



HAL
open science

Interroger les implicites de l'évaluation des résultats des élèves lors d'une expérimentation didactique

Bertrand Daunay

► To cite this version:

Bertrand Daunay. Interroger les implicites de l'évaluation des résultats des élèves lors d'une expérimentation didactique. Cora Cohen-Azria, Nathalie Sayac. Questionner l'implicite. Les méthodes de recherche en didactiques (3), Septentrion, pp.201-210, 2009, 9782757401224. hal-01354226

HAL Id: hal-01354226

<https://hal.science/hal-01354226>

Submitted on 18 Aug 2016

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

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

Interroger les implicites de l'évaluation des résultats des élèves lors d'une expérimentation didactique

Bertrand Daunay
Université Charles de Gaulle
ÉA Théodile – Lille 3

Introduction

Je me propose de revenir, pour en interroger quelques implicites méthodologiques, sur une expérimentation menée entre 2002 et 2004, qui portait sur l'écriture d'invention, introduite peu de temps auparavant dans les programmes de lycée en France.

À l'issue de cette expérimentation et des publications de ses résultats, j'ai fait plusieurs retours critiques, de nature épistémologique (s'agissant de la manière dont le savoir visé par la séquence est construit par ses concepteurs) ou méthodologique (à propos des modes d'analyse des performances des élèves). Ce souci du retour critique ne tenait pas à la prise de conscience ultérieure d'erreurs qui auraient été commises (de ce point de vue, notre travail ne semble pas déparer les recherches de même nature publiées dans notre champ), mais voulait exhiber les questions épistémologiques et méthodologiques que posait notre démarche, dans un double but :

– conforter la possibilité d'une discussion scientifique de nos résultats en montrant à *quelles conditions* ils sont recevables et donc capitalisables dans la communauté de recherche qui les a accueillis ;

– contribuer à la discussion méthodologique en didactique, auquel concourt particulièrement le séminaire international annuel sur les méthodes de recherche en didactique, pour participer là encore à une possible capitalisation des méthodes propres à un champ disciplinaire.

Cette exigence de discussion et de capitalisation des méthodes comme des résultats de recherche me semble devoir être confortée actuellement en didactique, notamment en didactique du français.

Je me propose ici, après avoir présenté rapidement dans ses grandes lignes le cadre de l'expérimentation, de m'arrêter sur trois questions cruciales qui la concernent – et qui concernent aussi, dans mon esprit, toute expérimentation en didactique – à propos de trois aspects de la mesure des résultats des élèves dans le cadre du protocole expérimental : quel sujet cerne-t-on ? Dans quelle temporalité ? Pour quel savoir ?

1. Le cadre de l'expérimentation

Sans entrer dans le détail, il suffit de rappeler les grandes lignes du cadre méthodologique : l'expérimentation a concerné 39 classes et 960 élèves (de 3^e et de 2nde, soit des élèves de 14 à 15 ans). Notre hypothèse était qu'écrire en début de séquence et exploiter l'écrit produit facilitait l'acquisition de savoirs et de savoir-faire ; la variable de notre expérimentation était donc la *place* (au sens de *moment* et de *rôle*) de l'écriture dans le processus d'apprentissage.

L'introduction dans les derniers programmes de français du lycée en France de *l'écriture*

d'invention était l'occasion de questionner une conception didactique de l'écriture comme outil *et* objet d'apprentissage, considérée non comme le couronnement (non appris) d'autres apprentissages (*cf.* Halté, 1992, p. 61 *sqq.*) mais comme élément central des apprentissages et comme moyen de construire des savoirs et des savoir-faire, dans le domaine de la lecture comme de l'apprentissage de notions.

Le protocole expérimental envisageait des classes expérimentales proprement dites, qui réalisaient la séquence expérimentale, avec une écriture d'invention en ouverture et en clôture (classes A, au nombre de 17) et des classes témoins, qui ne commençaient pas la séquence par une écriture d'invention ; mais nous avons distingué, parmi ces classes témoins, des classes qui terminaient la séquence par une écriture d'invention (classes B, au nombre de 14) et des classes où n'étaient réalisées aucune écriture d'invention (classes C, au nombre de 8). Le reste de la séquence (l'étude d'un corpus de textes) était identique pour les trois types de classes. Les trois types de classes subissaient le même test d'entrée et de sortie, destiné à mesurer leur maîtrise du savoir visé par la séquence : le discours indirect libre.

En résumé :

Classes A	Classes B	Classes C
Test d'entrée		
Écriture d'invention (1) Étude d'un corpus de textes Écriture d'invention (2)	Étude d'un corpus de textes Écriture d'invention	Étude d'un corpus de textes
Test de sortie		

Nous avons décidé de viser un savoir comme objet d'apprentissage – et les compétences de lecture ou d'écriture afférentes à cette notion. Le choix du discours indirect libre sera discuté plus loin, mais il reste à préciser les domaines d'apprentissage visés par la séquence didactique et, en partie, par les tests d'entrée et de sortie :

	Réception	Production
Approche spontanée	Perception de la polyphonie dans texte comportant du discours indirect libre	Utilisation intuitive du discours indirect libre pour faire entendre une autre voix que celle du locuteur
Approche consciente	Repérage et identification du discours indirect libre, étude des effets produits par le discours indirect libre	Utilisation consciente du discours indirect libre pour produire des effets

Pour faciliter l'appropriation de la séquence par les enseignants, nous leur avons laissé une grande liberté : c'est que nous voulions que cette séquence soit réalisée dans les classes au plus près des habitudes d'enseignement des professeurs sollicités. Cette exigence « écologique » engage, évidemment, de nombreux biais possibles dans la démarche d'expérimentation, mais c'est là un aspect normal de toute expérimentation en didactique (*cf.* Reuter *et al.*, 1994, p. 16). Ce que nous voulions évaluer (les effets de la place de l'écriture) ne devait en outre pas être connu des

enseignants, pour qu'ils agissent à l'égard de cette séquence comme avec celles qu'ils conçoivent eux-mêmes pour leurs classe.

Un simple mot sur les résultats : si l'on prend les résultats globaux, les effets de la séquence d'apprentissage sont plus visibles dans les classes expérimentales, ce qui va dans le sens de notre hypothèse. Si l'on entre dans le détail, l'interprétation est plus complexe : concernant l'utilisation spontanée du discours indirect libre en production, les effets de l'entrée par l'écriture ne sont pas significatifs en classe de 3^e, mais le taux de réussite est significativement supérieure dans les classes expérimentales en 2nde. Pour ce qui est de la perception intuitive du discours indirect libre, les effets de l'entrée par l'écriture semblent importants en classe de 3^e, mais pas en classe de 2nde. S'agissant enfin de la reconnaissance consciente du discours indirect libre, il n'y a pas de différence nette selon les classes. Cette spécification des résultats interroge d'ailleurs un de nos *a priori* selon lequel les effets d'une écriture en début de séquence seraient de même nature en collège et en lycée.

2. Qu'est-ce qu'un « bon résultat » en didactique ?

Nos choix méthodologiques découlent d'un positionnement épistémologique qui s'inscrit dans un débat alimenté régulièrement au sein de la didactique du français, dont on trouve les échos dans l'ouvrage collectif édité par M. Marquilló Larruy (2001). Admettant l'impossibilité de la maîtrise totale d'un protocole expérimental pour les questions éducatives (Van der Maren, 1996/2004, p. 33 *sq.* ; 1999/2003, p. 53 *sqq.*), nous avons opté pour une *quasi-expérimentation*, selon une distinction que reprennent la plupart des traités méthodologiques (y compris Gagné, Lazure, Sprenger-Charolles, Ropé, 1989, p. 51). Cette « quasi-expérimentation » ne prétend pas reproduire les conditions d'un travail de laboratoire, mais respecte la dimension « écologique » qui nous semble devoir être celle de toute approche expérimentale en didactique. Cela n'enlève rien à la légitimité épistémologique d'une telle approche, nécessaire à la construction de connaissances en didactique, comme le dit Jean-Louis Dufays (2001, p. 20) :

Même si les recherches quasi-expérimentales ne peuvent donner qu'une vision restreinte et schématique des réalités qu'elles cherchent à saisir, leur méthode s'avère précieuse pour objectiver un tant soit peu des modélisations ou des propositions qui, sans elle, risquent de rester confinées dans le flou de la subjectivité.

Ce choix de mettre en œuvre un protocole quasi-expérimental dans un contexte « écologique » nous a amenés à renoncer à maîtriser un certain nombre de paramètres, qui deviennent de fait des variables latentes, sans être précisément interrogées.

Mais, de surcroît, ayant fait au départ le choix d'une approche quantitative courante pour poser comme « bon résultat » le progrès d'un plus grand nombre d'élèves de telle ou telle catégorie de classes (A, B ou C), nous avons logiquement renoncé à considérer comme variables certaines caractéristiques des élèves (sexe, retards scolaires, professions et catégories socio-professionnelles, etc.), mais aussi des enseignants, des établissements, voire des choix d'organisation de la séquence didactique.

Ce choix est bien sûr volontaire et s'explique par le fait que la variable que nous introduisons (la *place* de l'écriture dans le processus d'apprentissage) était de nature didactique (en ce qu'elle concernait un mode de conception de la construction du savoir dans la classe) et méritait d'être traitée pour elle seule, sans autre considération – qui aurait en revanche été pertinente dans d'autres champs de recherche.

Notons au passage que présidaient aussi à ce choix des conditions de faisabilité, de deux

natures : le souci de ne pas trop solliciter les enseignants volontaires par des demandes de renseignements non strictement nécessaires et la nécessité de ne pas multiplier les variables croisées, afin de ne pas multiplier le nombre de catégories et de réduire ainsi l'effectif de chacune, au risque de perdre tout le bénéfice que l'on pouvait retirer du grand nombre d'élèves concernés au départ.

Pour autant, on doit interroger les postulats méthodologiques et épistémologiques, voire déontologiques qui ont présidé à ce choix. Pour emprunter les mots de Dominique Lahanier-Reuter lors de la présentation que nous avons faite de notre recherche devant l'équipe Théodile : « Qu'est-ce qu'un "bon" résultat (statistique) en didactique ? » L'élève moyen (celui que donne à voir notre traitement statistique des données, sans prise en compte du niveau des élèves) est-il le plus intéressant à cerner ? On pourrait aussi bien supposer qu'en didactique, l'élève moyen (statistiquement) n'est pas intéressant : un « bon » résultat serait alors celui qui montrerait que telle variable a des effets particulièrement significatifs sur les seuls élèves en difficulté (ce qui serait assez en résonance avec la vocation initialement affichée de la didactique – du français particulièrement – de lutter contre l'échec scolaire) ou sur les seuls meilleurs élèves...

Il se trouve que, par un réexamen ultérieur des données, nous avons pu montrer que les résultats ne sont plus interprétables de la même manière si l'on prend en compte l'appartenance des élèves à des classes « faibles » ou non : par exemple, alors que dans nos résultats globaux, la séquence expérimentale montrait, sur plusieurs plans, une plus grande réussite des élèves aux tests de sortie, c'est au contraire la séquence témoin (plus « traditionnelle ») qui aurait plus de réussite auprès des élèves des classes faibles. Qu'en serait-il de nos résultats et de leur assez grande conformité avec nos hypothèses, si l'on prenait en compte le niveau de chaque élève ?

Bien sûr, on n'est pas obligé de chercher un « bon résultat » et l'on peut donner des résultats diversifiés en prenant en compte les variables que l'on voudra, sans décider quel est le « bon » : mais outre que l'on rencontre à nouveau le problème de la faisabilité même d'une telle démarche d'analyse, outre encore qu'on n'épuisera jamais les catégories potentiellement intéressantes susceptibles d'être construites dans l'analyse des résultats, on voit bien que l'on rencontre au passage un problème épistémologique d'ampleur : en effet, en s'attachant à d'autres variables que strictement didactiques (autrement dit déterminées par la théorie didactique), on en arrive à la conclusion qu'une variable spécifiquement didactique (c'était précisément le cas de ce que nous visions ici : la place de l'écriture dans une séquence) ne peut faire l'objet d'une expérimentation indépendamment d'une autre...

C'était d'ailleurs cette intuition qui nous a amenés au choix d'interroger les résultats moyens des élèves : choix conscient et raisonné, donc, mais qui laissait dans l'implicite la question de l'élève que l'on cherchait à cerner.

3. Question de temporalité

Ma deuxième question n'est pas seulement un clin d'œil au précédent séminaire sur les méthodes de recherche en didactique (Lahanier-Reuter, Roditi *dir.*, 2007), mais une vraie question liée aux choix méthodologiques et épistémologiques privilégiés d'une discipline de recherche. Si l'on affirmait sérieusement que la séquence didactique est l'espace temporel pertinent pour évaluer les effets d'un dispositif didactique, on serait sans doute désavoué par beaucoup.

C'est pourtant ce qui implicitement ressort de l'organisation de notre protocole expérimental (au demeurant très classique et très répandu), quand nous isolons un certain nombre de

compétences pour les évaluer par des tests d'entrée et de sortie ; même si l'on reconnaît la validité de l'évaluation, que dit-elle du devenir de ces compétences ? Or quelle pertinence y aurait-il eu, pour répondre à cette question, à faire repasser des tests (les mêmes ou d'autres) à un ou deux mois ou années de distance ? Qu'aurions-nous mesuré alors : le même savoir ou un autre ? Et quel sérieux aurait l'affirmation selon laquelle la variable de notre expérimentation était encore mesurable à une telle distance ? Bref, la compétence mesurée risque de n'être que celle construite au cours de la séquence.

Il n'y a là rien d'aberrant, dans la mesure où ce n'est pas un « mode de travail didactique » qui était évalué – ce qui effectivement aurait nécessité une étude longitudinale et une comparaison à plus grande échelle : c'était un dispositif didactique spécifique, propre à la temporalité d'une séquence, et qui pouvait être évalué à l'issue d'une séquence, *mais* seulement à ce moment-là. Et c'est ce *mais* qui restait dans l'implicite quand on prétendait évaluer les compétences construites, en réception comme en production de textes (« conscientes » ou non), dans la maîtrise du discours indirect libre.

On voit bien comment une expérimentation de cette nature laisse dans l'implicite la temporalité de l'évaluation et ce que cet implicite charrie de présupposés non interrogés : si nous avons pu apporter des connaissances en didactique sur l'effet de l'écriture en début de séquence, c'est au prix de la négligence de la question du devenir de ces effets, au prix aussi de la naturalisation de l'idée que la séquence serait une temporalité didactique pertinente.

Plus généralement, cette question de la temporalité dans la mesure des apprentissages des élèves rencontre celle de la durée et de la nature du système didactique : on voit bien en effet les particularités de l'analyse d'un didacticien par rapport à celles d'autres chercheurs, dans la mesure où la prise en compte du système didactique l'amène à interroger à la fois la dimension collective (la classe) et temporelle (l'évolution du système). Élisabeth Nonnon (2005, p. 173) posait cette question à propos de l'oral, en observant la différence d'approche des acquisitionnistes, qui suivent un petit nombre de sujets et peuvent le faire sur un temps long, et des didacticiens, dont la tâche est rendue plus ardue par le fait de devoir suivre un ensemble d'élèves sur une durée qui ne peut, matériellement, être trop longue. Élisabeth Nonnon posait d'ailleurs à ce sujet le problème des « tensions liées aux contraintes de la temporalité » : l'oral particulièrement (en raison de l'inscription de toute réalisation orale dans le temps) demande que son évaluation soit inscrite dans une « temporalité » significative. Mais le problème est généralisable à tout apprentissage : c'est une autre question pertinente de se demander quelle est cette significativité : à quel moment un apprentissage est censé être évaluable ?

Du reste, cette question de la temporalité en croise une autre, celle du lien, très discutable (dans tous les sens du mot), entre l'évaluation d'une performance et celle d'une compétence et la validité accordée aux outils d'évaluation des performances et/ou des compétences des élèves, question particulièrement centrale dans l'interprétation des données verbales. Sur ce point, rien ne garantit toujours qu'une recherche se donne les moyens de s'assurer que l'analyse des performances des élèves dans le cadre d'une expérimentation ne relève pas d'intuitions issues de la pratique scolaire ordinaire.

4. Le savoir construit et évalué

Quelle est l'influence de l'objet étudié sur les performances des élèves et la validité de notre mesure de l'efficacité du dispositif didactique expérimental ? Nous avons décidé de choisir

comme objet d'apprentissage un savoir et non des compétences de lecture ou d'écriture. Nous n'avions pas choisi au hasard le savoir en jeu dans la séquence : nous nous étions même donné des critères très explicites pour le choisir, critères qui relevaient à la fois des contraintes propres à notre expérimentation et de notre conception déclarée de l'écriture d'invention, objet central de notre recherche. Il fallait selon nous que ce savoir :

- soit accessible au collège et au lycée ;
- ne relève pas de questions souvent travaillées en collège, ce qui introduirait le paramètre difficilement contrôlable des connaissances antérieures des élèves ;
- ne soit pas trop vaste (un genre littéraire, par exemple) ;
- soit assez facile à identifier et ne mette pas en jeu trop de composantes, trop de niveaux d'analyse (afin de pouvoir identifier plus facilement ce qui pose problème aux élèves et ce qui se construit) ;
- engage la question des valeurs ;
- soit intéressante pour la lecture comme pour l'écriture.

Le discours indirect libre nous a semblé répondre à ces critères : objet technique (qui à ce titre pourrait être traité d'un point de vue strictement linguistique), il touche cependant aux valeurs en ce qu'il questionne la propriété du discours et le rapport au langage de l'autre. Malgré son côté en apparence très « pointu », le discours indirect libre est, en collège comme en lycée, un point d'entrée dans des questions qui n'ont rien d'anecdotique.

Si donc le choix de ce savoir a été raisonné et se justifiait à plusieurs égards, sa légitimité est cependant discutable dans la mesure où certaines de ses caractéristiques, laissées dans l'implicite, peuvent se révéler non neutres dans ce que nous cherchions à mesurer et l'on peut se demander si ce choix ne donnait pas *a priori* un « avantage » important au dispositif didactique expérimental par rapport au dispositif mis en place dans les classes témoins, pour au moins deux raisons. D'une part, le savoir en question peut être l'objet d'une *pratique*, autrement dit être autant *procédural* que *déclaratif* – et cela seul peut entrer en résonance avec la variable introduite, la place de l'écriture d'invention ; en effet, on peut se demander si les résultats n'auraient pas été franchement différents s'il s'était agi d'un savoir plus nettement déclaratif, ne renvoyant pas *directement* à un savoir-faire. D'autre part, le discours indirect libre est à beaucoup d'égards indécidable, qu'il s'agisse de la définition de la notion ou de la reconnaissance de l'objet (en réception comme en production) ; et cela n'est pas sans lien avec la logique de construction du cours à partir des réactions des élèves et des discussions qu'elles engageaient.

Autrement dit, l'objet, implicitement, se révélait assez pertinent pour la séquence expérimentale, qui consistait à faire écrire les élèves dès le début et à les faire revenir sur leurs productions. Ce ne serait pas le cas pour un objet plus construit scolairement, dont le caractère normé ne fait pas de doute (dans le monde scolaire ou dans la théorie). Certes, dans le domaine de la didactique du français, il n'est pas sûr qu'aucun objet ait une absolue stabilité, mais il est au moins des objets que l'école a construits comme stables, ce qui n'est pas le cas du discours indirect libre : dès lors, outre le fait que nous introduisions là un biais dans notre expérimentation, nous nous interdisions quasiment la possibilité d'un transfert de notre situation expérimentale à d'autres situations didactiques. Finalement, si tant est que nos résultats soient acceptables, pourrait-on en conclure autre chose que l'efficacité de telle séquence sur le seul apprentissage du discours indirect libre ?

La question du transfert d'une expérimentation, qui rejoint celle de sa reproductibilité comme de la généralisation de ses résultats, ne concerne pas seulement notre expérimentation,

évidemment : mais il est intéressant de voir comment des chercheurs, ayant pourtant défini des critères assez stricts pour le choix de l'objet d'apprentissage visé par la séquence, ont finalement laissé dans l'implicite certains traits du savoir en jeu, qui l'auraient disqualifié sans doute s'ils avaient été explicités avant la mise en œuvre du protocole expérimental.

Conclusion

Les trois questions posées ici sur la mesure des résultats des élèves dans le cadre d'un protocole expérimental interrogent en réalité un implicite majeur : la valeur accordée *a priori* à l'approche expérimentale ou quasi-expérimentale en didactique... Il me semble que les trois questions posées ci-dessus concourent à la même conclusion : l'intérêt d'une approche (quasi)expérimentale en termes de résultats de recherche serait finalement exactement proportionnel au degré de réduction de l'objet investigué ; plus on maîtrise ce que l'on mesure, moins on donnerait d'informations éclairant la complexité des phénomènes que vise à comprendre la recherche didactique.

Quand on interroge en effet quel est l'élève que l'on évalue, quelle temporalité est concernée par l'évaluation, quelle est la nature des objets mêmes de l'évaluation, nombreuses sont les questions qui apparaissent et qui interrogent l'empan de validité d'une telle expérimentation. Le cadre théorique et méthodologique préalablement défini légitime certes l'approche et il ne s'agit pas de remettre en cause la démarche entreprise, même si elle a fait ici l'objet d'un questionnement serré : mais les implicites qui demeurent, dont rien ne saurait permettre de dire qu'ils ne sont pas finalement consubstantiels à toute démarche de cette nature, illustrent une fois de plus l'extrême complexité d'un traitement didactique des phénomènes d'apprentissage.

Références

- BRASSART Dominique Guy, 1995/2005, « Didactique du français langue maternelle : approche(s) "cognitivist(e)s" »?, dans CHISS Jean-Louis, DAVID Jacques, REUTER Yves *dir.*, *Didactique du français. Fondements d'une discipline*, Paris, Nathan, p. 95-118.
- CAUTERMAN Marie-Michèle, DAUNAY Bertrand, COGET Clémence, DENIZOT Nathalie, VANDERKELEN Brigitte, 2004, *Construction d'un « objet didactique » : l'écriture d'invention au collège et au lycée*, rapport de recherche R/RIU/03/04, IUFM Nord – Pas-de-Calais.
- DAUNAY Bertrand, 2004a, « Écriture d'invention et apprentissage du discours indirect libre », *Pratiques* n° 123-124, *Dialogisme et Polyphonie*, Metz, CRESEF, p. 213-248.
- DAUNAY Bertrand, 2004b, « Traces d'apprentissage : que reste-t-il d'une séquence didactique ? », *Recherches* n° 41, *Traces*, 2004-2, p. 141-170 ;
- DAUNAY Bertrand, 2006a, « Écrire d'abord. Étude comparative selon les milieux d'enseignement », dans LEBRUN Marlène *éd.* (2006) *Littérature et pratiques d'enseignement-apprentissage : difficultés et résistances* [Actes du colloque international à l'IUFM d'Aix-Marseille, 22-23 octobre 2005, édités en ligne : <http://www.fse.ulaval.ca/litactcolaix/>]
- DAUNAY Bertrand, 2006b, « Écrire d'abord. L'expérimentation d'un principe didactique », dans FALARDEAU Érick, FISHER Carole, SIMARD Claude, SORIN Noëlle *éd.*, *Le français : discipline plurielle ou transversale ?*, Québec, Presses de l'Université Laval, p. 185-202.

- DAUNAY Bertrand, 2006c, « Évaluer les traces des apprentissages d'élèves dans le cadre d'une expérimentation », dans LECLERCQ Véronique, REUTER Yves coord., *Cahiers d'études du CUEEP n° 57, L'Évaluation : regards croisés en didactiques*, Lille, USTL-CUEEP, p. 7181.
- DAUNAY Bertrand, 2008, « Performances et apprentissages disciplinaires », *Cahiers Théodile n° 8*, Villeneuve d'Ascq, Université Charles de Gaulle – Lille 3, p. 7-23.
- DUFAYS Jean-Louis (2001), « Quelle(s) méthodologie(s) pour les recherches en didactique de la littérature ? Esquisse de typologie et réflexions exploratoires », *Enjeux n° 51-52*, juin/décembre 2001, *Recherches en didactique de la littérature*, Namur, CEDOCEF.
- GAGNÉ Gilles, LAZURE Roger, SPRENGER-CHAROLLES Liliane, ROPÉ Françoise, 1989, *Recherches en didactique et acquisition du français langue maternelle*, Tome 1, Cadre conceptuel, thésaurus et lexique des mots-clés, Bruxelles, De Boeck-Wesmael ; Paris, Éditions Universitaire-INRP ; Montréal, Université de Montréal-PPMF.
- GENETTE Gérard, 1969, *Figures II*, Paris, Seuil, collection « Points ».
- GENETTE Gérard, 1982, *Palimpsestes*, Paris, Seuil.
- HALTÉ Jean-François, 1992, *La didactique du français*, PUF, coll. « Que-sais-je ? ».
- LAHANIER-REUTER Dominique, RODITI Éric dir., 2007, *Questions de temporalité. Les méthodes de recherche en didactiques (2)*, Villeneuve d'Ascq, Presses universitaires du Septentrion.
- MARQUILLÓ LARRUY Martine éd., 2001, *Questions d'épistémologie en didactique du français (langue maternelle, langue seconde, langue étrangère)*, AIRDF-Université de Poitiers.
- NONNON Élisabeth, 2005, « Entre description et prescription, l'institution de l'objet : qu'évalue-t-on quand on évalue l'oral ? », *Repères n° 31, L'évaluation en didactique du français : résurgence d'une problématique*, p. 161-188.
- REUTER Yves éd., 1994, *Les interactions lecture / écriture*, Peter Lang.
- REUTER Yves éd., 2007b, *Une école Freinet. Fonctionnements et effets d'une pédagogie alternative en milieu populaire*, Paris, L'Harmattan.
- VAN DER MAREN Jean-Marie, 1996/2004, *Méthodes de recherche pour l'éducation*, Bruxelles, De Boeck.
- VAN DER MAREN Jean-Marie, 1999/2003, *La Recherche appliquée en pédagogie : des modèles pour l'enseignement*, Bruxelles, De Boeck.