

Constitution de corpus scolaires et universitaires : vers un changement d'échelle ?

Marie-Laure Elalouf



Édition électronique

URL : <http://journals.openedition.org/pratiques/1702>

DOI : 10.4000/pratiques.1702

ISSN : 2425-2042

Éditeur

Centre de recherche sur les médiations (CREM)

Édition imprimée

Date de publication : 15 juin 2011

Pagination : 56-70

Référence électronique

Marie-Laure Elalouf, « Constitution de corpus scolaires et universitaires : vers un changement d'échelle ? », *Pratiques* [En ligne], 149-150 | 2011, mis en ligne le 16 juin 2014, consulté le 19 avril 2019. URL : <http://journals.openedition.org/pratiques/1702> ; DOI : 10.4000/pratiques.1702

Constitution de corpus scolaires et universitaires : vers un changement d'échelle ?

Marie-Laure ELALOUF

Université de Cergy-Pontoise-IUFM
Ecole, mutations, apprentissages, ÉA 4507

Si la constitution de grands corpus, écrits et oraux, a profondément renouvelé l'appréhension des faits de langue en sciences du langage, les recherches en didactique du français ne connaissent pas encore un tel changement d'échelle. Les raisons en sont sans doute multiples : les unes sont liées au contexte socio-politique dans lequel s'effectuent aujourd'hui ces recherches et notamment la déstabilisation des équipes associant chercheurs, formateurs et enseignants, en IUFM et à l'INRP, mais elles tiennent également à la nature de ces corpus : leur recueil exige disponibilité et précautions déontologiques ; leur présentation nécessite l'articulation de supports sémiotiques différents et doit permettre de saisir conjointement des évolutions individuelles et des dynamiques collectives, tant du point de vue de l'enseignant que des élèves ; leur réception s'expose aux paradoxes du rapport à la norme en milieu scolaire.

Pourtant, des archives se sont constituées pour cerner au plus près ce qui est effectivement enseigné et qui produit des effets différenciés en termes d'apprentissage : travaux du GRIDIFE (groupe de recherche sur les Interactions Didactiques et la Formation des enseignants, Toulouse), de l'ERTÉ *les langages, conditions d'entrée dans les apprentissages*, du laboratoire PAEDI, sans compter les travaux portant sur d'autres systèmes éducatifs francophones. La publication d'articles en ligne, en même temps que leur version papier, rend accessibles des corpus plus vastes. Ce sont trois séries de textes – une sélection de 39 productions écrites au total – que proposent Soledad Valera-Kummer et Caroline Masseron (2009) dans leur analyse d'écrits scientifiques de collégiens. Parallèlement, des outils de formation exploitant des corpus ainsi constitués sont annoncés (Bucheton, 2009). La plate-forme *néopass*, qui repose sur la mise à disposition de ressources réalisées à partir de recherches fondées sur le travail réel des enseignants, a vocation à se développer.

Autant de signes qui indiquent qu'un saut qualitatif devient nécessaire et possible : passer de corpus constitués à des fins de recherches internes à une équipe à

des corpus mis à disposition de la communauté scientifique dans ses missions de recherche, de formation et de contribution à la politique éducative. Réunir des corpus représentatifs, comparables, cumulables est un enjeu essentiel pour la production de connaissances en didactique du français. Par ailleurs, la réduction de l'alternance dans la formation des enseignants oblige à concevoir des dispositifs alternatifs qui confrontent des étudiants à des productions d'élèves, des interactions en classe, des pratiques effectives, toutes choses qu'ils ne trouveront pas s'ils tapent dans un moteur de recherche « poésie CM1 ». Enfin, à un moment où la politique curriculaire est profondément remaniée sans réelle prise en compte des travaux de didactique, la collecte de données empiriques sur une échelle plus large peut mettre à jour l'inanité d'une définition a priori des paliers du socle commun.

J'étudierai dans cet article comment des travaux récents ont problématisé la constitution de corpus à travers l'étude de quatre thèses ; je m'interrogerai sur l'extension et la mise en réseau possible de corpus grâce à la contribution d'écrits professionnels élaboré en master « formation de formateurs », et conclurai sur les perspectives d'élaboration d'un grand corpus sur et pour la formation.

1. La conception du corpus dans quatre thèses récentes

Mon choix s'est porté sur quatre thèses qui ont en commun de présenter des corpus susceptibles d'être lus et exploités de façon quasi autonome, tout en étant très fortement articulés à la thèse par une réflexion méthodologique. Toutes appellent *annexes* ce qui constitue en fait de véritables corpus, l'annexe n'étant qu'une partie des documents de référence ⁽¹⁾, ceux qui servent plus particulièrement à l'exploitation scientifique.

1.1. Présentation des travaux et principes d'élaboration des corpus

La thèse de Bernadette Lafourcade (2008 ⁽²⁾), *Traitement des contraintes formelles liées au genre et au médium de production par des scripteurs novices – Étude didactique*, compte 295 pages d'annexes. Elle comporte les deux versions successives d'une suite de nouvelle proposée à 77 lycéens de seconde (43 en traitement de texte, 34 en support papier), à 7 élèves suivis en aide individualisée ayant travaillé successivement sur les deux supports, et à 30 scripteurs experts à titre de comparaison. Entre la première et la seconde version, les élèves ont connaissance de l'*explicit* ⁽³⁾ de la nouvelle qui appelle une réorientation générique. Ce corpus hétérogène joue donc sur deux variables articulées aux deux hypothèses

(1) On pourra s'appuyer sur les travaux de F. Rastier à propos de la linguistique de corpus pour les transférer et lui emprunter les distinctions entre « archive », « corpus de référence », « corpus d'étude » et « sous-corpus de travail ». Voir F. Rastier, 2005, « Enjeux épistémologiques de la linguistique de corpus », pp. 31-46, in G. Williams, *La Linguistique de corpus*, PUR.

(2) Soutenue à l'université Paris-Est sous la direction de Marie-Emmanuelle Diéval et Sylvie Plane.

(3) *explicit*, antonyme d'incipit.

ses que formule la thèse : l'une postule le caractère facilitateur de la connaissance des genres, la seconde l'effet du traitement de texte sur la révision des productions écrites.

L'ensemble des analyses manuelles et outillées (grâce aux logiciels *Tropes* et *Médite*) sont fournies dans les annexes : il est ainsi permis de confronter les critères élaborés dans le cadre de la thèse à une lecture tabulaire des productions écrites.

La thèse de Marie-Françoise Fradet-Le Coz (2009⁽⁴⁾), intitulée *La construction de la fiction dans l'écriture de textes narratifs à visée littéraire à l'entrée du collège : quand le dialogue pédagogique interfère avec le dialogue intérieur chez les jeunes scripteurs* comporte deux volumes d'annexes de 773 pages au total, et un DVD permettant de naviguer d'un texte à l'autre et d'entendre les interactions entre l'enseignante et les élèves lors de l'élaboration d'un premier jet d'écriture. L'auteur étudie comment une écriture en trois jets, avec évaluation critériée, évolue grâce au dialogue, instauré par les annotations, entre le concepteur de la consigne et l'élève d'une part et d'autre part grâce à la relecture par l'élève de son propre texte. Elle met en évidence les difficultés que suscitent les textes en décalage avec les attentes de l'institution, tant du point de vue des habiletés discursives que de l'invention.

Les principes retenus pour la conception du dispositif permettent des comparaisons rigoureuses : une consigne identique pour les cinq classes mais un jeu sur trois variables : le choix du texte inducteur, le degré d'explicitation de la grille d'évaluation et le changement de médiateur – le chercheur dans les deux premières campagnes d'écriture, puis le professeur titulaire de la classe dans la dernière.

La thèse de Patrice Gourdet (2009⁽⁵⁾), intitulée *L'enseignement de la grammaire à l'école élémentaire : le cas du verbe en CE2, Des savoirs à enseigner aux savoirs appris : un facteur efficient, l'enseignant intègre* dans le corps de la thèse une partie du corpus, qui est complété par 65 pages d'annexes. L'étude des effets d'apprentissage s'appuie sur un travail d'observation et d'expérimentation mené sur deux années scolaires dans cinq classes de CE2 de l'Essonne et concerne 108 élèves en CE2 et 97 en CM1. Le corpus réunit trois outils auxquels les élèves sont quotidiennement confrontés : les affichages collectifs sur le verbe, l'outil référent (ou cahier de règles) dont dispose chaque élève et les exercices proposés, ce qui autorise des synthèses sur l'environnement grammatical de chaque classe pointant la cohérence ou les décalages entre les différents supports.

La passation du même exercice de l'évaluation nationale en septembre, en juin et l'année suivante, en regard des définitions successives du verbe données par les mêmes élèves, permet de mettre en perspective les performances des élèves avec l'environnement grammatical qui a été le leur et de formuler des hypo-

(4) Soutenue à l'université Paris-Est sous la direction de Marie-Emmanuelle Diéval et Sylvie Plane.

(5) Soutenue à l'université Paris-Ouest-Nanterre-La Défense, en cotutelle avec l'université Roma Tre, sous la direction de Danielle Leeman et Emma Nardi.

thèses sur les éléments constitutifs d'un environnement grammatical propice à l'apprentissage du verbe.

La thèse de Claude Beucher (2010⁽⁶⁾) intitulée *L'accompagnement de l'écriture de nouvelles et de fables dans l'enseignement primaire et secondaire. Etudes de pratiques françaises et belges* réunit en deux volumes d'annexes (1026 pages) les dossiers constitués pour chacun des six enseignants suivis. Chaque dossier comporte la transcription complète de la séance observée, avec en vis-à-vis le codage des gestes d'étayage selon une typologie dont l'élaboration est l'enjeu théorique de la thèse. Cette typologie croise quatre dimensions : le degré d'intervention de l'enseignant (de l'absence de réponse à une sollicitation à la correction imposée), le moment de l'étayage (avant, pendant ou après la production écrite), le canal (écrit, oral ou non verbal) et les « lieux d'intervention didactique » (Turco : 1987 et Séguy : 1989) : pragmatique, sémantique, morphosyntaxique et matérielle. À la transcription s'ajoutent l'ensemble des documents fournis aux élèves pour la séance d'écriture (texte, grilles, différentes versions des productions écrites, transcription des entretiens avec l'enseignant et les élèves, questionnaire proposé aux élèves sur leur perception du cours.

Le choix d'enseignants ayant moins de cinq années d'exercice, exerçant à des niveaux différents (du CM1 à la fin du secondaire) et dans deux systèmes éducatifs différents – la France et la Belgique – permet d'interroger la portée heuristique du concept de *modèle disciplinaire en acte* (Garcia-Debanc : 2009).

Ces quatre thèses ont pour caractéristique commune de fournir au lecteur tous les éléments permettant de refaire les analyses présentées et de vérifier la pertinence des choix relatifs à la conception du corpus au regard de la thèse soutenue.

1.2. Modalités de recueil et place du chercheur

La collecte de données dans des classes est coûteuse en temps et son interprétation dépend étroitement des moyens dont dispose le chercheur pour accéder à leur intelligibilité. Les travaux observés procèdent de deux choix différents : soit le chercheur est à la fois celui qui conçoit et met en œuvre – totalement ou partiellement – le dispositif et il se donne les moyens de mettre à distance son expérience ; soit il sollicite des enseignants auprès desquels il recueille son corpus et le complète par des entretiens ou par d'autres données susceptibles de contribuer à son interprétation.

Le premier choix est celui de Bernadette Lafourcade et Marie-Françoise Fradet. Elles conçoivent un dispositif didactique à partir de leurs hypothèses. La mise à distance passe par le codage des scripteurs et la transcription génétique des textes, cohérente avec le cadre théorique retenu pour l'analyse des productions écrites. Cette transcription impose une lecture littérale des textes d'élèves, attentive à leurs graphies, à leurs suppressions, ajouts, modifications et à leur choix de mise en page. Dans le travail de Marie-Françoise Fradet, ce sont également les annotations de l'enseignante qui font l'objet d'une transcription néces-

(6) Soutenue à l'université de Toulouse 2 le Mirail, en cotutelle avec l'université catholique de Louvain, sous la direction de Claudine Garcia-Debanc et Jean-Louis Dufays.

saire à l'analyse du dialogue pédagogique et du dialogue intérieur qui se nouent dans la réécriture.

Ces transcriptions sont complétées par un choix de copies numérisées qui offrent une image de l'écriture manuscrite des élèves et des modalités d'intervention du professeur sur la copie.

Le second choix est celui de Patrice Gourdet et Claude Beucher. Chez Patrice Gourdet, l'observation est écologique. L'étude du verbe est centrale en CE2 et la stabilité des pratiques à l'école primaire fait que toute classe ordinaire comportera des affichages, un cahier de règles et un cahier d'exercices. En revanche, la passation de la même épreuve de l'évaluation CE2 à trois reprises relève d'une intervention du chercheur. Le recueil de séances comparables a conduit Claude Beucher à préférer un mode de recueil plus proche de l'ingénierie didactique : un même album a été proposé aux enseignants qui avaient la liberté de sélectionner certaines images comme inducteur d'écriture, les modalités de travail (écriture individuelle, en groupes ou collective) étant laissées à leur appréciation. Le chercheur étant celui qui filme, il doit faire des choix susceptibles de réduire sa part de subjectivité, ici celui de suivre le professeur avec la caméra pour adopter son point de vue.

Dans chacun des cas, la place du chercheur est pensée en même temps que le recueil des données qui en tient compte.

1.3. Comparabilité des corpus

La comparabilité des corpus est une question à double détente. C'est de la rigueur méthodologique avec laquelle est établie la comparabilité des sous-corpus que découle la possibilité de comparer un corpus à d'autres.

Chez Bernadette Lafourcade, le corpus en écriture manuscrite et celui sur traitement de texte sont recueillis dans des conditions identiques, ce qui autorise des comparaisons portant sur la forme (nombre de production achevées, de variantes, de mots par texte produit) et le contenu (thèmes récurrents, personnages, chronotope, narrateur, genres de discours). En revanche, la consigne de réécriture propose un choix entre trois explicites, ce qui introduit une hétérogénéité calculée : il s'agit de restituer les inférences faites par les élèves selon le choix effectué et d'accéder à leur perception des contraintes génériques.

Dans la thèse de Marie-Françoise Fradet, la détermination de trois sous-corpus permet une lecture en trois étapes, donnant lieu à l'application de grilles d'analyse et de graphes de synthèse :

- « *le sous-corpus A*, qui réunit les premiers jets des deux premières campagnes d'écriture. Sa lecture a pour but de relever les composantes de la fiction construite par les élèves et de les comparer avec celles qu'impose ou que n'impose pas la commande d'écriture. Cette comparaison permet de faire une première synthèse des "réponses" des jeunes scripteurs, de mesurer le degré de conformité de leurs fictions à la consigne mais aussi l'étendue de leur singularité.
- *le sous-corpus B* qui rassemble les deux premiers jets annotés ainsi que le troisième, sans annotation, de ces mêmes campagnes. On y observe les mo-

difications apportées par le dialogue qui s'est noué avec l'enseignant-expérimentateur à travers les annotations de copies et les grilles d'évaluation. Un nouveau panorama des "réponses" des jeunes scripteurs est ensuite dressé pour rendre compte du rôle joué par les interventions du professeur sur la construction de la fiction des élèves.

– *le sous corpus C* enfin, qui regroupe cette fois-ci les trois jets (les deux premiers annotés, le troisième sans annotation) de la troisième campagne. Ils sont lus en même temps car on vise principalement dans cette dernière étude à mesurer l'impact du changement de médiateur ⁽⁷⁾ sur les modifications constatées à l'examen du *sous-corpus B*. Un dernier bilan des "réponses" des jeunes scripteurs est établi afin de comparer les résultats obtenus par le titulaire à ceux obtenus par l'enseignant-expérimentateur. (p. 154) »

Dans le doctorat de Patrice Gourdet, chaque sous-corpus est constitué des mêmes éléments hétérogènes, qui ne sont pas nécessairement pensés en cohérence par l'enseignant et dont il s'agira d'établir la comparaison : les affichages et « l'outil référent » (ou cahier de règles) qui constituent l'environnement grammatical explicite et les exercices qui forment l'environnement grammatical implicite, c'est-à-dire les propriétés du verbe que les élèves sont censés induire des tâches proposées. C'est le croisement avec le codage des propriétés du verbe qui permet de caractériser chaque sous corpus :

- « la classe A : un environnement grammatical pluridimensionnel avec quelques décalages ;
- la classe B : un environnement grammatical orienté ;
- la classe C : un environnement grammatical réduit mais riche, et cependant déséquilibré ;
- la classe D : un environnement grammatical chaotique ;
- la classe E : un environnement grammatical disjoint. »

Au terme d'une analyse systémique, des décalages apparaissent en effet entre outils au sein d'une même classe, voire au sein du même affichage, comme on peut le voir en annexe pour la classe A (p. 360).

Dans la thèse de Claude Beucher, le principe de comparabilité se fonde sur la transcription de séances filmées en continu, c'est-à-dire sans montage. Le choix d'une transcription *in extenso*, comportant trois observables – les discours oraux, les modifications des écrits par les annotations et les éléments non verbaux – est justifié ainsi dans la partie méthodologique : « L'ensemble de notre corpus est codé, rares sont les passages qui n'entrent pas dans les analyses et les comptages. Ce qui semble inintéressant de prime abord trouve souvent un intérêt plus tard, nous ne voulions pas courir le risque de perdre de l'information » (pp. 230-31). Le codage des locuteurs et des tours de parole, l'indication du déroulement du temps sont autant d'éléments qui, couplés avec le codage des différentes formes d'étayage, autorisent la comparaison des sous-corpus.

Les choix méthodologiques posés dans ces différents travaux permettent d'envisager la collecte de nouveaux sous-corpus, reprenant les mêmes protocoles

(7) En effet, les textes des deux premières campagnes sont évalués par le chercheur et ceux de la dernière par le professeur de la classe.

les ou jouant sur une variable, de façon à éprouver, discuter, amender les thèses défendues dans ces travaux. Des comparaisons peuvent aussi s'établir à un autre niveau. En effet, par-delà la diversité des corpus présentés ici, des constantes apparaissent. Un premier principe, qui reste implicite, mais fonde l'approche didactique est que l'unité d'observation est la classe. Aucune production n'est écartée. L'unité de temps correspond aussi à la temporalité des apprentissages scolaires : séquence d'écriture, conceptualisation attendue à l'issue d'une année scolaire.

On constate au terme de cette étude que l'hétérogénéité des corpus n'est pas un obstacle à leur comparabilité mais relève d'une nécessité épistémologique, celle de restituer un contexte complexe, poly-sémiotique et de comprendre le jeu des logiques subjectives dans un cadre didactique.

1.4. Singularité/ représentativité des corpus

Les travaux présentés ici ont la richesse d'études de cas sans pour autant se limiter à une succession d'études de cas. Pour autant, ils ne peuvent prétendre à une représentativité statistique. L'intérêt de ces approches qualitatives est de donner à voir la singularité des pratiques, du côté de l'enseignant comme des élèves, tout en dégagant des éléments de typicité. C'est ainsi que Bernadette Lafourcade, en croisant son analyse des textes de lycéens avec les résultats du logiciel *MEDITE*, définit des profils de scripteurs contrastés. Par exemple, « Caroline : supprimer, insérer et remplacer pour assurer une meilleure cohérence ; Emilie : supprimer et remplacer en négligeant le problème de la cohérence » (pp. 324-7). Marie-Françoise Fradet dégage les caractéristiques des textes qui s'accordent à la consigne, directement, ou au bout de deux ou trois jets et celles des textes qui ne s'y accordent pas, avant d'extraire à des fins d'analyses deux productions qui illustrent des postures différentes. Du côté de l'enseignant, cette thèse met en évidence la grande stabilité des pratiques de correction avec leur palette limitée et polyvalente d'annotations (*mal dit, impropre*, etc.).

Dans la thèse de Claude Beucher, la recherche d'invariants se fait en comparant des séances de différents niveaux, dans deux systèmes éducatifs différents au sein d'un même genre et d'un genre à l'autre. Patrice Gourdet quant à lui exploite les explications linguistiques des élèves pour distinguer les profils explicatifs des classes selon leur environnement didactique, mais aussi les profils explicatifs des élèves dont l'évolution différenciée milite en faveur d'un environnement grammatical articulant explicitement plusieurs niveaux d'analyse dont la syntaxe. De façon complémentaire, les performances aux mêmes exercices de l'évaluation nationale portant sur le verbe (en début de CE2, fin de CE2 et fin de CM1) permettent de caractériser des profils de classe, en relation avec l'environnement grammatical qui est le leur et de dégager l'effet de l'enseignant. L'articulation entre données qualitatives et quantitatives est rendue possible par le recours à cet étalon qu'est l'évaluation nationale.

1.5. La reproductibilité des analyses

La validité des analyses dépend de leur reproductibilité. Or, comme le rappellent Soledad Valera-Kummer et Caroline Masseron, on ne peut refaire une expé-

rimentation en sciences humaines car les conditions ne sont pas reproductibles mais on peut fournir les éléments permettant de refaire les analyses. C'est ce qu'elle fait dans l'article déjà cité : « Nous avons pris le parti de consigner en annexe plus de textes que nous ne saurions en étudier dans le cadre de l'article : 11 descriptions, 15 comparaisons et 13 rapports d'expérience. Nous avons souhaité, ce faisant, laisser une certaine autonomie de jugement au lecteur, et, par ailleurs, livrer le corpus nécessaire pour forger des indicateurs » (Soledad Valera-Kummer et Caroline Masseron, 2009 : 88).

Les quatre thèses étudiées font le même choix, que Claude Beucher formule ainsi (p. 231) « Nous voulions aussi offrir la plus grande connaissance possible de ce qui a été effectivement dit et fait dans la classe à un observateur extérieur qui voudrait critiquer des points de nos analyses ». Outre les transcriptions grâce auxquelles les analyses peuvent être refaites, les études intermédiaires, manuelles ou outillées sont présentes dans les corpus de Bernadette Lafourcade et Marie-Françoise Fradet, sous forme de tableaux et de graphes. Quant à Patrice Gourdet, il fournit systématiquement le codage des propriétés du verbe dans les explications linguistiques des élèves et des adultes, après un double traitement en aveugle, ainsi que celui des réponses aux exercices de l'évaluation. Tous sont conscients que les principes méthodologiques qui ont présidé à l'élaboration de leur corpus devraient permettre son extension et/ou son exploitation pour d'autres problématiques.

2. La construction de corpus comme moyen de formation

Si la constitution de corpus a pour finalité la production de connaissances dans des travaux de doctorat, elle peut viser d'autres finalités dans des écrits professionnels : développer conjointement une initiation à la recherche et une intelligence de sa propre pratique. C'est l'option qui a été retenue dans un diplôme universitaire de formateur d'enseignant en didactique du français, ouvert en 2008 sous ma responsabilité, à l'université de Cergy-Pontoise et intégré en 2010 dans le parcours « formation de formateurs » de la spécialité « conseil, éducation, management, formation, recherche » de la mention « éducation et formation ».

2.1. L'écrit professionnel, comme modalité de formation

Dans cette formation qui correspond à un master 2 en deux ans, deux écrits professionnels jalonnent l'itinéraire réflexif proposé aux enseignants-formateurs (désormais désignés par le terme générique d'étudiants), de l'analyse de leur propre pratique à la conception d'un outil d'aide à un débutant. Le premier écrit doit leur permettre une analyse de leur pratique professionnelle, en relation avec les apprentissages des élèves. À partir d'une question de didactique abordée en formation et qui rencontre une préoccupation professionnelle de l'étudiant, il s'agit de constituer un corpus selon une méthodologie explicitée, en s'appuyant sur la littérature professionnelle. La constitution de ce corpus, son organisation et sa transcription, accompagnées dans le cadre d'ateliers, confron-

tée aux choix des pairs, installent un double mouvement d'objectivation de la pratique et d'élucidation des schèmes d'action (Goigoux & Vergnaud, 2005).

L'étape de la transcription s'est révélée décisive en termes de formation : dans le cas des transcriptions d'oral, elle permet la réévaluation des spécificités de la langue orale, au-delà des stéréotypes, la mise en évidence de la place prise par la parole du maître et des dysfonctionnements dans les interactions ; dans le cas des transcriptions d'écrit, elle exige une attention à la lettre des écrits d'élèves et simule une entrée dans le processus d'écriture. On a pu mesurer la différence entre des corpus qui accumulent les documents numérisés, avec un effet d'authenticité certes, mais un grain d'analyse assez grossier, et des corpus s'appuyant sur des transcriptions reprises dans le corps du mémoire et croisées à des références théoriques.

Le second écrit professionnel accompagne le changement de posture de l'enseignant vers le formateur. Il doit permettre de concevoir, si possible en binôme ou à plusieurs, un outil de formation pour un enseignant débutant, dont l'objectif est de développer chez celui-ci une attitude réflexive. Cet outil de formation est conçu à partir des observations faites en suivant un débutant volontaire au premier semestre de la seconde année, en lien avec les cours sur la formation des pratiques professionnelles. Il est également constitué d'un corpus, mais destiné cette fois à développer une démarche d'observation et d'analyse chez le débutant.

Qu'il s'agisse du premier ou du second écrit professionnel, la constitution du corpus est le moment où se construit une représentation de la pratique effective, loin des rationalisations *a posteriori* uniquement centrées sur le maître, et où sont rendues possibles des analyses référées à des travaux de recherches.

2.2. Des corpus qui enrichissent des recherches

Comment les étudiants s'emparent-ils de ces recherches ? La recherche d'autorités est une pente naturelle chez ces professionnels qui ont plus couramment accès à des travaux de seconde main ou des manuels. Le changement de posture imposé par la constitution de corpus appelle un autre rapport aux travaux de recherches, comme outils d'analyse. C'est chez Aïssatou Bah, le choix de reprendre un dispositif présenté dans *Apprentissages progressifs de l'écrit* (2000) – le travail sur l'album *Sissi la souris* – pour analyser la façon dont le rappel de récit permet à des enfants de moyenne section de progresser dans la compréhension du langage écrit, en articulant des activités en classe entière, en atelier et en aide personnalisée. Mais le travail ne se réduit pas à une simple application. À partir de la transcription des rappels de récits de quatre élèves suivis en aide individualisée, à trois dates différentes, l'étudiante construit une grille d'évaluation des compétences narratives et de l'emploi des outils linguistiques qui lui permet de cerner des évolutions, de les confronter à celles observées en classe entière et d'identifier, parmi les plans symbolisant les procédures de compréhension de langage écrit entendu par un enfant, ceux qui méritent une attention toute particulière dans le contexte de sa classe. L'année suivante, Marie-Hélène Le Berre reprend une problématique très proche sur deux albums différents dans un mémoire intitulé : *En quoi des situations d'apprentissage visant le rappel de récit*

peuvent être un tremplin pour des apprentissages langagiers en TPS-PS (très petite section – petite section) ? Le corpus comportant lui aussi des transcriptions de rappel de récit en classe et en aide individualisée, il devient possible d'envisager une étude longitudinale.

Dans d'autres cas, la réflexion se porte sur la transposition d'un outil conçu pour un niveau donné dans un autre contexte. Ingrid Duplaquet fait par exemple le choix de travailler le lexique avant une production écrite, en s'appuyant sur les principes énoncés par Jacqueline Picoche dans le *Dictionnaire du français usuel* (2002). Elle demande à ses élèves de 3^e DP⁽⁸⁾ de retenir quatre mots caractérisant la nouvelle à partir de laquelle ils veulent écrire et leur propose d'abord une recherche dont les entrées sont lexicologiques (définition, verbes associés, synonymes et antonymes, expressions). L'analyse des premiers jets lui montre que les termes recherchés, sur lesquels il y a eu pourtant des débats manifestant une réflexion métalinguistique, sont peu repris dans les textes. Elle se tourne alors vers un dispositif expérimenté à l'école primaire, la corolle lexicale, qui semble répondre au problème qu'elle a identifié : « la construction d'une telle corolle par les élèves installe des compétences métalinguistiques dans la mesure où ils élaborent eux-mêmes les critères de regroupement des mots collectés autour du mot-thème » (Porcar : 111, 2000). Le corpus réunissant les tableaux à entrée lexicologique, les corolles lexicales et les trois versions du texte écrit par les élèves permet de mesurer les effets de chaque dispositif et la conscience que les élèves en ont (puisqu'ils soulignent eux-mêmes les mots et expressions de la corolle qu'ils ont repris dans leur texte).

Dans cette recherche d'outils pour l'analyse de corpus, des thèses dont l'orientation méthodologique correspond à celles présentées plus haut sont un atout. Pour ne citer qu'un exemple, on évoquera, dans celle de Bernadette Lafourcade, les tableaux des productions lycéennes et étudiantes qui permettent de confronter pour tous les scripteurs la forme et le contenu. Les entrées de ces tableaux s'appuient sur les choix théoriques développés dans la thèse : genèse du texte pour les tableaux portant sur la forme, caractères génériques pour les tableaux portant sur le contenu. Ces mêmes outils peuvent être réutilisés dans un écrit professionnel : appliqués à des textes du même genre ou de genres différents, dans des dispositifs d'écriture comparables, ils rendraient possibles des comparaisons sur un ensemble plus vaste de données.

Une modalité de formation s'est révélée propre à faire comprendre la solidarité méthodologique entre une problématique en didactique du français et la constitution d'un corpus : la lecture d'articles de recherche, en remontant des conclusions aux principes de constitution de corpus permettant de établir ces dernières. L'article dirigé par Claudine Garcia-Debanc (2009) a été travaillé dans cet esprit. Il ne présente en effet qu'un texte d'élève, jugé représentatif des productions écrites obtenues après un travail lexical sur les verbes de déplacement, les contraintes de publication ne permettant guère plus d'une illustration. Cet exemple atteste de la faisabilité de la séquence et de son effet sur l'amélioration de la production écrite, et fonctionne comme une ouverture sur une expérimentation

(8) 3^e découverte professionnelle : classe implantée dans un lycée professionnel, accueillant des élèves en échec scolaire en 4^e et qui disposent d'un an pour appréhender la voie professionnelle et choisir une orientation.

qui reste à faire : « Mais la validation de la pertinence de cette ingénierie pour les apprentissages visés [...] suppose une mise à l'épreuve dans des classes nombreuses et diverses, d'une part en vue d'observer les éventuels heurts ou contradictions entre les activités ainsi proposées et les pratiques ordinaires des enseignants, d'autre part en vue d'évaluer les effets sur les apprentissages visés à court, moyen et long terme. » (*op. cit.* p. 210) C'est précisément ce travail qui peut être mené dans le cadre d'un écrit professionnel, adossé une réflexion sur les fondements de cette ingénierie didactique et les caractéristiques du corpus susceptible de l'éprouver : quels moyens se donner pour observer les travaux de groupe sur le classement des verbes et comprendre leur logique sous-jacente ? pour analyser la mise en commun ? comment transcrire les différentes productions d'élèves et quels critères élaborer pour évaluer les effets des apports lexicaux sur le fonctionnement des textes ? Autant de questions qui méritent d'être posées pour garantir la solidité des résultats de l'expérimentation.

Pour passer d'études qualitatives, à visée méthodologique, à des études plus quantitatives (sans atteindre d'emblée la représentativité statistique), la mise en réseau de travaux de recherche et d'écrits professionnels produits par des enseignants expérimentés ou des formateurs pourrait lever une difficulté signalée par Soledad Valera-Kummer et Caroline Masseron (2009 : 89) : celle qui réside « dans la généralisation du propos et dans son application à d'autres analyses comparables, conduites sur de nouveaux textes ».

2.3. Des corpus qui renvoient des questions à la recherche

Retravailler des outils issus de la recherche pour analyser sa propre pratique exige un processus d'élaboration théorique qui ne saurait se réduire à la simple application. Mais il y a plus. De ce processus peuvent naître de nouvelles questions posées à la recherche. J'en donnerai deux exemples.

Dans un mémoire intitulé *Faire évoluer des productions d'écrit en groupe de suivi de module FLS (français langue seconde)*, Catherine Bosredon s'empare des travaux de Jean-Charles Chabanne et Dominique Bucheton (2002) pour donner un étayage théorique à sa pratique d'enseignante convaincue de l'éducabilité des élèves, des pouvoirs du langage pour « penser, apprendre et se construire » (*op. cit.*, 2002), et refusant par conséquent de faire de la maîtrise de la langue un préalable à des activités d'écriture complexes. Si son écrit professionnel se caractérise comme l'aboutissement d'une démarche réflexive bien antérieure, il marque aussi un tournant avec l'invention d'un concept dont l'importance est rappelée en conclusion : « [...] c'est l'adoption de ce rôle de *lecteur intermédiaire* qui, à proprement parler, a modifié mon positionnement. Cela m'a permis de faire des choix plus précis, plus différenciés pour aider les élèves à faire évoluer leurs textes. » Ce concept s'articule à celui de *texte intermédiaire*, emprunté aux auteurs cités, dont l'étudiante a au départ une conception restrictive, minorant leur dimension réflexive, et qui évolue à travers les différentes versions du mémoire, à mesure que s'explicitent la posture qu'elle adopte vis-à-vis de ces textes intermédiaires : « Il s'agit ici de montrer qu'on peut se situer par rapport à un état provisoire de l'écrit, ce qui dédramatise ; et qu'ensuite l'enseignant va se placer en position d'intercesseur pour aider l'élève à tenir compte de son lecteur, à avoir plus d'exigences et donc pour cela à éprouver le besoin d'une meilleure

maîtrise de la langue ; mais pour tenter d'arriver à cela, il faut absolument aider l'effort de l'élève donc cibler des passages, des points précis, lui poser des questions précises. Il faut en fait lui proposer le questionnement et la planification qu'il n'est pas encore en mesure d'opérer lui-même. »

Dans un mémoire intitulé *La mise en voix aide-t-elle à une compréhension des connecteurs grammaticaux à l'école primaire*, Catherine Bacos analyse un débat survenu à l'issue de la mise en voix d'un conte parodique dans une classe de CM1. Ce conte, une réécriture moderne du *Petit chaperon rouge*, a été écrit par des élèves de 5^e à l'intention de cette classe. Or un élève signale son incompréhension d'une phrase comportant un connecteur en segmentant la phrase « La petite racaille rouge prend l'escalier A tandis que le loup prend l'escalier Z » de la façon suivante : « La petite racaille rouge prend l'escalier attendit que le loup prend l'escalier Z ». Il recherche un mot connu dans la chaîne sonore, à un temps prototypique du conte, *attendit*, là où d'autres élèves ont associé *tandis* et *que*. Les explications que proposent alors ces derniers recourent à différents canaux : non verbaux (le geste des mains qui se rencontrent) et verbaux : ils s'appuient sur l'interprétation du conte (les intentions du loup, l'itinéraire parallèle des deux personnages) pour proposer des équivalences avec *alors que* et *pendant que* qui sont ensuite nuancées. Une formulation, en apparence maladroite, « ça veut dire qu'ils vont prendre pas le même passage », montre qu'un élève nouvellement arrivé en France perçoit la congruence de *tandis que*, connecteur dont Guimier montre qu'il présente les deux facettes complémentaires d'une même question à travers deux propositions qui entrent en contraste mais sont orientées vers le même but.

Ce mémoire a immédiatement été convoqué à la lecture de l'article de Soledad Valéra-Krummer et Caroline Masseron (2009) : « Stratégie scripturale et activité conceptuelle : analyse de quelques indicateurs langagières dans des écrits scientifiques de collégiens à visée comparative ». Les auteurs y comparent des écrits dans lesquels des collégiens doivent comparer les virus et les bactéries et observent des choix de planifications et des stratégies scripturales très différentes opposant deux groupes d'élèves, les uns se destinant à des études longues, et les autres à la voie professionnelle : « Les productions les plus réussies au moins les plus efficaces sont celles qui n'hésitent pas à reconduire le même moule syntaxique : ordre des thèmes et répétition de *tandis que*. » Dans les autres productions « le texte manque du *lien cohésif* propre au genre d'écrit pour lequel il est sollicité (exposition méthodique et critériée des savoirs en biologie). La ressource d'un cadre syntaxique, intuitif et importé d'une pratique langagière spontanée, n'est pas suffisante ici, ainsi qu'en témoignent les effets de rupture, gérés au coup par coup » (*op. cit.* : 101). Qu'en est-il pour ces élèves de la compréhension du connecteur dans un texte narratif ? de son emploi dans un récit ? à quelles conditions peut s'effectuer son réemploi dans un écrit scientifique ? Sans apporter une réponse à l'ensemble de ces questions, la connexion des deux corpus permet au moins de les poser.

La mise en relation d'écrits professionnels et d'écrits de recherche dans le cadre d'un grand corpus permettrait à la fois d'accroître les données sur une même problématique et de susciter de nouvelles recherches.

3. Perspectives : vers l'élaboration d'un grand corpus sur/pour la formation

Dans le contexte de la mastérisation de la formation des enseignants, la constitution d'une telle somme de corpus apparaît comme une nécessité. Les étudiants se destinant à l'enseignement ont désormais une expérience professionnelle plus limitée et moins directe, voire inexistante, puisque l'entrée dans le travail de recherche se fait en amont de la pratique. Alors que dans la formation en IUFM, l'alternance a pu être utilisée comme levier pour faire évoluer les représentations des élèves, de l'apprentissage et du métier, dans le dispositif actuel, d'autres moyens doivent être inventés, dans le cadre d'une initiation à la recherche à visée professionnelle. Il faudrait par exemple que les étudiants puissent lire de nombreux textes d'élèves de différents niveaux dans des genres différents avant d'aborder leur stage de pratique accompagnée, ce qui peut être proposé à partir d'une banque d'écrits scolaires réalisée selon une méthodologie explicite (Elalouf, 2005).

On peut faire l'hypothèse que la mise à disposition des formateurs de corpus accompagnés de guides d'utilisation favoriserait, en première année de master, la centration sur les apprentissages des élèves dans un contexte institutionnel donné. De même, on peut penser que des données recueillies dans le cadre de masters de formation de formateurs aideraient des étudiants de master 2, préparant leur stage en responsabilité, à prendre la mesure des marges de manœuvre possibles, en tenant compte des contraintes institutionnelles et de la culture du métier. On peut enfin faire le pari que des étudiants ayant eu travaillé sur différents corpus dans le cadre de leur formation seront plus nombreux à adopter dans leur mémoire professionnel la posture de recherche qu'implique la constitution d'un corpus.

Il n'en reste pas moins que bien des questions demeurent en suspens, déjà posées dans le cadre du colloque *Les didactiques en questions* : « Cette initiation à la recherche peut-elle développer la professionnalité de futurs enseignants, faire évoluer leurs représentations des élèves, du milieu professionnel et de l'acte d'enseigner ? Comment éviter que ne surgissent des “monstres théoriques” » (Elalouf *et al.*, 2011) ?

C'est la raison pour laquelle ce grand corpus pour la formation ne saurait se concevoir sans son complémentaire : un grand corpus sur la formation, qui interrogerait les incidences de la mise en place des masters d'enseignement, dans leur grande diversité, sur la professionnalité des enseignants issus de ces formations. Nous disposons pour la période antérieure de recherches menées dans les ERTE de Toulouse et Montpellier ainsi que dans le laboratoire PAEDI (Carnus *et al.*, 2008, Bucheton 2009, Goigoux *et al.* 2009). Dans le cadre de l'équipe EMA, un suivi de cohorte d'étudiants inscrits dans un master « enseignement des Lettres » s'engage. En se fondant sur les conclusions d'un suivi antérieur (1998-2004) de la préprofessionnalisation à l'entrée dans le métier (Elalouf, 2007), on peut faire l'hypothèse que la façon dont l'enseignant débutant mobilise dans sa pratique des savoirs disciplinaires, didactiques et professionnels est étroitement liée à la façon dont sa formation lui a permis de les recomposer et de les travailler avec les savoirs issus de l'expérience. Un nouveau suivi longitudinal permettrait à la

fois d'accéder à une meilleure compréhension des pratiques des débutants mais aussi d'éclairer la relation entre la conception de la formation et ses effets sur les pratiques.

Peut-être sommes-nous à un moment charnière où, malgré un contexte défavorable, la mise en réseau de recherches, dans le cadre d'un projet commun à plusieurs laboratoires, pourrait faire franchir à la didactique du français un saut qualitatif. Il s'agirait de réunir les conditions pour que les corpus deviennent effectivement cumulables mais aussi d'anticiper sur la façon dont ils pourront être utilisés en recherche, en formation collective ou personnelle, en intégrant les inévitables distorsions qui en résulteront.

Bibliographie

- BRIGAUDIOT, M. (coord.) (2000) : *PROG, Apprentissages progressifs de l'écrit à l'école maternelle*. Paris, Hachette éducation-INRP.
- BUCHETON, D. (dir.) (2009) : *L'agir enseignant, des gestes professionnels ajustés*. Toulouse, Octarès.
- CARNUS, M.-F., GARCIA-DEBANC, C., TERISSE, A. (2008), *Analyse des pratiques des enseignants débutants, approches didactiques*. Grenoble, La pensée sauvage.
- CHABANNE, J.-C., BUCHETON, B. (dir.) (2002) : *Ecrire en ZEP*. Paris, Delagrave.
- ELALOUF, M.-L. (dir.) (2005) : *Écrire entre 10 et 14 ans, un corpus, des analyses, des repères pour la formation*. CRDP de l'académie de Versailles.
— (2007) : *Apprentissages de l'écrit et réflexion métalinguistique, étude qualitative de corpus scolaires et universitaires*. Document de synthèse présenté pour l'obtention de l'Habilitation à Diriger des Recherches, sous la direction de Bernard Combettes, Nancy 2.
- ELALOUF, M.-L., ROBERT, A., BELHADJIN, A. BISHOP, M.-F. (2011) : *Les didactiques en questions : État des lieux et perspectives pour la recherche et la formation*. Bruxelles, De Boeck (à paraître).
- GARCIA-DEBANC, C. (2008) : « De la configuration didactique au modèle disciplinaire en acte ». *Pratiques* n°137-138, *La didactique du français, hommage à Jean-François Halté*, pp. 39-56.
- GARCIA-DEBANC, C., DUVIGNEAU, K., DUTRAIT, C., GAGNEUX, M. (2009) : « Enseignement du lexique et production écrite. Un travail sur les verbes de déplacement à la fin de l'école primaire », *Pratiques* n°141-142, *La synonymie*, pp. 208-232.

- GOIGOUX, R. & VERGNAUD, G (2005) : « Schèmes professionnels », in BERNIÉ J.-P. et GOIGOUX, R. (coord.). *Les gestes professionnels. Dossier de la Lettre de l'AIRDF* n° 36, pp. 7-10.
- GOIGOUX, R., RIA, L. TOCZEK-CAPELLE, M.-C. (2009) : *Les parcours de formation des enseignants débutants*. Clermont-Ferrand, Presses universitaires Blaise Pascal.
- PICOCHÉ, J. & ROLLAND, J.-C. (2002) : *Dictionnaire du français usuel*. Bruxelles : De Boeck.
- PORCAR, M.-H. (2000) : « De l'usage de la corolle lexicale au cycle 3 et en 6^e », *Le français aujourd'hui* n°131, *Construire les compétences lexicales*, pp. 110-116.
- VALERA-KUMMER, S. & MASSERON, C. (2009) : « Stratégie scripturale et activité conceptuelle : analyse de quelques indicateurs langagiers dans des écrits scientifiques de collégiens à visée comparative ». *Pratiques* n° 143-144, *Écrits de savoirs*, pp. 83-110.

questions de communication 18-2010

Les non-usagers des T I C

Dossier coordonné et présenté par Catherine Kellner, Luc Massou et Pierre Morelli

Les non-usagers de l'internet. Axes de recherche passés et futurs

Sally Wyatt

Le « non-usage » de l'internet : reconnaissance, mépris et idéologie

Fabien Granjon

Pourquoi certains n'adoptent-ils pas l'internet ? L'influence de la vie quotidienne et de la culture de résistance en Grèce

Panayiota Tsatsou

Des usages limités des T I C chez des professionnels de l'éducation et du conseil dans le social

Catherine Kellner, Luc Massou, Pierre Morelli

Non-usages du téléphone portable : au-delà d'une position binaire usagers/non-usagers

Corinne Martin, Thilo Von Pape

Interdisciplinarité, interdisciplinarités

Béatrice Fleury, Jacques Walter

Quand l'interdisciplinarité est un « état d'esprit » critique et heuristique

Roselyne Koren

De l'infra-conceptuel des données à horizon de pertinence focalisé

Sylvie Leleu-Merviel

Analyse du discours et champ disciplinaire

Dominique Maingueneau

PRIX AU NUMÉRO 20 euros (frais de port de 3,25 euros en sus)

ABONNEMENT (1 an, 2 numéros) 32 euros (frais de port de 6,50 euros en sus)

Revue publiée avec le soutien du Centre de recherche sur les médiations (Université Paul Verlaine-Metz, Nancy Université, Université de Haute-Alsace) du Conseil Régional de Lorraine et bénéficiant de la reconnaissance du CNRS

Abonnement : Presses universitaires de Nancy • pun@univ-nancy2.fr