 10.1075/scl.21
- Gombert, J.-É. (1996). "Activités métalinguistiques et acquisition d'une langue". *Acquisition et interaction en langue étrangère*, n° 8. pp. 41-55. <https://journals.openedition.org/aile/1224> DOI : 10.4000/aile.1224
- Hulstijn, J. (2007). "Psycholinguistic perspectives on second language acquisition". In Cummins, J. & Davison, C. (eds.). *The international handbook on English language teaching*. Norwell, MA : Springer. pp. 701-713.
- Johns, T. (1988). "Whence and Whither Classroom Concordancing?". In Bongaerts, T., de Haan, P., Lobbe, S. & Wekker, H. (eds.). *Computer Applications in Language Learning*. Dordrecht : Foris. pp. 9-27.
- Johns, T. (1991). "Should you be persuaded—two samples of data-driven learning materials". In Johns, T. & King, P. (eds.). *Classroom concordancing. English language research journal*, n° 4. pp. 1-13.
- Leblay, C. (2014). "Les écritures intermédiaires réflexives en littérature avancée". *Le Français Aujourd'hui – Pratiques de l'écrit en formation*, n° 184. pp. 103-115.
- Mangenot, F. & Nissen, E. (2016). "Éditorial". In Mangenot, F. & Nissen, E. (éd.). *Scénarios pédagogiques et interactions en ligne – CJLT RECAT*, vol. 42, n° 2. <https://www.cjlt.ca/index.php/cjlt/article/view/27474/20221>
- Matthey, M. (2010). "Interaction : lieu, moyen ou objet d'acquisition ?". In Vargas, C., Calvet, L.-J., Gasquet-Cyrus, M., Véronique, D. & Vion, R. (éd.). *Langues et sociétés : Approches sociolinguistiques et didactiques*. Paris : L'Harmattan. pp. 31-42.
- Mayaffre, D. (2010). "Corpus et web-corpus. Réflexion sur la corporalité numérique". *Cahiers de praxématique*, vol. 54-55. pp. 233-248. DOI : 10.4000/praxématique.1170
- McEnery, T., Xiao, R. & Tono, Y. (2006). *Corpus-based language studies*. Oxford : Routledge.
- Merriam-Webster*. <https://www.merriam-webster.com/>

- Meunier, F. (2011). "Corpus linguistics and second/foreign language learning: exploring multiple paths". *Revista Brasileira de Linguística Aplicada*, vol. 11, n° 2. <http://dx.doi.org/10.1590/S1984-63982011000200008>  
DOI : 10.1590/S1984-63982011000200008
- Murison-Bowie, S. (1993). *MicroConcord Manual*. Oxford : Oxford University Press.
- Narcy-Combes, J.-P. (2005). *Didactique des langues et TIC – Vers une recherche-action responsable*. Paris : Ophrys.
- Paas, F., Renkl, A. & Sweller, J. (2003). "Cognitive Load Theory and Instructional Design: Recent Developments". *Educational Psychologist*, vol. 38, n° 1. pp. 1-4.  
DOI : 10.1207/S15326985EP3801\_1
- Pastré, P., Mayen, P. & Vergnaud, G. (2006). "La didactique professionnelle". *Revue française de pédagogie*, n° 154. pp. 145-198. <http://rfp.revues.org/157>  
DOI : 10.4000/rfp.157
- Pekarek Doehler, S. (2000). "Approches interactionnistes de l'acquisition des langues étrangères : concepts, recherches, perspectives". *Acquisition et interaction en langue étrangère*, n° 12. <http://aile.revues.org/934>
- Piaget, J. (1937). *La construction du réel chez l'enfant*. Neuchâtel : Delachaux et Niestlé, 1967.
- Pinto, M.-A. & El Euch, S. (2015). *La conscience métalinguistique – Théorie, développement et instruments de mesure*. Québec : Presses de l'Université Laval.
- Poudat, C. & Landragin, F. (2017). *Explorer un corpus textuel – Méthodes – pratiques – outils*. Louvain-la-Neuve : De Boeck Supérieur.
- Rastier, F. (2004). "Enjeux épistémologiques de la linguistique de corpus". *Texte !* [http://www.revue-texto.net/Inedits/Rastier/Rastier\\_Enjeux.html](http://www.revue-texto.net/Inedits/Rastier/Rastier_Enjeux.html)
- Reppen, R. (2010). *Using corpora in the language classroom*. New York, Cambridge, Melbourne : Cambridge University Press.
- Römer, U. (2011). "Corpus research applications in second language teaching". *Annual review of applied linguistics*, vol. 31, pp. 205-225.  
<https://pdfs.semanticscholar.org/e29a/dc27e2a158ef1d95a0a1d2a86838b46bc4aa.pdf>  
DOI : 10.1017/S0267190511000055
- Sketch Engine*. Learn how language works. <https://www.sketchengine.eu/>
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam, Philadelphia : John Benjamins.  
DOI : 10.1162/coli.2002.28.4.583a
- TreeTagger*. A part-of-speech tagger for many languages. <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>
- TXM*. <http://textometrie.ens-lyon.fr/>
- Trévisse, A. (1994). "Représentations métalinguistiques des apprenants, des enseignants et des linguistes : un défi pour la didactique". *Bulletin VALS-ASLA* (Association suisse de linguistique appliquée), n° 59. pp. 171-190. [https://doc.rero.ch/record/23104/files/Trevisse\\_Anne\\_-\\_Repr\\_sentations\\_m\\_talinguistiques\\_des\\_apprenants\\_20110527.pdf](https://doc.rero.ch/record/23104/files/Trevisse_Anne_-_Repr_sentations_m_talinguistiques_des_apprenants_20110527.pdf)
- Tyler, A. E. & Ortega, L. (2018). "Usage-inspired L2 instruction through three lenses". In Tyler A. E., Ortega, L., Uno, M. & Park, H. (eds). *Usage-inspired L2 instruction—Researched pedagogy*. Amsterdam, Philadelphia : John Benjamins. pp. 3-26.
- Vergnaud, G. (1989). "La formation des concepts scientifiques. Relire Vygotski et débattre avec lui aujourd'hui". *Enfance*, vol. 42, n° 1. pp. 111-118. [https://www.persee.fr/doc/enfan\\_0013-7545\\_1989\\_num\\_42\\_1\\_1885](https://www.persee.fr/doc/enfan_0013-7545_1989_num_42_1_1885)  
DOI : 10.3406/enfan.1989.1885
- Vygotski, L. S. (1934). *Pensée et langage*. (Trad. par Sève, F.). Paris : La Dispute, 1997.
- WordReference*. <https://www.wordreference.com/fr/>
- WordReference Random House Learner's Dictionary of American English*. <https://forum.wordreference.com/threads/wordreference-random-house-learners-dictionary-of-american-english.3263298/>
- Wortschatz-Portal*. <http://wortschatz.uni-leipzig.de/de>
- Zanettin, F., Stewart, D. & Bernardini, S. (eds) (2003). *Corpora in Translator Education*. New York, London : Routledge.  
DOI : 10.4324/9781315760162

---

## Notes

1 Il est précisé que les statistiques sont non exemptes d'erreurs.

2 Tels que l'enseignement programmé, les formations hybrides, la simulation globale et le tandem.

3 Sont cités les outils suivants : la vidéo, Internet, les cédéroms, les baladeurs, les outils d'exploration de corpus, les outils sociaux du web 2.0, les jeux vidéo sérieux.

4 Les travaux contemporains (Audin, 2008 ; Bereiter & Scardamalia, 1983 ; Bernardini, 2004 ; Bigot, 2016 ; Bronckart, 1997 ; Castellotti, 2015 ; Chanquoy et al., 2007 ; Cobb, 2006 ; Gombert, 1996 ; Ericsson & Simon, 1987 ; Hulstijn, 2007 ; Johns, 1988, 1991 ; Leblay, 2014 ; Mangenot & Nissen, 2016 ; Matthey, 2010 ; Mayaffre, 2010 ; Murison-Bowie, 1993 ; Narcy-Combes, 2005 ; Paas, Renkl & Sweller, 2003 ; Pastré, Mayen & Vergnaud, 2006 ; Pekarek Doehler, 2000 ; Pinto & El Euch, 2015 ; Poudat & Landragin, 2017 ; Rastier, 2004 ; Trévisse, 1994 ; Tyler & Ortega, 2018 ; Vergnaud, 1989) sont mis en rapport avec les classiques (Bruner, 1960, 1967 ; Piaget, 1937, Vygotski, 1934 et Bachelard, 1934).

5 Si, en allemand, comme le dit l'auteure, un seul dictionnaire, celui de *WordReference* est proposé, les définitions de deux dictionnaires en anglais sont disponibles : le *WordReference* et le *Collins Concise English Dictionary*, ainsi qu'un renvoi vers le *Merriam-Webster* et le *Dictionary.com*. Comme on peut le constater, grâce au nombre de ressources et à la disponibilité de la version anglaise et américaine, l'anglais est richement documenté.

---

## **Pour citer cet article**

### *Référence électronique*

Elena Berthemet, « Analyse de *Corpus et didactique de l'allemand. La langue à bras le corps* », *Alsic* [En ligne], Vol. 22, n° 2 | 2019, mis en ligne le 28 décembre 2019, consulté le 23 février 2023. URL : <http://journals.openedition.org/alsic/4082> ; DOI : <https://doi.org/10.4000/alsic.4082>

---

## **Auteur**

### **Elena Berthemet**

Enseignante contractuelle à l'université de Bretagne Occidentale. Ses travaux de recherche sont engagés dans deux directions : linguistique générale et didactique des langues étrangères (didactique du vocabulaire phraséologique, corpus et apport qu'ils peuvent avoir dans l'enseignement-apprentissage du lexique, interculturalité, lexicographie didactique).

**Affiliation** : Chercheur associé au Centre de Linguistique en Sorbonne (CeLiSo – EA 7332), Sorbonne Université et Héritages et Constructions dans le Texte et l'Image (HCTI – EA 4249), université de Bretagne Occidentale, Brest.

**Courriel** : [elenaberthemet@gmail.com](mailto:elenaberthemet@gmail.com)

**Adresse** : Université de Bretagne Occidentale, 20 rue Duquesne, 29200 Brest, France.

---

## **Droits d'auteur**



Creative Commons - Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International - CC BY-NC-ND 4.0

<https://creativecommons.org/licenses/by-nc-nd/4.0/>