

Enjeux des corpus en didactique de la grammaire pour la formation des enseignants. Présentation du numéro

Camussi-Ni, Marie-Armelle

CELLAM EA 3206

Université Rennes 2

Pour citer cet article :

Camussi-Ni M.-A. (2024). Enjeux des corpus en didactique de la grammaire pour la formation des enseignants. *Scolagram* n°11. Novembre 2024. https://scolagram.u-cergy.fr/index.php/content_page/item/1349-presentation-du-numero-enjeux-des-corpus-en-didactique-de-la-langue-pour-la-formation-des-enseignants

La linguistique outillée a pris une place très importante dans le champ de la recherche en linguistique et l'utilisation de corpus y semble devenue incontournable au risque d'ailleurs de lui accorder un crédit illimité¹. S'il importe de ne pas confondre l'outil destiné à investiguer le réel avec le réel lui-même, il est indéniable que l'accent mis sur les corpus de linguistique a permis de mettre au jour des caractéristiques de la langue peu identifiables sans cela. Il reste que les corpus ne sont pas directement lisibles pour les profanes, ils demandent à être soutenus par une description de la langue ne serait-ce que pour la confronter à l'usage et d'une réflexion sur les usages statistiques pour en mesurer la part subjective.

Un des postulats des travaux sur corpus est d'accéder à une représentation plus authentique de la langue, de pouvoir mettre à l'épreuve des représentations que le locuteur voire le spécialiste en a. Il n'est donc pas étonnant que les chercheurs en didactique de la langue s'en emparent à leur tour, soit parce qu'ils ont une conscience aigüe que la description de la langue enseignée est très peu authentique et que les représentations véhiculées par la grammaire traditionnelle enseignée à l'école qui exclut une partie de la réalité de la langue par sa normativité (Cappeau *et al.*, 2021 ; Gadet & Guérin, 2022) aboutissent à des définitions au moins en partie erronées, soit parce qu'ils souhaitent mieux appréhender les étapes de développement en milieu scolaire.

Conscients de l'enjeu scientifique, les chercheurs en didactique de la langue se sont donc saisis à leur tour des corpus pour évaluer leur potentiel dans l'enseignement – et l'appréhension des objets de recherche par les corpus de langue

¹ De nombreux chercheurs abordent cette question épistémologique (Milner, Charaudeau, Cori et David) devenue de plus en plus vive, semble-t-il et la discussion des limites de la linguistique basée sur des corpus est devenue un passage obligé dans les thèses de linguistique actuelles.

investit aussi aujourd'hui la recherche en didactique de la langue française. Cette recherche sur les corpus en didactique de la langue correspond à trois grands axes identifiables selon le type de corpus convoqué, souvent en interrelation comme en témoignent les articles de ce numéro : l'exploitation de grands corpus du français, écrit ou parlé ; l'analyse des obstacles de l'apprentissage et des étapes de développement à partir de corpus d'apprenants ou corpus scolaires ; et enfin, la création de corpus d'apprentissage ou corpus d'enseignement.

La démarche d'exploitation des grands corpus du français écrit ou parlé (comme Eslo, Frantext, le CNRTL, BDTS, BELTEXT, CERF, CFPP, CIEL, CLAPI, MPF, *L'Est Républicain* ...) est une démarche bien implantée en Français Langue Étrangère (FLE). Elle vise à l'utilisation de ces corpus par des apprenants allophones sous l'appellation de *data-driven learning* (Johns, 1986) ou d'apprentissage sur corpus (Boulton et Tyne, 2014) et elle a donné lieu à la création de corpus adaptés pour les apprenants du FLE (citons par exemple, CLAPI-FLE développé par l'équipe d'ICAR à Lyon, le corpus Fleuron (André, 2020) qui inclut l'analyse des méta-données ou encore FULS sur le corpus Scientex (Cavalla et Loiseau, 2013). Ces grands corpus du français étaient encore peu sollicités dans le cadre de la didactique en Français Langue Première (FLP), il y a une dizaine d'années, selon M.-L. Elalouf (2012), et le restent probablement. Peut-être faut-il mettre ce constat en relation avec le poids de la norme dans l'enseignement du FLP.

Inversement, la constitution de corpus d'apprenants, elle aussi plus ancienne en FLE, a aujourd'hui le vent en poupe en FLP. En didactique du FLE, les travaux sur corpus d'apprenants se sont dégagés des inventaires d'erreurs pour aller soit vers l'exploitation qualitative de corpus longitudinaux dans le cadre d'approches acquisitionnistes qui peinent encore à être didactisées (Véronique, 2017 ; Véronique *et al.*, 2017) soit vers l'exploitation quantitative des corpus d'apprenants, dans l'objectif de développer des outils ciblés, moins développée en français qu'en anglais à ce jour (De Cock et Tyne, 2014).

En didactique du FLP, malgré les sérieux défis que peut poser la constitution de corpus non normés d'apprenants (Boré & Elalouf, 2017 ; Doquet *et al.*, 2017), nombreux sont les chercheurs en didactique de la langue qui consacrent une partie de leur temps de recherche à la constitution de corpus d'apprenants, dans une démarche souvent artisanale constituée de repérages et codages manuels dont l'investissement en temps souligne *in fine* l'intérêt des équipes d'enseignants-chercheurs pour cette exploration.

Depuis 2005 et l'initiative de Marie-Laura Elalouf et Catherine Boré qui ont appelé à la diffusion de corpus scolaires, plusieurs corpus scolaires de grande envergure ont été élaborés et diffusés ou sont en cours d'élaboration et de diffusion, en général sur *Ortolang*. Par exemple, le corpus *Ecriscol* (David et Doquet), le corpus ÉMA, écrits scolaires (Boré, Roubaud et Elalouf), le corpus *Resolco* (Garcia-Debanc et Bonnemaïson), le corpus *Dynascript* (Vogüé *et al.*) ou le corpus *Scoledit* (Ponton, Brissaud, Totereau et Wolfarth)², le corpus *Littératie avancée* (Rinck, Jacques et

² Cette liste s'appuie sur l'état des lieux exhaustif et récent et effectué par Claire Wolfarth dans sa thèse (2019).

Boch). La plupart de ces corpus sont consacrés aux écrits d'élèves et cherchent à donner une image de l'écriture scolaire dans différentes dimensions, dont celle, notable, de la progression de la compétence scripturale.

Même si certains de ces corpus d'écriture scolaire en contexte écologique sont utilisés pour décrire la production linguistique des élèves (Cappeau et Roubaud, 2018), la démarche complémentaire visant à construire des corpus scolaires de grammaire pour donner une visibilité des compétences grammaticales des élèves et de leur progression est plus débutante. Et cela, alors que les enjeux du développement de ces compétences et de la façon dont elles sont enseignées et mises en œuvre ont pris une importance renouvelée dans l'économie générale de la formation des élèves en français et en particulier dans leur accession à la conceptualisation de la langue. Or, rares sont encore les grands corpus en français destinés à l'analyse de la conceptualisation grammaticale des élèves et aux obstacles qu'elle rencontre.

Ce n'est que récemment, en 2019, que l'équipe de *Realang* pilotée par Patrice Gourdet, Morgane Beaumanoir-Secq et Jean-Pierre Sautot (Sautot *et al.*, 2021) s'est lancée dans l'élaboration de deux corpus quantitativement imposants, l'un en 2019-2020, l'autre en 2022-2023. Ceux-ci regroupent des données autour de l'appréhension des catégories du verbe et de l'adjectif par les élèves de CE2, CM1 et CM2, qui se complètent de données en 6^e et en 3^e en 2022-2023. Le premier de ces corpus avec son outillage statistique a déjà permis de mettre au jour des découvertes sur ces compétences scolaires et de réaliser de nombreuses publications (voir *Scolagram* 8) ; le second est en cours d'exploitation.

Cependant, une revue de littérature permet de rendre compte de l'existence de plus petits corpus, rarement diffusés, constitués au gré de thèses ou de recherches plus ciblées. Citons par exemple, Voiriot-Cordary (2005), Geoffre (2013) Beaumanoir-Secq (2016), Bonnal (2016), Camussi-Ni (2024), Le Levier (2019), Furman (2022), Tallet (2012) mais aussi les travaux de Danièle Cogis (2004) qui s'appuie sur un corpus constitué par des évaluations nationales en 6^e, ceux de Catherine Brissaud (2009) (Brissaud et Sandon, 1999). (Brissaud *et al.* 2006) qui interroge la grammaire à travers l'orthographe, ceux de Jean-Pierre Sautot et Solveig Lepoire-Duc (2017) et, au Québec, de Marie-Claude Boivin (Boivin *et al.* 2017) (Boivin et Pinsonneault, 2018). Ces quelques exemples témoignent de l'intérêt pour la constitution de corpus constitués de productions grammaticales d'élèves souvent associées à leurs commentaires métalinguistiques.

Enfin, l'intérêt des chercheurs en didactique de la langue se porte aussi aujourd'hui sur la création de corpus d'enseignement, d'autant que cette démarche est préconisée par les Instructions Officielles en France. Fabriquer un corpus d'enseignement, c'est en somme extraire de la langue un « échantillon » qui rende compte des caractéristiques d'un fait de langue, caractéristiques qui seraient diluées dans le corpus plus large, celui de l'usage, auquel est soumis l'élève ; ce peut être aussi repérer et déconstruire une difficulté récurrente dans l'appréhension de la

langue. Jean-Louis Chiss et Claude Muller définissent ainsi cette démarche :

L'activité métalinguistique en classe suppose de problématiser des données de langue au lieu d'en fournir immédiatement la description réglée, de s'interroger sur l'appareil notionnel existant et donc de prévoir, à côté des leçons magistrales, des démarches effectives de recherche apprenant à poser un problème à partir d'un corpus, d'une règle, d'une définition et à réunir les conditions permettant de le résoudre. (Chiss et Muller, 1993, p. 47)

En FLES, l'adaptation de grands corpus pour en faire des objets didactiques peut relever de cette démarche de création de corpus. Mais certains corpus ont également été directement créés pour cet objectif pédagogique comme le corpus FLORALE à Lausanne (Surcouf et Ausoni, 2018). Il existe aussi d'autres initiatives de corpus forgés pour déconstruire les difficultés de la langue pour des apprenants allophones (Camussi-Ni et Coatéval, 2013).

En FLP, la réflexion sur les corpus d'exemples forgés proposés aux élèves émerge depuis une quinzaine d'années (Brissaud et Grossman, 2009). Elisabeth Nonnon (2010) souligne que les corpus utilisés face aux élèves le sont souvent sans conscience des enjeux :

L'analyse préalable doit passer aussi par un questionnement sur les corpus pour étudier un point de langue : en contrôler la pertinence par rapport au fonctionnement visé, distinguer occurrences courantes et rares, simples et complexes, repérer les cas problématiques pouvant être stimulants pour l'apprentissage et ceux qui sont aporétiques. Les corpus pour étudier la langue, les exemples des grammaires sont souvent simples, prototypiques même aux niveaux avancés : ils ne confrontent pas les élèves à des cas problématiques où interfèrent plusieurs modèles ou stratégies. Un traitement peu contrôlé de la progressivité des corpus peut ainsi interférer avec les enseignements explicites et leur visée. (Nonnon, 2010, p. 14)

De leur côté, Parisi et Grossmann (2009) analysent très finement les écueils de la démarche inductive attachée à ces corpus d'enseignement et en particulier son exigence pour l'enseignant :

La démarche dite inductive, correspond bien en didactique à une stratégie visant à ce que les élèves réfléchissent et raisonnent « pour de vrai ». Entendue en ce sens, elle est particulièrement exigeante, puisqu'elle requiert du maître qu'il envisage avec précision les obstacles possibles au cheminement intellectuel de l'élève, y compris en ce qui concerne les éléments liés à la notion linguistique elle-même. (Parisi et Grossmann, 2009, p. 168)

Pour autant, la plupart des recherches en didactique de la langue mettent l'accent sur l'apprenant, l'efficacité et les conditions de son apprentissage. Rares sont en définitive les études portant sur la formation des enseignants³, comme si la

³ Des travaux focalisés sur l'enseignant existent cependant qui peuvent produire des préconisations pour leur formation. Citons-en quelques-uns : Goigoux (2007), Bucheton (2009), Elalouf et al (2005, 2017), Schneuwly & Dolz (2009). Plus rares sont ceux qui sont centrés sur l'analyse de la formation

« centration sur l'apprenant » préconisée dans les démarches pédagogiques venait occulter le travail du maître, ses compétences, ses savoirs et en définitive, la transposition didactique qu'il met en œuvre. Cet oubli de l'enseignant et de sa nécessaire formation est encore plus à risque dans le cadre de l'exploitation de corpus. Ceux-ci peuvent en effet véhiculer avec eux les mythes de la technologie triomphante et en l'occurrence, celui d'outils permettant de se passer de passeurs. Or, sauf à postuler une société où les élèves seraient enseignés par des robots (l'IA dirons-nous aujourd'hui), il est urgent de réfléchir aux transmetteurs que sont les enseignants et à leur formation. Si l'objectif de tout enseignement est bien de conduire à une autonomie de l'apprentissage, cette autonomie ne peut s'acquérir seul et d'emblée, et elle demande d'activer une compétence d'autonomisation qui met en jeu un enseignant et un enseigné. L'article de Taous *et al.* dans ce numéro décrit bien ce que peut être cette démarche d'autonomisation d'un apprenant quand il s'agit de savoir exploiter un corpus. Ce numéro de *Scolagram* vient lever cette illusion d'un outil tout puissant permettant de négliger la formation des enseignants. Il met en lumière l'exigence en termes de connaissances linguistiques, d'analyses didactiques que requièrent l'utilisation de corpus, qu'ils soient forgés ou exploités. Dans l'autre sens, l'utilisation réelle de corpus dans les classes, y compris dans les classes de FLE, reste minoritaire comme l'analysent Cavalla et Loiseau (2013) soulignant le manque de temps des enseignants pour s'emparer de ces outils mais aussi le manque de formation :

Les filières FLE universitaires françaises ne proposent que peu de cours présentant ces outils et donc les techniques pour exploiter les corpus en classe de FLE.[...] Ainsi, malgré l'enthousiasme autour des corpus dans l'enseignement des langues, et la diversité des usages avérés, on se heurte à des difficultés dans la mise en place des activités, dans le choix des outils et les possibilités d'interrogation. (Cavalla et Loiseau, 2013, p 3 et 5)

Dans ce domaine aussi, la différence est grande entre les avancées de la recherche en didactique et la réalité du terrain.

Aussi la volonté de ce numéro de *Scolagram* était-elle d'explorer l'intérêt de l'usage des corpus pour la formation en mettant l'accent sur la formation de l'enseignant qui permette de le mettre véritablement en œuvre, avec l'exigence d'une telle démarche qui demande d'activer une forme de mise en abîme, celle qui parie sur les conséquences auprès des élèves de la formation des enseignants. Quelle place les corpus peuvent-ils avoir dans la formation des enseignants en didactique de la grammaire ?

Les auteurs ayant contribué à ce numéro ont relevé ce défi en interrogeant l'usage qu'ils font des corpus auprès d'enseignants, que ce soit en formation initiale ou en formation continue. Ils nous ouvrent généreusement leurs archives, dévoilant les secrets de fabrication de leur séquence de formation d'enseignants, les obstacles qu'ils rencontrent, en termes de pré-acquis, de préjugés ou d'impacts de questions matérielles comme le temps alloué, les situations de formation, en présence ou à

distance, le cadre institutionnel de la formation. Les articles sont souvent accompagnés d'annexes qui donnent à voir les corpus qui sont traités et qui sont autant de précieux témoignages. Il pourra être utile dans certains cas d'explorer ces annexes avant même de lire ou relire l'article pour en saisir toute la richesse.

Ce numéro regroupe huit articles qui retracent des expériences variées dans des contextes différents et des temporalités différentes : formation initiale d'enseignants du primaire ou du secondaire ou encore d'enseignants de FLES, formation continue d'enseignants du primaire ou de lycée. Tous ont en commun de réinterroger l'enseignement de la grammaire en utilisant des corpus, en confrontant les représentations ordinaires sur des notions grammaticales à la réalité d'énoncés de corpus. Les notions explorées dans ce numéro sont naturellement variées : de la définition de l'adjectif pour l'article de Solveig Lepoire-Duc et de Marie-Noëlle Roubaud, des questions induites par le morphème *on* dans celui de Marie-Laure Elalouf, de la différence d'usage entre les deux prépositions *parmi* et *entre* pour l'article d'Houssein Taous, Emmanuelle Canut et Henry Tyne, des définitions réductrices de la phrase complexe dans l'article de Magali Durieu-Gardelle et celui d'Anne Sardier et de Dominique Vergne, des accords dans le groupe verbal et le groupe nominal dans l'article d'Agnès Furman, des unités morphologiques dans celui de Marie-Armelle Camussi-Ni, de l'interrogation et de la négation dans celui d'Isabelle Vallejo-Malgouyres.

Il est notable que la majorité des auteurs de ce numéro, répondant peut-être au sentiment d'un manque sur le terrain, aient cherché à mettre l'accent sur la constitution de corpus pour l'enseignement que ce soit sous forme de didactisation d'extraits d'un grand corpus ou de création du corpus. L'article d'Houssein Taous, Emmanuelle Canut et Henri Tyne rend bien compte des opérations qui doivent être réalisées à partir d'un corpus brut (celui du journal *l'Est Républicain*) pour le didactiser : choix des énoncés extraits du corpus pour le condenser en un échantillon représentatif de la langue ou de la notion à déduire, organisation de ces énoncés pour rendre l'induction possible, élaboration de questions sur l'échantillon de corpus sélectionné. Isabelle Vallejo-Malgouyres, Magali Durieu-Gardelle, Agnès Furman, Solveig Lepoire-duc et Marie-Noëlle Roubaud rendent compte d'expériences menées auprès d'enseignants ou de futurs enseignants autour de la fabrication de corpus d'enseignement. D'autres confrontent les étudiants à différents corpus : corpus d'apprenants pour Marie-Armelle Camussi-Ni ainsi que pour Solveig Lepoire-Duc et Marie-Noëlle Roubaud ; corpus d'exemples forgés et de textes authentiques pour Marie-Laure Elalouf d'une part, Anne Sardier et Dominique Vergue, d'autre part. Mais beaucoup des auteurs optent pour une confrontation de différents types de corpus, avec des logiques différentes. Il peut s'agir de repérer des obstacles à l'apprentissage pour fonder la création de corpus d'apprenants (Camussi-Ni, Lepoire-Duc et Roubaud, Vallejo-Malgouyres) ou inversement de déconstruire des préjugés de futurs enseignants en les confrontant à une réalité plus scientifique de la langue par le biais de corpus authentique avant de les conduire à constituer des corpus d'enseignement (Elalouf).

De fait, plusieurs études présentées parient sur l'isomorphisme de la situation d'enseignement faite aux enseignants, supposant que l'expérimentation de corpus les

amènent à prendre conscience d'un autre rapport à la langue et à la grammaire pour eux-mêmes et pour leurs élèves. Et en effet, les études présentées dans ce numéro de *Scolagram* montrent les corpus comme susceptibles de bousculer, transformer l'appréhension de la langue et de son enseignement, à la façon dont la linguistique de corpus est venue bousculer certaines représentations de la langue bien établies et continue de le faire :

- Marie-Laure Élalouf initie la réflexion de ses étudiants à partir d'un usage considéré comme « non normé » par la grammaire traditionnelle en cours à l'école, celui du morphème « on », ce qui le rend quasiment invisible car directement évacué. La convocation de textes différents et l'injonction à l'analyser selon différentes entrées suscite du débat, une verbalisation. Ainsi, les futurs enseignants sont-ils amenés à faire un pas de côté par rapport à la posture normative et à tester un autre rapport à la grammaire, celui qui, avec ses enjeux d'élucidation, en fait un objet passionnant de débat.

- La démarche de Solveig Lepoire-Duc et de Marie-Noëlle Roubaud consiste à faire passer aux futurs enseignants les mêmes tests que ceux passés par les élèves du premier degré dans l'étude *Realang*, avant de les confronter à des extraits de ce corpus, soigneusement choisis pour leur potentiel de remise en question des définitions. L'étude montre la proximité des représentations initiales des étudiants et des élèves et tout l'intérêt du corpus pour venir les confronter.

- Isabelle Valléjo-Malguyres explore également les représentations initiales sur la grammaire d'enseignants de lycée en activité, au moment où l'introduction récente de la grammaire dans les programmes en France vient bousculer les enseignants de lettres. Son utilisation de corpus consiste également en une forme de déstabilisation de postures transmissives et de savoirs sur la langue peu interrogés par les enseignants. C'est l'image même de la grammaire auprès des enseignants qui est mise en jeu par l'utilisation de corpus dans son article.

- L'article d'Anne Sardier et Dominique Vergne met en cause également les corpus forgés proposés dans des documents de l'Éducation Nationale en soulignant comme leur simplification vient faire obstacle à l'appréhension de la phrase complexe dans toute sa complexité. Leur expérience auprès d'un grand groupe d'étudiants tend à montrer qu'ils viennent annihiler la capacité à prendre en compte la complexité de la subordination, comme la récursivité des enchâssements.

- L'article de Magali Durieu-Gardelle traite également de l'enseignement des phrases complexes et des questions sur la langue qui surgissent au cours de discussions sur l'emploi de corpus, alors que ces questions seraient occultées dans un enseignement de la grammaire « ordinaire ». Sa recherche collaborative amène à pointer la question intéressante des risques encourus par les élèves dans leur conceptualisation de la langue du fait de la simplification des notions enseignées dans l'enseignement reçu au premier degré. Le corpus et les définitions limités destinés aux élèves deviennent ceux

des enseignants et les verbalisations suscités par ces corpus réélaborés dans des échanges collaboratifs viennent réouvrir ces limites.

- L'article d'Houssein Taous, Emmanuelle Canut et Henri Tyne vient également remettre en cause l'illusion de synonymes parfaitement équivalents. Il déconstruit cette notion de synonymie à propos de l'usage des prépositions *entre* et *parmi* et cherche à cerner le potentiel de corpus en fonction des types de guidage.

- L'article de l'auteur de ces lignes (Marie-Armelle Camussi-Ni) vient délibérément utiliser les corpus d'apprenants allophones pour déconstruire et reconstruire les représentations sur la langue d'étudiants en didactique du FLES. Les exemples *a contrario* fournis par un corpus de locuteurs allophones mettent au jour des caractéristiques de la langue ignorées par la grammaire traditionnelle mais reconnues par des descriptions linguistiques.

- L'étude qualitative d'Agnès Furman montre les écarts entre les attentes des étudiantes testant leur corpus forgé et les réflexions des élèves. Il souligne ainsi que les corpus ne sont pas des objets dociles mais bien des objets heuristiques susceptibles de cheminements voire de découvertes inattendues. L'usage des corpus est donc un lieu d'apprentissage de la souplesse ... et de la réflexion.

Quel que soit le public visé *in fine* et la notion que vise à construire cette exploration de l'usage des corpus de grammaire par des enseignants du français, un certain nombre de points forts ressortent également de la lecture conjointe de ces articles :

- Se former à l'utilisation d'un corpus auprès d'apprenants qu'ils soient francophones ou allophones, ne va pas de soi. Les auteurs s'interrogent sur la fabrication de corpus, qu'elle corresponde à créer de toutes pièces des énoncés destinés à exposer toutes les caractéristiques d'une notion ou à extraire d'un corpus authentique, quelques énoncés, il y a une démarche artificielle qui se met en place qui correspond à la didactisation d'un objet, didactisation nécessaire mais qui demande d'en analyser les contours. Le traitement de données authentiques ne se fait pas sans compétences linguistique, technique et didactique. La fabrication de corpus à valeur heuristique pour des élèves n'est pas naturelle, elle s'enseigne et cet apprentissage demande du temps. Les articles qui témoignent d'expériences dans des temporalités variées permettent d'esquisser ce temps nécessaire.

- Les articles ici réunis témoignent d'un intérêt pour les corpus comme objets permettant de mettre à distance un enseignement transmissif, descendant et parfois éloigné de la réalité de la langue. Les corpus conduisent à un changement de posture de l'enseignant. L'enseignant mis en situation d'explorer un corpus, teste la démarche inductive, qui conduit à faire verbaliser les élèves. Les corpus montrent leur capacité à mettre les futurs enseignants en débat pour leur permettre de s'extraire de leur « corset de l'étiquetage » (Elalouf).

- Les corpus viennent souvent interroger les savoirs des enseignants. L'utilisation de corpus permet de montrer l'aspect réducteur de ce qui est présenté dans les manuels destinés aux élèves qui sont pourtant, le soulignent

Elalouf et Vallejo-Malgouyres, les outils de référence des enseignants qui ont du mal à se saisir d'outils plus scientifiques. La confrontation aux corpus, de différents types, a des effets sur des points de vue occultés, de fausses représentations des enseignants.

Au total, cette collection d'articles décrit plusieurs démarches possibles d'exploitation de corpus à mener dans le cadre d'une formation d'enseignants et surtout souligne l'intérêt d'introduire de telles démarches dans un modèle de formation des enseignants encore à créer en France.

Il reste quelques questions ouvertes à l'issue de ce parcours : si l'utilisation de corpus peut permettre de changer le regard sur la langue et en particulier de valider l'existence de variantes, dans quelle mesure les corpus d'oral spontané ont-ils leur place dans une formation à l'enseignement de la grammaire ? L'appréhension de la grammaire de la langue par l'oral, par exemple des flexions audibles à l'oral, peut-elle s'appuyer sur des corpus ?

Est-il envisageable de viser une autonomisation de l'apprenant en langue première en lui apprenant à utiliser des grands corpus du français disponible comme l'ambitionne la recherche en didactique du FLE ? Le modèle d'enseignement *data-driven learning* utilisé en Français Langue Étrangère (Boulton et Tyne, 2014) peut-il fournir des pistes pour la formation en FLP ou pour la formation des enseignants ?⁴

Qu'en est-il des résultats extraits de l'analyse de corpus par des linguistes ? Ces résultats sont-ils transposables dans une démarche didactique utilisant des corpus, comme ici pour l'adjectif (Delapoire-Duc et Roubaud) ou la collaboration menée avec Juliette Thuilier à partir de sa thèse (Thuilier, 2012 ; Camussi-Ni *et al.*, 2016) ? De la même façon qu'on peut souligner l'intérêt de collaborations entre des domaines différents de la recherche en didactique, comme celle du FLE et du FLP, il est souhaitable de voir se développer plus de collaborations entre didacticiens de la langue et linguistes, acquisitionnistes pour conforter la scientificité et la progression de ce qui est enseigné.

Bibliographie

- André, V. (2020). Faire de la linguistique de corpus avec des apprenants de français langue étrangère, in P. Larrivée & F. Lefevre. La didactisation du français vernaculaire. Caen : Presses Universitaires de Caen. 37-66.
- Beaumanoir-Secq, M. (2016). *Le tri de mots : pour une grammaire utile aux élèves, dans la continuité et la cohérence*. [Thèse de doctorat, Université de Cergy-Pontoise], sous la direction de M.-L. Elalouf.
- Beaumanoir-Secq, M. (2018). *Conceptualiser les classes de mots - A la recherche d'une grammaire utile aux élèves, dans la continuité et la cohérence*. Gramm-R, Peter Lang.

⁴ Le numéro 24/2023 de la revue *Corpus* est ainsi consacré à la question de la formation des enseignants à l'usage des corpus numériques pour la didactique des langues.

- Bernié, J.-P. et Goigoux, R. (coord.) (2005). Les gestes professionnels. *La lettre de la DFLM* 36. https://www.persee.fr/issue/airdf_1776-7784_2005_num_36_1.
- Bonnal, K. (2016). *L'orthographe telle qu'elle s'enseigne : pratiques d'enseignement de l'accord sujet-verbe observées à la fin de l'école primaire*. [Thèse de doctorat, Université de Toulouse 2], sous la direction de C. Garcia-Debanc.
- Boivin, M.-C., Roussel, K. et Pinsonneault, R. (2017). Phrases complexes et maturité syntaxique : une comparaison entre des écrits d'élèves de 13 et 16 ans. *Lidil* 55. <http://lidil.revues.org/4206>
- Boivin, M.-C. et Pinsonneault, R. (2018). Les erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale des élèves québécois en contexte de production écrite. *Revue canadienne de linguistique appliquée* 21(1), 43-78. <https://journals.lib.unb.ca/index.php/CJAL/article/view/25121>
- Boré, C. et Elalouf, M.-L. (2017). Deux étapes dans la construction de corpus scolaires : problèmes récurrents et perspectives nouvelles. *Corpus* 16. 31-63.
- Boré, C., Roubaud, M.-N. et Elalouf, M.-L. (2022). Corpus ÉMA, écrits scolaires [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage). www.ortolang.fr, v3, <https://hdl.handle.net/11403/ema-ecrits-scolaires-1/v3>.
- Boulton, A. et Tyne, H. (2014). *Des documents authentiques aux corpus : Démarches pour l'apprentissage des langues*. Paris. Didier.
- Bras M. , Vieu L., Joret M., Pépin-Boutin A., Poujade C., et al. (2021). Vers un corpus de textes d'élèves annoté en relations de discours. *Langue française* 211 (3). Écrire de l'école à l'université : corpus, traitements, analyses outillées. 115-129.
- Bras M. et Vieu, L. M. J.-A. P.-B. C. P. C. R. (2021). Vers un corpus de textes d'élèves annoté en relations de discours. *Langue Française*, 211(3):115–129.
- Brissaud, C. (2009). Quels enseignements tirer de quatre-vingt-dix-huit dictées de Troisième ? Du décalage entre prescription et acquisition des élèves. Brissaud, C., Grossmann, F. (coord.) *Repères* 39, La construction des savoirs grammaticaux.
- Brissaud, C., Chevrot, J.-P. et Lefrançois, P. (2006). Les formes verbales homophones en /E/ entre 8 et 15 ans : contraintes et conflits dans la construction des savoirs sur une difficulté orthographique majeure du français. *Langue française* 151. 74-93.
- Brissaud, C. et Grossmann, F. (dir.). (2009). *Repères* 39. « La construction des savoirs grammaticaux ».
- Brissaud, C. et Sandon, J.-M. (1999). L'acquisition des formes verbales en /E/ à l'école élémentaire et au collège, entre phonographie et morphographie. *Langue Française*. 124, 40-57.
- Bucheton, D. (2009). *L'agir enseignant : des gestes professionnels ajustés*. Toulouse : Octarès.
- Camussi-Ni, M.-A. (2024). Le traitement de l'erreur : comment introduire la linguistique dans l'enseignement de la grammaire au cycle primaire ?. *Le Français Aujourd'hui*

225. 97-108.

- Camussi-Ni, M.-A. et Coatéval A. (2013). *Comprendre la grammaire. Une grammaire à l'épreuve de la didactique du FLE*. Presses Universitaires de Grenoble.
- Camussi-Ni, M.-A., Coatéval, A. et Thuilier, J. (2016). Exploiter un travail de corpus sur la place de l'épithète pour élaborer une progression didactique en FLE. Dans A. Kamber & M. Dubois (dirs.). *Corpus, grammaire et français langue étrangère : une concordance nécessaire*. *Linguistik online* 78. 41-54.
<https://bop.unibe.ch/linguistik-online/article/view/2949>.
- Cappeau, P. (dir.) (2021). *Une grammaire à l'aune de l'oral*. Presses Universitaires de Rennes.
- Cappeau P. et Roubaud M.-N. (2018). *Regards linguistiques sur les textes d'élèves (de 5 à 12 ans)*. Clermont-Ferrand : Presses universitaires Blaise-Pascal.
- Cavalla, C. et Loiseau, M. (2013). Scientext comme corpus pour l'enseignement. In Agnès Tutin; Francis Grossmann. *L'écrit scientifique : du lexique au discours. Autour de Scientext*, Presses Universitaires de Rennes.163-182.
- Charaudeau, P. (2009). Dis-moi quel est ton corpus, je te dirai quelle est ta problématique. *Corpus* 8. <http://journals.openedition.org/corpus/1674>.
- Chiss, J.-L. et Muller, M. (1993), *Recherches en didactique de la langue et des discours*, Paris, INRP.
- Cogis, D. (2004). Une approche de la morphographie. L'exemple d'une séquence sur l'accord de l'adjectif. *Lidil* 30.
- Cori, M. et David, S. (2008). Les corpus fondent-ils une nouvelle linguistique ?. *Langages* 171, 111-129. <https://www.cairn.info/revue-langages-2008-3-page-111.htm>.
- De Cock, S. et Tyne, H. (2014). Corpus d'apprenants et acquisition des langues. *Recherches en didactique des langues et des cultures* 11-1. DOI : 10.4000/rdlc.1716.
- Dezutter, O., Elalouf, M.-L. et Le Goff, F. (2021). « Discipline », « progression », « gestes didactiques » Échanges croisés autour de trois concepts stratégiques pour les recherches actuelles en didactique du français. *Pratiques* 189-190. DOI : <https://doi.org/10.4000/pratiques.10060>
- Dolz, J. et Gagnon, R. (dirs). (2018). *Former à enseigner la production écrite*. Villeneuve-d'Ascq : Presses universitaires du Septentrion.
- Doquet, C. et David, J. (2018). Collecter, interpréter, enseigner autrement l'écriture. Analyse linguistique des écrits d'élèves. *Repères* 57.
- Doquet, C., Enouï, V., Fleury, S. et Maziotti, S. (2017). Problèmes posés par la transcription et l'annotation d'écrits d'élèves. *Corpus* 16. <http://www.univ-paris3.fr/corpus-ecriscol-300513.kjsp?RH=1416243625396>
- Elalouf, M.-L. (dir.) (2005). *Écrire entre 10 et 14 ans. Un corpus, des analyses, des repères pour la formation*. Paris, Scérén, CRDP Versailles, CDDP Essonne.

- Elalouf, M.-L. (2012). La didactique de la grammaire dans 20 ans de la revue *Repères*. *Repères* 46. <https://doi.org/10.4000/reperes.86>
- Elalouf M.-L. (2011). Constitution de corpus scolaires et universitaires : vers un changement d'échelle ? », *Pratiques* 149-150. 56-70.
- Elalouf, M.-L. (2017). Utiliser le terme *de* prédicat : quel cout pour les enseignants ? quels gains pour les élèves ?, *Pratiques* 175-176.
DOI : <https://doi.org/10.4000/pratiques.3759>
- Elalouf, M.-L., Gomila, C., Bourhis, V., Peret, C., Avezard-roger, C. et Gourdet, P. (2017). « Le rôle du maitre dans l'étude de la langue au CP : description et analyse de pratiques différenciées », *Repères* 55, 183-204. DOI : [10.4000/reperes.1167](https://doi.org/10.4000/reperes.1167)
- Furman, A. (2022). *L'orthographe des verbes dans les dictées et les rédactions des collégiens*. [Thèse de doctorat, Sorbonne Université], sous la direction de S. Plane.
- Gadet, F. et Guérin, E. (2022). Décrire le 'français tout court' : pourquoi et à quelles conditions ?. *Travaux de linguistique* 84-85, 159-172.
<https://doi.org/10.3917/tl.084.0159>
- Gagnon, R. (2013). De l'analyse de productions écrites d'élèves et de ses usages potentiels pour la formation des enseignants du secondaire en grammaire. *Lidil* 47. <http://journals.openedition.org/lidil/3261>
- Garcia-Debanc, C. (2009). Quand les enseignants débutants enseignent la relation sujet/verbe. De l'analyse didactique de pratiques observées de professeurs des écoles débutants à la détermination d'éléments d'expertise professionnelle pour l'enseignement grammatical. In Simard C., Dolz J. (coord.) *Les pratiques d'enseignement de la grammaire dans la francophonie*. Québec : Presses de l'université Laval. 99-124.
- Garcia-Debanc, C. et Bonnemaïson, K. (2014). La gestion de la cohésion textuelle par des élèves de 11-12 ans : réussites et difficultés, *Actes du Congrès Mondial de Linguistique Française (CMLF 2014)*, Berlin, Germany, 961-976.
- Garcia-Debanc, C., Ho-Dac L.-M., Bras, M. et Rebeyrolle, J.. (2018). Vers l'annotation discursive de textes d'élèves. *Corpus*, 16. Mis en ligne le 09 janvier 2018.
URL : <http://journals.openedition.org/corpus/2783>
- Garcia-Debanc, C., Paolacci, V., Benaïoun-Ramirez, N., Bessagnet, P., Gangneux, M., Beucher, C. et Dutrait, C. (2010), Penser la progressivité de l'enseignement grammatical au cycle 3 de l'école primaire : discours, programmations et préparations de formateurs et de professeurs des écoles stagiaires. *Repères* 41. En ligne : <https://doi.org/10.4000/reperes.297>.
- Garcia-Debanc C. et Sanz-Lecina E. (2008). De l'analyse des modèles disciplinaires en acte à la détermination de schèmes professionnels. L'exemple de l'enseignement de la grammaire au cycle 3 par des professeurs des écoles débutants. In Carnus M.-F., Garcia-Debanc C. et Terrisse A. *Analyse des pratiques des enseignants débutants : approche didactique*. Grenoble : La Pensée sauvage Éditions. 151-170.

- Geoffre, T. (2013). *Vers le contrôle orthographique en cycle 3 de l'école primaire. Analyses psycholinguistiques et propositions didactiques*. [Thèse de doctorat, Université de Grenoble], sous la direction de C. Brissaud.
- Goigoux, R. (2007). Un modèle d'analyse de l'activité des enseignants. *Éducation et didactique* 1 (3). DOI : 10.4000/educationdidactique.232
- Gourdet, P. et Roubaud, M.-N. (2022). Les discours grammaticaux sur le verbe par des élèves de CE2 et de CM2. *Scolia*, 36. <https://doi.org/10.4000/scolia.2225>
- Jacques, M.-P. et Rinck, F. (2017). Un corpus de littéracie avancée : résultat et point de départ. *Corpus* 16.
- Johns T. (1986). Micro-concord: a language learner's research tool, *System* 14(2). 151-162.
- Lefrançois, P. (2009). Évolution de la conception du pluriel des noms, des adjectifs et des verbes chez les élèves du primaire. *Repères* 39.
- Le Levier, H. (2019). *Mise en œuvre et perception de l'orthographe française chez des élèves de collège et des étudiants de sections de technicien supérieur*. [Thèse de doctorat, Université de Grenoble], sous la direction de C. Brissaud.
- Nonnon, É. (1999). "Tout un nuage de philosophie condensé dans une goutte de grammaire" : interactions verbales et élaboration de règles dans la mise en œuvre d'une "démarche inductive" en grammaire. *Pratiques* 103/104, 116-148.
- Nonnon, E. (2010). La notion de progression au cœur des tensions de l'activité d'enseignement. *Repères* 41. 5-34. <https://doi.org/10.4000/reperes.276>
- Paolacci, V. et Garcia-Debanc, C. (2005). Comment former à l'enseignement de la ponctuation. Analyse de pratiques effectives de formation initiale. *Pratiques* 125-126, 85-114. DOI : [10.3406/prati.2005.2062](https://doi.org/10.3406/prati.2005.2062)
- Parisi, G. et Grossmann, F. (2009). Démarche didactique et corpus en classe de grammaire : le cas du discours rapporté. *Repères* 39.
- Roubaud, M.-N. (2017). Le français écrit: transcription et Edition. Le cas des textes scolaires. *Corpus* 16.
- Roubaud, M.-N. et Garcia-Debanc, C. (2014). L'approche d'"anomalies" dans des textes narratifs d'élèves de fin d'école primaire (10-11 ans). Dans Avanzi M. et alii (dirs.). *Enseignement du français : les apports de la recherche en linguistique*. Peter Lang, Bruxelles-Bern, 307-326.
- Sautot, J.-P. et Lepoivre-Duc, S. (2017). Saisir les relations temporelles dans des récits d'élèves : quels enjeux pour l'analyse de corpus ? *Corpus* 16, 265-287.
- Sautot, J.-P., Gourdet, P. et Beaumanoir-Secq, M. (dirs.) (2021). Méthodologie pour l'étude des rendements des classes. *Scolagram* 8, Realang.
- Schneuwly, B. et Dolz, J. (éds) (2009). *Des objets enseignés en classe de français. Le travail de l'enseignant sur la rédaction de textes argumentatifs et sur la subordonnée relative*. Rennes : Presses universitaires de Rennes.

- Sémidor, S. (2015). La genèse orthographique chez l'élève de cours préparatoire : description de la langue et pratiques d'enseignement. *Éducation*. Université de Bordeaux.
- Surcouf, C., et Ausoni, A. (2018). Création d'un corpus de français parlé à des fins pédagogiques en FLE – La genèse du projet Florale. *Études en didactique des langues* 31, 71-91. https://serval.unil.ch/fr/notice/serval:BIB_6F4C4454FC17
- Tallet, C. (2012). *Mots-outils homophones hétérographes. Leur enseignement à l'école primaire*. [Thèse de doctorat, Université Paris 3], sous la direction de Danièle Manesse.
- Thuillier, J. (2012). *Contraintes préférentielles et ordre des mots en français*. Thèse de doctorat, Université de Diderot-Paris VII.
- Vaubourg, J.-P. (2017). *Apprendre à réaliser les accords au cycle 3 de l'école primaire. Aspects linguistiques, psycholinguistiques et didactiques*. [Thèse de doctorat, Université Paris 4], sous la direction de Sylvie Plane.
- Véronique, G.-D. (2017). La grammaire en français langue étrangère : questions d'acquisition et d'intervention. *Lidil* 56. DOI : <https://doi.org/10.4000/lidil.4734>
- Véronique, D. (dir.) (2017). *L'acquisition de la grammaire du français langue étrangère*. Didier.
- Vogüé, S. D., Espinoza, N., Garcia, B., Perini, M., Sitri, F. et Watorek, M. (2017). Constitution d'un grand corpus d'écrits émergents et novices: principes et méthodes. *Corpus* 16.
- Voiriot-Cordary, N. (2005). *Acquisition et gestion de la morphologie verbale flexionnelle en français à l'entrée au lycée*. [Thèse de doctorat, Université de Dijon], sous la direction de Ghislaine Haas.
- Wolfarth, C., Ponton, C. et Totereau, C. (2017). Apports du tal à la constitution et à l'exploitation d'un corpus scolaire. *Corpus* 16.
- Wolfarth, C. (2019). *Apport du TAL à l'exploitation linguistique d'un corpus scolaire longitudinal*. [Thèse de doctorat, Université Grenoble-Alpes], sous la direction de C. Brissaud et C. Ponton.

Corpus :

- Clapi (n.d.). Corpus de langue parlée en interaction. <http://clapi.icar.cnrs.fr>
- Clapi-FLE (n.d.). Plateforme de Clapi dédiée à l'enseignement.
http://clapi.icar.cnrs.fr/FLE/projet_clapi_fle.php
- Eslo (Enquêtes sociolinguistiques à Orléans) (n.d.). <http://eslo.huma-num.fr>
- FLORALE. <https://florale.unil.ch/>
- Scientext : https://corpora.aiakide.net/scientext19/?do=SQ.setCorpus&corpus=sqCorpus__sctexts-fr__sctexts-fr&view=tex



Revue de Didactique de la Grammaire
<http://scolagram.u-cergy.fr/>

Ecriscol : <http://syled.univ-paris3.fr/ecriscol/CORPUS-TEST/>

ÉMA : <https://www.ortolang.fr/market/corpora/ema-ecrits-scolaires/>

Littératie avancée : <https://www.ortolang.fr/market/corpora/litteracieavancee>

REALang : https://scolagram.u-cergy.fr/index.php/content_page/75-realang

Resolco : <http://redac.univ-tlse2.fr/corpus/resolco/>

Scoledit : <http://scoledit.org/scoledition> (CP-CM2 - accès libre)
<http://scoledit.org/scoledit> (CP-CE1)