L'impact des corpus numériques sur la transmission et l'évaluation des connaissances en didactique des langues

Clive HAMILTON
CLILLAC-ARP URP 3967, Université Paris Cité

Résumé : Cette contribution se présente sous la forme d'un retour d'expérience et d'une réflexion critique dans le cadre d'une formation en langue de spécialité. Nous partons du constat que les corpus numériques ne sont pas encore pleinement intégrés dans les pratiques didactiques contemporaines, et ce, malgré une progression conséquente depuis les années 1980. Trois mouvements didactiques orientés vers l'utilisation progressive des corpus ainsi que le résultant changement de paradigme sur la co-construction et la transmission du savoir sont présentés. L'impact de ce changement de paradigme sur le rôle des enseignants et des étudiants dans les cours de langue aussi bien que l'évaluation de ces derniers soient également abordés.

Mots-clés : corpus numériques, langue de spécialité, co-construction du savoir, médiation, transmission

Abstract : Digital corpora have not been fully integrated into contemporary teaching practices, despite significant progression since the 1980s. Using a context of specialised language training as a case in point, this chapter describes three movements indicative of progressive corpus use and discusses the resulting paradigm shift on shared knowledge construction and transmission. The impact of this paradigm shift on teacher and student roles as well as the issue of language evaluation are also addressed.

Keywords : digital corpora, specialised language, shared knowledge construction, mediation, transmission

1 Démystifier le numérique

Qu'il soit pensé en termes informatiques comme média, plateforme, support ou dispositif, l'intégration du numérique au sens large n'est pas un phénomène récent dans les pratiques pédagogiques de l'enseignement supérieur (Dejean-Thircuir & Nissen, 2013 ; Demaizière, 2007 ; Howard & Mozejko, 2015 ; Qotb, 2019 ; Soubrié, 2020). Cependant, la façon dont le numérique a réussi à s'immiscer dans de nombreux domaines disciplinaires et à se rendre indispensable nous oblige à réfléchir collectivement à cette perméabilité. De plus, comme en témoigne le recours massif à l'enseignement à distance

avec ses corollaires, à l'échelle mondiale, en raison de la covid-19, il nous parait d'autant plus judicieux d'identifier et de délimiter l'influence incommensurable qu'exerce le numérique sur nos pratiques pédagogiques actuelles : et ce, de manière à ne pas subir le changement mais à l'anticiper et à le mettre à profit de manière optimale.

Cette contribution se présente sous la forme d'un retour d'expérience et d'une réflexion critique. Nous proposons ainsi de réfléchir à la manière dont le numérique s'est immiscé dans les cours de langue et, en particulier, à la manière dont le recours aux corpus et aux outils d'interrogation de corpus est devenu une niche pour certains enseignants et un objet de méfiance ou de désintérêt pour d'autres (Debaisieux, 2009 ; Ciekanski, 2014). Ces deux perspectives diamétralement opposées ont des incidences directes et indirectes sur l'avancement des théories et pratiques pédagogiques actuelles et entraînent, par conséquent, des changements qui méritent que nous nous y attardions. Parmi les aspects concernés par ces changements, nous retrouvons la participation active des étudiants dans la découverte du savoir et les modalités d'évaluation de ces mêmes étudiants.

Après une brève désambiguïsation terminologique, nous aborderons les tenants et aboutissants du numérique vis-à-à-vis des points déjà évoqués et nous tacherons de répondre aux questions suivantes.

- Quelles sont les nouvelles pratiques et compétences nécessaires à l'utilisation des corpus ?
- Quelle(s) incidence(s) sur la transmission et co-construction du savoir ?
- Quelle(s) incidence(s) sur les rôles et rapports entre étudiants et enseignants ?
- Quelle(s) incidence(s) sur les pratiques d'évaluation ?

Nous nous appuierons également sur une des formations (voir section 2.1) qui est à l'origine de notre réflexion critique.

1.1 Définition et usages du numérique en didactique des langues

Le caractère polysémique dont jouit le terme « numérique » se manifeste dès lors que l'on cherche à le définir. Une consultation rapide dans un dictionnaire et une base de données terminologiques, respectivement le Trésor de la Langue Française informatisée (TLFI) et Termium Plus, révèle plus de six acceptions possibles, chacune liée à un champ d'application différent. Or, aucune ne correspond réellement au sens généraliste, tel que le terme est employé en didactique des langues. Demaizière & Grosbois (2014) et Grosbois (2015) relèvent cet imbroglio terminologique mais ne proposent pas de définition uninotionelle, au sens de Pavel & Nolet (2001), et optent pour des acceptions « très englobantes ». Cette difficulté à identifier les traits distinctifs et à proposer une définition claire suggère une certaine généricité.

Ce cheminement conceptuel n'est pas anodin puisque l'ambiguïté terminologique se manifeste dans les publications traitant d'une part le numérique en didactique des langues sous l'appellation technologies de l'information et de la communication pour

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

l'enseignement (TIC/TICE) et l'apprentissage des langues médiatisé par les technologies (ALMT). Or, les définitions englobantes mises en avant par les auteures ci-dessus laissent penser que ces deux catégorisations relèvent bien du numérique au service de la didactique. De ce fait, la multiplicité des appellations interpelle et il nous parait donc légitime de nous demander s'il s'agit d'un champ avec un objet d'étude identique.

Certains, comme Guichon (2012) et Soubrié (2020), retracent l'évolution de l'acronyme ALMT, à travers des changements successifs de positionnements épistémologiques, mettant en avant le pouvoir de la précision terminologique. En effet, ces derniers soutiennent que le changement, par ordre chronologique, de l'enseignement assisté par ordinateur (EAO), l'apprentissage des langues assisté par Ordinateur (ALAO), l'apprentissage des langues médiatisé par les technologies (ALMT) s'accompagne d'une certaine précision et d'un élargissement de l'objet d'étude au fur et à la mesure que de nouveaux composants numériques intègrent l'environnement didactique. Cet élargissement s'accompagne systématiquement d'un regain d'intérêt par les chercheurs souhaitant examiner l'apport de chacun d'entre eux.

Partant de ce constat, il convient d'affirmer que la conception d'un terme, à l'égard de la formulation et de la définition, a une incidence prépondérante sur les usages, et que les changements nominatifs suggèrent un dépassement de la vision de l'ordinateur comme matériel unique ou central de la médiation du savoir entre enseignant et étudiant. Autrement dit, les appellations consécutives font apparaitre une perspective distincte entre matérialité séparée et matérialité intégrée : il est question d'apprentissage avec un outil qui présuppose une certaine complémentarité didactique d'une part ou à travers un dispositif qui traduit une instrumentalisation didactique d'autre part. Nous maintenons alors que la définition et les usages du numérique sont directement liés à la conception épistémologique à laquelle s'inscrit l'enseignant. Par conséquent, les potentialités du numérique renvoient à l'ensemble des dispositifs, au sens large, exploités dans le cadre didactique : qu'il s'agisse de ce qui est pensé initialement pour une exploitation pédagogique ou de ce qui est didactisé pour des besoins spécifiques.

Ainsi remarquons-nous l'existence d'une littérature et d'un terrain d'étude riches et foisonnants sur chaque élément numérique au service de la didactique des langues, tant vis-à-vis des usages que des usagers (en l'occurrence ici les enseignants et les étudiants). Nous souhaitons de ce fait apporter une contribution au débat actuel en nous intéressant aux corpus numériques et à l'impact qu'ont ces derniers sur la transmission du savoir en didactique des langues.

1.2 Numérique et corpus

Comme nous l'avons soutenu, en raison de son caractère générique, voire hyperonymique, aborder le numérique en didactique des langues ouvre un vaste champ de possibilités. Il semble avoir une centration sur « l'outil » que représentent l'ordinateur et l'Internet, présenté comme une panacée didactique, dans laquelle l'accès à la langue et la didactisation des médias jouent un rôle clé dans la conceptualisation pédagogique. En effet, il est désormais question en contexte institutionnel d'allier et d'explorer des avancées technologiques ainsi que les tendances et pratiques numériques évolutives des

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

usagers : l'objectif ultime est de prendre en compte ces avancées, de manière ludique, dans un esprit de modernisation continue et constante des formations en langue.

Nous pouvons maintenant aborder l'intégration progressive des corpus, parmi les nombreuses avancées technologiques qui ont été intégrées dans la didactique des langues. En effet, soulignons d'emblée que l'intégration des nouvelles technologies en didactique est synonyme d'une accessibilité accrue à la technologie dans la population au sens large. L'avènement du corpus peut être comparé à un changement de paradigme et également à une certaine démocratisation dans les théories linguistiques, dans lesquelles l'importance du contexte et de l'utilisation des faits de langue authentiques ont été démontrés. Ce changement dans la recherche linguistique s'est rapidement avéré fructueux et, par conséquent, s'est progressivement infiltré dans les cours de langue et les approches pédagogiques qui en découlent : ces dernières mettant l'accent sur le fait d'opter pour des données linguistiques attestées, où et quand cela est possible. S'il faut reconnaître que les approches pédagogiques dans ces formations sont aussi variées que les théories linguistiques desquelles elles s'inspirent, l'appel à l'inclusion de corpus – quelle que soit la taille des données utilisées – trouve également un écho chez de nombreux didacticiens qui n'adhèrent à aucun dogme pédagogique ou linguistique. En dépit de ce constat, nous tentons de regrouper et de résumer succinctement ce que nous appelons des mouvements didactiques orientés vers l'utilisation progressive des corpus.

Tab. 1 : Mouvements didactiques intégrant des corpus numériques

Mouvement	**Usager**	**Phase de l'utilisation des corpus**			
		Compilation	Interrogation	Didactisation	Transmission
N°1	Enseignant	optionnelle	optionnelle	active	active
N°2	Enseignant	optionnelle	active	active	active
	Etudiant	N/A	sélective	N/A	N/A
N°3	Enseignant	active	active	active	active
	Etudiant	active	active/sélective	N/A	sélective

Mouvement n° 1. L'avènement de l'utilisation de corpus en didactique des langues peut être retracé dans les années 1980 et 1990 (Boulton & Tyne, 2014 ; Frankenberg-Garcia, 2014 ; Higgins & Johns, 1984). A cette époque, l'accessibilité des corpus est un problème majeur en termes de données et de méthodes d'analyse, cette dernière étant restreinte à certains initiés des milieux linguistiques. Dès lors, l'utilisation des corpus est réservée à ceux qui ont reçu une formation linguistique explicite sur le maniement des corpus. Qu'ils aient exploré des données préalablement compilées ou qu'ils les aient compilées eux-mêmes, les résultats obtenus à partir des interrogations de corpus sont didactisés et mis à profit dans le matériel pédagogique sous forme de règles descriptives (par opposition à prescriptives) à apprendre, d'exercices, etc. Cette période voit l'arrivée de résultats d'analyses de corpus destinés aux étudiants et présentés sous diverses formes pour n'en citer que quelques-uns, les dictionnaires (*cf.*

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

COBUILD), les grammaires et les manuels scolaires basés sur corpus. De nombreuses approches didactiques sont issues de ce mouvement ou en ont hérité directement : par exemple, l'approche bien connue de « *Data-driven learning* » ou apprentissage sur corpus proposée par Johns (1991). De plus, malgré un démarrage relativement lent (et le décloisonnement progressif de la niche identifiée chez les linguistes), cette étape a connu une avancée considérable au cours des 30 dernières années et est désormais parmi la plus reconnue par ceux qui utilisent du matériel pédagogique obtenu à partir d'interrogation ou d'exploration de corpus.

Mouvement n° 2. L'accessibilité et l'utilisation croissantes des technologies numériques en classe ont ouvert la voie à l'intégration de tâches d'exploration de corpus dans le cadre du processus d'apprentissage des langues. Si ce mouvement requiert des compétences techniques de la part de l'enseignant en termes de savoir-faire en matière d'interrogation de corpus, les étudiants ne sont pas contraints d'acquérir une nouvelle compétence technique pour participer au cours. Les activités liées à l'exploration de corpus se déroulent principalement *via* des interfaces Web faciles à utiliser. Deux exemples de ce type sont la plateforme English-Corpora[1] ou la plateforme ScienQuest[2] qui requièrent un apprentissage minimal du langage d'interrogation de corpus afin d'identifier des occurrences linguistiques spécifiques. Quelques exemples d'interrogation possibles avec ces deux outils incluent, sans s'y limiter, les collocations, les colligations, la conjugaison, l'utilisation discursive d'éléments lexicaux spécifiques (patrons lexicaux, n-grammes) et d'autres phénomènes lexico-grammaticaux.

Un autre outil fréquemment utilisé à ce stade est le logiciel AntConc (Antony, 2020) : ce dernier peut être téléchargé et utilisé pour interroger un corpus « stocké » sur son dispositif personnel (ordinateur, tablette). De nombreuses études ont mis en évidence les avantages de cet outil dans diverses classes de langue, avec des niveaux de compétence linguistique variables. En outre, un nombre croissant d'études indiquent un recours accru à AntConc ou à des « outils d'interrogation de corpus hors ligne » similaires. Par ailleurs, l'évolution rapide des outils et des méthodes d'interrogation de corpus a vu l'arrivée de sites Web qui stockent de multiples corpus et proposent des outils sur mesure pour l'interrogation directement en ligne de données préexistantes ou pouvant être ajoutées par l'utilisateur. Pour des exemples concrets et une discussion détaillée sur des outils Web de ce type, comme SketchEngine, IMS Workbench, et CPQ Web voir respectivement Kilgarriff *et al.* (2014), Kübler (2014) et Kübler & Hamilton (2018). Tout au long de ce deuxième mouvement, l'enseignant reste responsable de la transmission du savoir et choisit les corpus, les outils et les logiciels en fonction des objectifs linguistiques spécifiques du cours. L'étudiant devient un participant actif en interagissant et en s'engageant avec la deuxième langue davantage que dans un cadre traditionnel de formation ainsi qu'en effectuant des requêtes spécifiques sur corpus soigneusement sélectionnés par l'enseignant. Deux retombées immédiates

[1] English-Corpora (anciennement BYU-COCA) est un outil en ligne, hébergeant quelques-uns des corpus les plus grands – en termes de taille – et les plus exploités virtuellement dans le monde anglophone : à savoir, entre autres, l'iWeb, le COCA, COHA, le BNC. Il permet aux utilisateurs non spécialistes du TALN d'interroger l'ensemble des corpus hébergés.

[2] ScienQuest est un outil en ligne, permettant aux utilisateurs non spécialistes en TALN d'interroger plusieurs corpus annotés en français et en anglais.

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

du mouvement n° 2 sont la participation active accrue et l'observation des faits de langue authentiques en contexte réel, dont il est démontré qu'elles ont toutes deux un impact positif sur la réception et la production langagière.

Mouvement n° 3. Cette étape adopte une perspective similaire en ce qui concerne l'exploration et l'interrogation des corpus linguistiques *via* des outils en ligne ou hors ligne. Une différence notable concerne la constitution des corpus. En effet, une tendance récurrente dans la littérature récente va dans le sens d'une recrudescence au niveau de l'utilisation de corpus spécialisés dans l'enseignement des langues, au lieu de corpus généraux plébiscités notamment dans le mouvement n° 1. En plus de l'évolution progressive vers des corpus spécialisés, c'est-à-dire restreints à certaines situations de communication homogènes, un changement de paradigme didactique observé tend vers l'incitation à utiliser de corpus spécialisés créés par les étudiants eux-mêmes (Charles, 2012, 2014 ; Kübler & Hamilton, 2018 ; Loock, 2019). Ces changements ont de nombreuses implications pour les pratiques d'enseignement et seront discutés plus en détail dans la section 3.2. Notons toutefois que le processus de compilation n'est plus la « chasse gardée » des enseignants et que les bénéfices de ce mouvement sont de plus en plus documentés. Parmi les avantages notables, citons l'augmentation de l'attention, de la motivation et subséquemment de la participation des étudiants, ainsi qu'une meilleure adaptabilité du contenu des cours aux niveaux de compétence réelles et un alignement des formations en langue avec les débouchés professionnels immédiats ou futurs des étudiants.

En bref, il convient de noter que le mouvement n° 1 fait référence à une phase pendant laquelle l'enseignant compile ou utilise des données préalablement compilées, les analyse et les présente à l'étudiant sous la forme de contenu de cours d'une part, ou exploite des matériels pédagogiques réalisées par d'autres mais qui sont issus de corpus d'autre part. L'étudiant a un rôle passif tout au long de cette période. Dans le mouvement n° 2, l'enseignant incorpore des requêtes d'interrogations de corpus dans les activités de classe. Les étudiants participent à des interrogations de corpus soigneusement présélectionnées par l'enseignant. Dans le mouvement n° 3, en plus de la contribution de l'enseignant, les étudiants compilent et analysent leurs propres corpus : à travers ces tâches, ils deviennent acteurs actifs de leur apprentissage. De plus, les résultats de leurs analyses peuvent être utilisés à des fins d'évaluation ou peuvent être partagés avec la classe.

Rappelons que dans les trois mouvements, la taille du corpus diffère au fil du temps – du général au spécialisé et du plus grand au plus petit – et que les analyses de corpus dépendent des compétences de l'enseignant. Enfin, les mouvements ne sont ni séquentiels ni mutuellement exclusifs. Ils n'impliquent pas, de ce fait, une progression chronologique d'un mouvement à l'autre ni que les enseignants ne peuvent pas s'inscrire dans un ou plusieurs mouvements simultanément.

2 Susciter l'intérêt des étudiants

Au-delà des objectifs pédagogiques purement linguistiques tant sur la plan langagier, acquisitionnel, interculturel que pragmatique se trouvent des objectifs secondaires – qui ont trait à la gestion, à l'organisation et à la mise en place concrète du cours

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

– et qui ont pour but de favoriser le processus d'enseignement-apprentissage. Ainsi relevons-nous, parmi ces objectifs, les besoins liés à la motivation ou le fait de susciter l'intérêt des étudiants aussi bien au niveau de la participation active en réception et en production que de la sensibilisation. Il est donc important de garder à l'esprit que certains cours de langue visent à développer des compétences linguistiques spécifiques qui peuvent être directement observées, vérifiées et évaluées par le biais d'exercices ou de tâches de production, d'une part. D'autres visent à sensibiliser aux compétences générales qui ne sont souvent pas facilement observables dans les tâches de production au cours de la formation mais qui constituent la base des situations de communication pouvant être rencontrées ultérieurement dans un environnement plus professionnel.

Examinons les deux exemples suivants où la sensibilisation est essentielle. Selon Siepmann (2006), les écrits universitaires diffèrent d'une langue et d'une culture à l'autre, non seulement en termes d'éléments lexico-grammaticaux mais aussi de caractéristiques discursives de niveau supérieur, parmi lesquelles nous pouvons souligner le flux ou ordre informationnel et les styles rhétoriques. Ces différences ne sont pas nécessairement des « éléments enseignables » dans une classe traditionnelle axée principalement sur le développement de la compétence linguistique (Van der Yeught, 2016). Un deuxième exemple serait l'enseignement d'une langue de spécialité aux étudiants qui se destinent à des métiers de rédacteurs techniques multilingues ou de traducteurs spécialisés. Malgré une ouverture dans laquelle les étudiants peuvent s'exercer à identifier les traits discursifs et linguistiques spécifiques d'une communauté discursive particulière, il faut garder à l'esprit que les activités en classe ne permettent que de travailler sur ou de reproduire des bribes de discours authentique.

Sans aborder l'immense apport que peut avoir un enseignement basé sur corpus dans ces deux cas précis (voir Gledhill & Kubler [2016] pour une discussion détaillée), il nous parait essentiel de préciser que la sensibilisation passe par l'explicitation, en amont, du bien-fondé d'une formation. En effet, dès lors qu'une approche non traditionnelle est adoptée – en particulier lorsque cette approche semble nouvelle ou abstraite, que le cadre de la formation relève d'une transmission passive et que les tâches ne sont pas orientées vers la reproduction et l'évaluation des observables linguistiques –, il est important d'expliquer la méthodologie aux étudiants. Ces derniers doivent clairement comprendre les objectifs et les avantages de la formation afin de pouvoir capitaliser sur leur propre motivation extrinsèque « qui provient de stimuli extérieurs à l'individu » et intrinsèque « qui vient de l'individu lui-même » (Knoerr, 2005). Sans cette capitalisation, l'intérêt en berne des étudiants peut freiner le processus d'apprentissage s'ils ne voient pas le bien-fondé du changement de paradigme. Autrement dit, toute intégration de corpus numériques dans les formations en langue se doit de faire émerger ou du moins stimuler les deux types de motivation soulignée ci-dessus.

Examinons maintenant le cadre d'une formation intégrant des corpus numériques et qui constitue la pierre angulaire de notre réflexion.

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

2.1 Présentation d'un cadre de formation en langue de spécialité

Dans le cadre d'un enseignement de socle commun en première année de master aussi bien en Langues Étrangères Appliquées (LEA) qu'en Traduction & Interprétation, il est question d'initier les étudiants à la linguistique de corpus et aux langues de spécialité. Cette initiation, qui s'inscrit pleinement dans une optique d'apprentissage des langues médiatisé par les technologies (ALMT), a pour objectif de présenter aux étudiants une vue générale de la discipline en mettant plus particulièrement l'accent sur les applications directes dans leurs débouchés professionnels futurs ; à savoir, entre autres, l'industrie de la langue, la communication technique multilingue, la recherche linguistique. La formation, qui se veut particulièrement généraliste en raison du parcours hétérogène des étudiants, retrace les grandes étapes de la linguistique de corpus en tant que discipline et en tant que méthode de recherche d'une part et explore quelques outils d'interrogation et d'annotation de corpus de façon à permettre aux étudiants de se rendre compte de la diversité d'application de la discipline d'autre part. Bien entendu, un tel cours ne saurait rester purement théorique et trouve dans le domaine des langues de spécialité un terrain d'application propice au vu des débouchés des étudiants.

Il convient de noter que ce cours se complète par des modules en linguistique de corpus appliqués spécifiquement à plusieurs champs d'étude : à titre d'exemple, à la terminologie et à la traduction.

Parmi les objectifs d'apprentissage et les principes linguistiques sous-jacents se trouve la prise en compte du poids qu'occupent le genre et les discours spécialisés. Ces deux concepts peuvent être mis en parallèle avec la distinction faite par Michael Halliday entre *genre* et *registre* dans la linguistique systémique fonctionnelle. Dans ce cadre théorique, le genre est considéré comme l'objectif d'un texte, identifiable dans l'imaginaire culturel (Eggins, 2004) ou des formes spécifiques socialement reconnues (Bloor & Bloor, 2006). Si nous prenons à titre indicatif le secteur commercial, la notion de genre peut renvoyer, entre autres, à un cahier de charges, un plan d'affaires, un résumé opérationnel. Tandis que la notion de registre renvoie à une variété communicative fonctionnelle d'une langue dont les exemples pourraient inclure l'anglais des affaires, l'anglais scientifique, l'anglais juridique.

> *A register is a tendency to select certain combinations of meaning with certain frequencies and thus can be formulated as the probabilities attached to grammatical systems, provided such systems are integrated into an overall system network in a paradigmatic interpretation of the grammar* (Halliday, 1991)

En bref, les considérations discursives en termes de forme, d'objectif et de type relèvent du genre alors que celles portant tout particulièrement sur les réalisations ou instanciations lexico-grammaticales relèvent du registre.

Par ailleurs, si certains linguistes semblent écarter toute distinction formelle et regroupent les deux concepts en une analyse significative et mutuellement inclusive (voir Swales 1990 pour une discussion sur les analyses de genre), certains procèdent par

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

désambiguïsation. Cette désambiguïsation a donné lieu à différents types d'analyse : l'analyse de genre et l'analyse de registre, parmi les courants linguistiques adoptant – de manière partielle ou intégrale – une distinction fondée sur la tradition en linguistique systémique fonctionnelle. La désambiguïsation est également centrale dans notre cours, car il incombe aux étudiants de vérifier leurs connaissances conceptuelles en obtenant des observables linguistiques ayant trait à la fois au genre et au registre/langue de spécialité.

Enfin, il convient également de noter que le format du cours reflète les besoins pédagogiques sous-jacents en termes de transmission des connaissances et d'application pratique. À cette fin, celui-ci est divisé en trois étapes : une transmission passive, axée sur la théorie ; une étape active subdivisée en formation pratique axée sur la prise en main d'un outil d'interrogation de corpus d'une part, et en compilation et exploration de corpus d'autre part.

Cette dernière étape est expliquée dans la section 2.2.

2.2 Création et analyse d'un corpus comparable spécialisé

Dans le cadre de l'évaluation du cours de master, les étudiants doivent créer un corpus comparable spécialisé, c'est-à-dire un corpus composé de données provenant de deux langues : en l'occurrence le français et l'anglais et relevant d'une langue de spécialité. Ils choisissent entre l'approche ascendante (*corpus-based*) où l'objectif est d'utiliser un corpus pour « confirmer ou infirmer une hypothèse » ou l'approche descendante (*corpus-driven*) adoptant une approche « préférant explorer sans *a priori* les données » (Teubert, 2009). Rappelons à ce titre la définition consensuelle de ces deux termes avancée par McEnery & Hardie.

> *Corpus-based studies typically use corpus data in order to explore a theory or hypothesis, aiming to validate it, refute it or refine it. The definition of corpus linguistics as a* method *underpins this approach to the use of corpus data in linguistics. Corpus-driven linguistics rejects the characterisation of corpus linguistics as a method and claims instead that the corpus* itself *should be the sole source of our hypotheses about language. It is thus claimed that the corpus itself embodies a theory of language.* (McEnery & Hardie, 2011 : 6)

Les deux sous-corpus sont ensuite analysés pour au moins deux traits caractéristiques souvent choisis au niveau lexico-grammatical ou discursif. Un exemple du premier niveau peut comprendre la terminologie spécialisée tandis qu'au deuxième niveau il peut être question de positionnements énonciatifs à travers l'étude de l'emploi des pronoms personnels. Précisons, par ailleurs, que les consignes générales pour la compilation du corpus sont adoptées principalement de De Bonville (2006), Gries (2009), Hunston (2006), McEnery & Hardie (2011) et Sinclair (2005). Celles-ci sont présentées de manière succincte sous forme de critères ci-dessous.

1) Critères de sélection, considérés comme indispensables pour garantir l'authenticité du corpus

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

Il est important que l'ensemble du corpus soit choisi en lien direct avec la fonction communicative de la variété linguistique que l'on souhaite étudier. La variété doit être choisie au vu de son ancrage dans une situation de production réelle (critère externe). Ceci s'oppose au fait de choisir un corpus selon les spécificités linguistiques qu'un ou plusieurs texte(s) pourrai(en)t contenir (critère interne). Parmi les critères externes, nous pouvons citer le registre, le mode (oral, écrit), le genre textuel, la variété linguistique (dialecte, aires géographiques ou culturelles).

2) Critères d'échantillonnage, avec un résumé succinct tiré de Jean de Bonville

> « L'échantillon correspond à la fraction ou au sous-ensemble de messages tirés de la population étudiée et que le chercheur entend soumettre à l'analyse proprement dite. Comme la population d'où il provient, l'échantillon doit présenter les qualités de pertinence et d'homogénéité, même si, par définition, il ne peut être exhaustif. L'échantillon doit être représentatif de la population, c'est-à-dire que tous les caractères pertinents présents dans la population doivent se retrouver dans l'échantillon, en principe dans des proportions semblables. » (2006 : 104)

3) Critères de représentativité : essentiels notamment pour les corpus spécialisés

Notons d'emblée que la notion de représentativité n'est pas consensuelle en linguistique de corpus. En effet, Sinclair (2005) y voit un besoin de refléter la réalité d'usage effectif tandis que Biber (1993) prône ce qui s'apparente à une représentativité proportionnelle à l'ensemble des variations connues. Cependant, tout corpus se doit d'être représentatif, de la langue générale ou d'une langue de spécialité. Selon Leech (2007) la non-représentativité signifie que les résultats obtenus d'un corpus donné ne pourront en aucun cas être comparés ou généralisés à autre chose que le corpus duquel ils sont issus. Le corpus ne serait exploité que par ses concepteurs et est voué, par conséquent, à devenir ce que l'on appelle un « corpus fantôme » (Baude & Abouda, 2006).

3 Repenser la transmission et la co-construction du savoir

Dans les sections précédentes, nous avons souligné que le corpus occupe plusieurs fonctions. Couplé avec d'autres ressources, il peut être envisagé par l'enseignant, entre autres, en tant que sources de connaissances à extraire ou identifier et ensuite à transmettre aux étudiants aussi bien que « matériel didactisable » avec lequel l'étudiant doit interagir, notamment dans le cas des requêtes d'interrogations de corpus. Il peut également servir de méthodes d'évaluation de l'étudiant, tant au niveau des activités d'interrogation qu'à travers sa compilation, si la situation le permet, comme dans le cadre de notre formation en master. L'ensemble de ces rôles ont été mis en avant dans les sections précédentes, à l'exception de celui portant sur l'évaluation. Dans la section 3.1, nous nous intéressons à ce rôle et nous proposons un retour critique d'expérience sur les considérations et questionnements qui subsistent après plusieurs années à enseigner dans cette formation.

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

La section 3.2 s'intéresse à l'impact de changements observés dans notre formation à l'aune des rôles des enseignants, à travers notamment leur fonction de passeurs de connaissances et l'inévitable redynamisation ou changement de paradigme qui semble se profiler.

3.1 Prise de conscience à travers l'évaluation des étudiants

Précisons que les approches adoptées par les enseignants influencent considérablement les stratégies et méthodes d'évaluation mises en place dans les formations en langue. À titre d'illustration, si le numérique se limite à un rôle de support de cours, il apparait dérisoire d'évaluer l'étudiant sur son rôle de support. Si par contre il est envisagé en tant qu'outil de production, d'application ou d'entrainement langagier, le produit final de l'étudiant aussi bien que sa maîtrise de l'outil – si ce dernier a fait l'objet d'un enseignement explicite – peuvent faire l'objet d'une évaluation. Dans le cas de notre cours, qui s'inscrit dans le mouvement n° 3 présenté dans la section 1.2, nous sommes confrontés à la question de savoir sur quelles modalités ou sur quels éléments doit porter l'évaluation. En effet, s'il est question d'évaluer les connaissances acquises (théories et mises en applications linguistiques), il convient de noter qu'une mauvaise maîtrise des outils d'interrogation aussi bien qu'un corpus qui n'aurait pas été compilé correctement peuvent influer sur les résultats obtenus lors de la phase d'interrogation et subséquemment sur les analyses qui en découlent. Ainsi, afin de ne pas doublement pénaliser ceux qui n'auraient pas maitrisé convenablement toutes les étapes du processus d'apprentissage, il nous parait judicieux de scinder l'évaluation en deux : pour les compétences techniques (compilation et interrogation de corpus) d'une part et pour les compétences linguistiques (analyses et interprétation des résultats) d'autre part.

Bien que cela puisse sembler complexifier inutilement le processus d'apprentissage et d'évaluation, il est important de garder à l'esprit que l'enseignant décide et adapte le degré d'intégration et d'exploration des corpus en fonction de son niveau d'expertise ou de confort dans les outils d'une part et des compétences linguistiques de ses étudiants d'autre part.

Au-delà de ces premières considérations pédagogiques, rappelons également que les objectifs de cette formation résident principalement dans l'acquisition de nouvelles connaissances linguistiques et une sensibilisation accrue relatives à l'importance et l'interdépendance du genre et des discours spécialisés. Les connaissances linguistiques et techniques sont vérifiables et évaluées à travers les phases de compilation et d'interrogation du corpus. La sensibilisation n'est pas évaluée, à proprement parler, mais observable par le biais des mises en garde ou conseils que les étudiants laissent à des futurs rédacteurs techniques ou traducteurs spécialisés. En effet, cela fait partie des éléments à inclure dans le dossier final que les étudiants doivent soumettre.

Précisons, à titre d'information, que les exemples ci-dessous relèvent uniquement des dossiers reçus de la promotion 2020-2021, dans laquelle il y avait 73 étudiants. Les dossiers sont réalisés principalement par groupe de 2 ou 3 étudiants, comportent chacun une quinzaine de pages et peuvent être rédigés dans la langue de leur choix.

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

Prenons le cas du tableau 2. Rappelons tout d'abord que les étudiants compilent un corpus comparable, en anglais et en français, dans une langue de spécialité de leur choix. Ils effectuent un ensemble d'analyses du niveau lexico-grammatical au niveau discursif, en s'intéressant bien entendu au niveau textuel. Ils décrivent des points de convergence et de divergence, mettent en avant des éléments saillants en termes d'occurrences ou faits de langues distinctifs entre les deux langues. La reproduction de ces éléments constitutifs de la langue de spécialité constitue un facteur essentiel dans l'intégration des communautés discursives au sens de Swales (2016). Et sans tomber dans la prescription linguistique extrême, nous soutenons que toute utilisation différente en termes de présence ou absence effective d'un trait caractéristique peut impacter considérablement le genre en tant que forme socialement reconnue (Bloor & Bloor, 2006). C'est ainsi que dans le tableau 2, nous pouvons relever des préconisations différentes, portant sur un même phénomène linguistique en langue anglaise mais adaptées à trois genres différents. En effet, le dossier C03 souligne un contexte de communication dans lequel l'emploi des pronoms personnel est à privilégier, C07 indique que cet emploi doit être modéré tandis que C02 précise qu'il est tout simplement à bannir. Ces préconisations ne peuvent être formulées qu'après une interrogation minutieuse d'un corpus comparable. Précisons donc que les étudiants en fournissent pour tous les faits de langues relevés et pour tous les niveaux d'analyses effectués.

Tab. 2 : Différences dans l'emploi des pronoms

Genre/type de texte	source	Exemple
Horoscope	C03	We would also recommend addressing the lector as if the writer really knew him or her, to use personal pronouns such as you, tu, vous when writing.
Articles scientifiques en Musicothérapie	C07	Lorsque l'on traduit en anglais, nous pensons qu'il faut également être vigilant sur l'emploi des pronoms, beaucoup plus fréquent en français, le risque étant de produire un texte qui sonne trop 'personnel' alors que ça n'est pas ce qui est attendu en anglais.
Décisions de la CEDH et CJUE	C02	The use of a completely impersonal tonality, even in situations that would normally call for more empathetic language […]

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

Tab. 3 : Conseils généraux

Genre/type de texte	Source	Exemples
Recettes de pâtisserie	C04	French is said to use more nouns while English is said to use more verbs. For instance, the collocation "porter à ébullition" needs to be translated by a single verb : "until X/ it/ the water boils." These points would probably not lead to an incomprehensible recipe, […] however such codes are what make the texts authentic and belong to the genre which is what Language for Specific Purposes is about.
Critiques cinématographiques	C11	To begin with, that type of text should not be translated with an automated translation software, or at least not without a very careful post-editing process as it can be full of metaphors or expressions and references that that type of technology might not get.
Articles scientifiques	16	Another distinctive feature of scientific writings is the resort to hedgers and boosters. When conducting the analysis, we found out that hedging adverbs are used more frequently in English texts than in French ones. These results are in line with the idea that French authors, although they do not use the first-person pronoun, tend to be more authoritative in their writings.
Articles de vulgarisation scientifique	C17	Tout d'abord, peu importe le degré de spécialité du texte de vulgarisation, il est essentiel de se procurer ou de constituer un corpus de référence thématique et pertinent dans la langue cible et source. Ceci vous assistera dans vos démarches de traduction, notamment pour cerner les difficultés terminologiques et collocationnelles du thème, afin de restituer le message de manière claire et fidèle dans la langue cible.

Soulignons également que les étudiants fournissent des conseils d'ordre général sur des aspects linguistiques, des techniques de rédaction à mettre en place ou des pièges à anticiper ou à éviter par toute personne souhaitant s'initier aux langues de spécialité. Examinons brièvement quelques exemples du tableau 3. Le dossier C04 relève une différence de structure et de catégorie grammaticales dans la réalisation d'un même phénomène dans les deux langues et préconise, à juste titre, le recours au procédé bien connu de la transposition en traduction. C11 met en avant la fréquence des expressions figées et des figures de style et met en garde contre la tentation de faire de la traduction littérale ou automatique, ce qui pourrait avoir comme résultat un texte totalement abscons. Cette mise en garde n'est pas anodine. En effet, quelqu'un souhaitant s'initier au type et au style de discours sur lesquels C11 travaille doit garder à l'esprit que l'absence de ce trait distinctif et caractéristique aura une incidence directe sur la bonne réception de son texte – en raison de son incongruité. Le dossier C16 souligne un fait de langue qui a été amplement documenté dans la littérature scientifique (Hyland, 1996 ; Hartwell *et al.*, 2017) tandis que C17 rappelle l'importance de constituer un corpus de référence, c'est-à-dire un corpus représentatif qui constitue un point de repère en termes de type de productions langagières que l'on étudie ou que l'on cherche à reproduire.

L'ensemble des éléments des tableaux 2 et 3 indiquent une certaine prise de conscience de la perméabilité ou de l'indissociabilité entre genre, discours spécialisés et ses corollaires au niveau lexico-grammatical, discursif et textuel.

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

Signalons maintenant quelques points clés à retenir de cette formation.

- Tout d'abord, les résultats bruts obtenus par les étudiants lors de la phase d'interrogation de corpus sont tout aussi valables que ceux que nous avons obtenus, notamment quand les compétences techniques sont bien acquises. À ce titre, il faut veiller à ne pas les « corriger » et les écarter comme des simples productions d'étudiants.

- Les choix de corpus créés par les étudiants fournissent souvent des champs de discours spécialisés peu exploités dans la littérature scientifique, notamment en termes de genre et langue de spécialité. Il faut donc accepter de ne pas maîtriser, en amont, le genre ou le discours spécialisé de tous les dossiers soumis. Cela ne doit donc constituer un frein en soi. Par contre, il est important que la méthode d'évaluation soit robuste. Pour la phase de compilation et d'interrogation, il est important d'avoir des critères clairement énoncés d'une part et une analyse des résultats qui privilégie la pertinence des faits de langue relevés ainsi que l'application des connaissances linguistiques dans l'interprétation des résultats d'autre part.

- Le dernier point concerne la progression du cours. Les trois étapes (pour rappel, une transmission traditionnelle, la prise en main d'un outil d'interrogation de corpus et la phase exploratoire du corpus en autonomie, voir la section 2.1.) doivent être repensées de façon à éviter ou écourter la passivité observée dans la phase transmissive. Une solution sera de commencer avec les étapes actives et transformer l'étape transmissive en une mise en commun collective.

3.2 Évolution et adaptation d'un paradigme didactique

Bien que cela puisse être souligné dans toute approche didactique, il convient de noter que les approches qui découlent des trois mouvements que nous avons mis en avant affectent grandement la dynamique de ce que nous enseignons, la manière dont nous enseignons et celle dont nous évaluons nos étudiants. La section 3.1 a exploré les impacts de l'évaluation. Abordons maintenant l'impact du corpus sur le contenu de l'enseignement et sur les pratiques pédagogiques. En effet, comme nous l'avons vu dans la section 1.2 le recours aux corpus en didactique des langues peut être globalement regroupé en trois mouvements distincts : le premier où les données du corpus sont extraites et didactisées ; le deuxième où le corpus devient une plateforme dans laquelle différentes activités peuvent être réalisées – principalement des requêtes d'exploration ; le troisième où les étudiants participent à la construction et à l'analyse de leur propre corpus. Cependant, même si ces mouvements ne sont pas mutuellement exclusifs ni nécessairement séquentiels dans une classe donnée, ils affectent la manière dont le contenu du cours est conceptualisé et transmis.

Dans sa conception la plus large, le principe fondateur de la pédagogique transmissive est encore présent dans les pratiques pédagogiques modernes, malgré le fait que le concept soit désuet ou renvoie à des préconisations d'un autre temps, comme le laisse entendre Paveau (2006). En effet, malgré les nombreuses approches qui ont modelé

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

la didactique des langues dans l'enseignement supérieur français – en passant par les méthodes audio-visuelles jusqu'à l'approche actionnelle –, le modèle majoritaire demeure résolument transmissif, linéaire et asymétrique. L'enseignant détient le savoir et le transmet à ses étudiants, en adoptant le format des cours aux courants auxquels il adhère. Ce modèle transmissif a fait ses preuves et n'est pas à écarter, comme l'a souligné Guillemette (2020). Cependant, dans les trois mouvements mis en avant – et notamment les deux derniers –, nous relevons un vaste champ de possibilités non exploitables en raison du fait que l'enseignant s'inscrit et s'enferme dans le modèle transmissif.

Prenons les quatre phases de l'utilisation de corpus (voir tableau 1). Dans une conception linéaire, les phases semblent séquentielles et relever du domaine de compétences de l'enseignant. Ce dernier compile et interroge son corpus, selon le mouvement ; mais le propre de sa fonction est de didactiser tout savoir qu'il souhaite transmettre et il lui revient également de décider des modalités de sa transmission. Cependant le mouvement n° 3 nous a montré que l'étudiant peut accéder aux phases de compilation et d'interrogation au même titre que l'enseignant. Désormais, chacun aura accès à la même information primaire et par conséquent la phase de didactisation se transforme en interprétation. En effet, si nous acceptons que l'exploitation du corpus passe par l'acquisition d'une compétence technique nouvelle de la part de l'enseignant, il convient d'accepter que les étudiants aient également à acquérir cette compétence – notamment dans le mouvement n° 3 et dans un moins mesure dans le mouvement n° 2.

Comme nous l'avons suggéré en abordant l'évaluation de nos étudiants, leur choix de faits de langue, genre et registre dépassent ce qui est traditionnellement présenté en tant que contenu théorique dans une seule formation. De ce fait, quand celles-ci sont correctement réalisées, leurs analyses constituent une nouvelle source de connaissances non négligeables qui pourront être partagées avec les autres étudiants. Ainsi, les étudiants participent à trois des quatre phases habituellement réservées aux enseignants : à savoir, compilation, interrogation, transmission. Il nous parait par conséquent déraisonnable de ne pas « réinjecter » ces nouvelles connaissances dans le cours, étant donné que le principe fondateur de notre formation est axé sur l'apprentissage sur corpus. Cette prise de conscience nous amène à nous interroger sur le rôle de l'enseignant face aux nouvelles technologiques dans les formations en langue.

En effet, si nous acceptons que certaines nouvelles technologies nécessitent l'acquisition d'une nouvelle compétence de la part de l'enseignant et de l'étudiant, cette compétence peut être de niveau symétrique ou asymétrique en faveur de l'un ou de l'autre. Par exemple, l'enseignant peut demander à ses étudiants de créer et de tenir un blog, s'il souhaite favoriser leur production orale et écrite en langue étrangère. S'il souhaite intégrer un aspect davantage ludique, il peut opter pour la gamification. Dans les deux cas, il est possible que certains de ses étudiants aient des compétences techniques dans ces deux domaines qui dépassent les siennes. Dans ce cas de figure, il n'y a pas de raison que cela constitue un problème. Cette asymétrie est possible notamment quand il est question de nouvelles technologiques qui ont été « détournées » de leur public principal afin d'être mises au profit de la didactique. Cela étant, comme

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

dans un cours de langue hétérogène – avec des étudiants de niveaux avancés et faibles – il est de la responsabilité pédagogique de l'enseignant de faire en sorte que tout le monde avance, quitte à s'appuyer d'une manière ou d'une autre sur ceux d'un niveau avancé pour faciliter certaines types d'activités. L'enseignant issu du mouvement n° 3 doit donc être capable de reconnaitre la valeur ajoutée de certains résultats des phases de compilation et d'interrogation de ses étudiants en tant que connaissances de source primaire et, si besoin, les incorporer dans ses modalités de transmission au profit de toute la classe.

Autrement dit, étant donné la nature des enseignements basés sur corpus notamment dans le mouvement n° 3, nous soutenons qu'il y aura de temps en temps de nouvelles découvertes non-prévues par l'enseignant mais tout aussi valables. Il convient dès lors de découvrir le moyen de dépasser le modèle singulièrement transmissif, de sorte que les « découvertes » des étudiants puissent être intégrées dans la mesure du possible dans l'apprentissage collectif.

4 En guise de conclusion

Comme nous l'avons soutenu, le numérique au sens large est devenu une partie intégrante des formations en langue en milieu universitaire. Cependant, les corpus en tant qu'objets numériques ne bénéficient pas du même engouement et ne sont pas encore pleinement intégrés dans les pratiques didactiques contemporaines et ce, malgré une progression conséquente depuis les années 1980. En dépit de ce constat, il est important de souligner une certaine démocratisation dans le recours aux corpus dans les trois mouvements didactiques que nous avons mis en avant.

Par ailleurs, nous avons montré que l'utilisation de corpus influe sur nos pratiques pédagogiques, en nous intéressant notamment à la notion de transmission du savoir et à quelques changements engendrés quand l'enseignement-apprentissage passe par une participation et implication accrues des étudiants. Nous nous sommes également interrogé sur l'impact de l'utilisation de corpus – à travers des phases de compilation et d'interrogation – sur les modalités d'évaluation. Dès lors, il nous parait important de repenser systématiquement l'identification des compétences et connaissances à évaluer, en fonction aussi bien des niveaux de langue des étudiants que des degrés d'intégration des corpus dans une formation donnée.

Enfin, nous avons souligné le rôle que joue l'exploration d'un corpus comparable dans la prise en compte des spécificités liées aux langues de spécialité. Par conséquent, nous soutenons qu'il faut désormais dépasser le modèle du savoir transmissif notamment en milieu universitaire afin de favoriser davantage et ancrer la démarche de co-construction du savoir linguistique dans nos pratiques pédagogiques : l'enjeu sera que les étudiants comprennent la validité des activités entreprises dans la formation et s'engagent et participent activement dans l'acquisition des compétences langagières.

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

Références bibliographiques

Anthony, L. (2020). AntConc (Version 3.5.9) [Logiciel]. Tokyo, Japan : Waseda University. URL : https://www.laurenceanthony.net/software

Baude, O. & Abouda, L. (2006) Constituer et exploiter un grand corpus oral : choix et enjeux théoriques. Le cas des ESLO. Actes du colloque Corpus en lettres et sciences sociales – des documents numériques à l'interprétation, 10-14 juillet 2006.

Biber, D (1993). "Representativeness in Corpus Design." In : Literary and Linguistic Computing 8.4, 241–257.

Bloor, T. & Bloor, M. (2006) The Functional Analysis of English : A Hallidayan Approach. Londres : Hodder Arnold.

Boulton, A. & Tyne, H. (2014). Corpus-based study of language and teacher education. In M. Bigelow & J. Ennser-Kananen (éds), The Routledge Handbook of Educational Linguistics. New York : Routledge, 301-312.

de Bonville, J. (2006). L'analyse de contenu des médias : De la problématique au traitement statistique. Louvain-la-Neuve : De Boeck Supérieur. https://doi.org/10.3917/dbu.bonvi.2006.01

Charles, M (2012). "Proper Vocabulary and Juicy Collocations" : EAP Students Evaluate Do-It-Yourself Corpus-Building. English for Specific Purposes, 31, 93-102.

Charles, M. (2014). Getting the Corpus Habit : EAP Students' Long-Term Use of Personal Corpora. English for Specific Purposes, 35, 30-40.

Ciekanski, M. (2014) Les corpus : de nouvelles perspectives pour l'apprentissage des langues en autonomie ?, Recherches en didactique des langues et des cultures 11 (1). DOI : https://doi.org/10.4000/rdlc.1710

Debaisieux, J.-M. (2009). "Des documents authentiques oraux aux corpus : un défi pour la didactique du FLE". Mélanges CRAPEL, 31, 35-56.

Dejean-Thircuir, C., Nissen, E. (2013). Évolutions technologiques, évolutions didactiques, Le Français dans le monde, Recherches et applications, « Mutations technologiques, nouvelles pratiques sociales et didactique des langues », n° 54, 28-40.

Demaizière, F. (2007) Didactique des langues et TIC : les aides à l'apprentissage, Alsic 10 (1). DOI : https://doi.org/10.4000/alsic.220

Demaizière, F. & Grosbois, M. (2014) Numérique et enseignement-apprentissage des langues en Lansad – Quand, comment, pourquoi ?, Alsic vol. 17 | DOI : https://doi.org/10.4000/alsic.2691

Eggins, S. (2004). An Introduction to Systemic Functional Linguistics. Londres : Continuum.

Frankenberg-Garcia, A. (2014) How language learners can benefit from corpora, or not, Recherches en didactique des langues et des cultures, 11-1. DOI : https://doi.org/10.4000/rdlc.1702

Gledhill, C. & Kübler, N. (2016) What can linguistic approaches bring to English for Specific Purposes ?, Asp, 69. DOI : https://doi.org/10.4000/asp. 4804

Gries, S. (2009) What is Corpus Linguistics ? Language and Linguistics Compass 3/5 (2009) : 1225–1241.

Grosbois, M. (2015) S'approprier le numérique : réussite ou échec ? Pour qui, pourquoi, comment ?, Recherche et pratiques pédagogiques en langues de spécialité vol. XXXIV N° 2. DOI : https://doi.org/10.4000/apliut.5168

Guichon, N. (2012) L'apprentissage des langues médiatisé par les technologies (ALMT) – Étude d'un domaine de recherche émergent à travers les publications de la revue Alsic », Alsic 15 (3). DOI : https://doi.org/10.4000/alsic.2539

Guillemette, F. (2020). Passer du modèle transmissif à un modèle de l'apprentissage guidé. Enjeux et société : Approches transdisciplinaires. 7 (2) 42–73.

Halliday, M. A. K. (1991) Corpus studies and probabilistic grammar. In K. Aijmer, & B. Altenberg (éds.), English corpus linguistics. Studies in honour of Jan Svartvik. Londres : Longman.

Hartwell, L., Esperança-Rodier, E., & Tutin, A. (2017). I think we need… : Verbal expressions of opinion in conference presentations in English and in French. CHIMERA : Revista De Corpus De Lenguas Romances Y Estudios Lingüísticos, 4(1), 35–60.

Higgins, J. & T. Johns. (1984), Computers in Language Learning, London : Collins.

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109

Howard, S. K. & Mozejko, A. (2015). Considering the history of digital technologies in education. In M. Henderson & G. Romero (éds.), Teaching and Digital Technologies : Big Issues and Critical Questions (p. 157-168). Port Melbourne : Cambridge University Press.

Hunston, S. (2006) Corpus Linguistics. In Brown, K (éd.) The Encyclopedia of Language and Linguistics. Oxford : Elsevier.

Hyland, K. (1996). "Writing without conviction ? Hedging in science research articles". Applied Linguistics, vol. 17, n° 4. p. 433-454.

Johns, T. (1991). "Should you be persuaded : two samples of data-driven learning materials". In Johns, T. & King, P. (dir.), Classroom concordancing. English Language Research Journal, vol. 4. p. 1-16.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine : Ten Years On. Lexicography, 1, 7-36.

Knoerr, H. (2005) TIC et motivation en apprentissage/enseignement des langues. Une perspective canadienne. Cahiers de l'APLIUT, vol. XXIV N° 2 | DOI : https://doi.org/10.4000/apliut.2889

Kübler, N. (2014) « Mettre en œuvre la linguistique de corpus à l'université », Recherches en didactique des langues et des cultures 11-1 | 2014, DOI : https://doi.org/10.4000/rdlc.1685

Kübler, N & Hamilton, C.E. (2018) L'exploitation des corpus numériques dans les formations d'anglais scientifique en ligne : une étude de cas, Lidil, vol. 58

Leech, G. (2007). New resources, or just better old ones ? The Holy Grail of representativeness. In M Hundt, N Nesselhauf & C Biewer (eds), Corpus Linguistics and the Web. Amsterdam : Rodopi 133-149.

Loock, R. (2019) La traductologie de corpus. Presses Universitaires du Septentrion.

McEnery, T. & Hardie, A. (2011) Corpus Linguistics : Method, Theory and Practice. Cambridge : Cambridge University Press.

Paveau, M. A. (2006). Chronique « Linguistique ». Leçon de vocabulaire (1). Le français aujourd'hui, 154, 121-128. https://doi.org/10.3917/lfa.154.0121

Pavel, S et Nolet, D (2001) Précis de terminologie/Handbook of Terminology. Ottawa, Travaux publics et services gouvernementaux du Canada.

Qotb, H. (2019) Apprentissage des langues et numérique : contextualisations, interactions et immersions. Dossier d'Habilitation à diriger des recherches, Université de Lorraine.

Siepmann, D. (2006) Academic writing and culture : An overview of differences between English, French and German. Meta : journal des traducteurs/Meta : Translators' Journal, vol 51 (1) | DOI https://doi.org/10.7202/012998ar

Sinclair, J. 2005. Corpus and Text – Basic Principles. In Developing Linguistic Corpora : a Guide to Good Practice (éd.) Wynne, M. Oxford : Oxbow Books, 1-16.

Soubrié, T. (2020) Penser le numérique : orientations pédagogiques et préoccupations de recherche en didactique des langues, Alsic 23 (1)| DOI : https://doi.org/10.4000/alsic.4687

Swales, J. (1990). Genre Analysis, Cambridge University Press.

Swales, J (2016). Reflections on the Concept of Discourse Community. ASp, 69, 7-19.

Teubert, W. (2009) La linguistique de corpus : une alternative [version abrégée], Semen 27 | DOI : https://doi.org/10.4000/semen.8914

Van der Yeught, M. (2016) A proposal to establish epistemological foundations for the study of specialised languages, ASp, 69 | DOI : https://doi.org/10.4000/asp. 4788

Published in : Louise Ouvrard, Catherine Brumelot (dir.), *Numérique et didactique des langues et cultures. Nouvelles pratiques et compétences en développement*, Editions des archives contemporaines, ISBN : 9782813004109, doi : 10.17184/eac.9782813004109