Evaluation, changement, représentation
Vers une interrogation de quelques allants de soi en didactique des langues

Emmanuelle HUVER, Université F. Rabelais (Tours), EA 4246 DYNADIV (Tours/Limoges)

Evoquer de manière exhaustive, ou, tout du moins panoramique, les questions relatives à la recherche sur l'évaluation en didactique des langues / cultures (DLC) ne va pas sans poser quelques difficultés dans les limites de taille imposées par un article, si l'on souhaite éviter à la fois la superficialité de l'analyse et son émiettement. Dans la mesure où cet « exercice » suppose de mener des réflexions plus finement étayées, qui viennent par ailleurs d'être réalisées (HUVER et SPRINGER 2011), j'ai plutôt choisi de proposer une lecture interprétative de l'argumentaire de ce numéro. J'en retiendrai plus particulièrement deux axes : d'une part, la problématique de la persistance et/ou du changement des pratiques évaluatives, qui pose de manière plus générale la question de l'innovation des pratiques didactiques, et d'autre part, la problématique de l'intégration des représentations aux formations et à la recherche sur l'évaluation, dans la mesure où cet aspect demeure encore trop marginal.

1. Pratiques d'évaluation et changement
1.1. Des changements dans le paysage de l'évaluation ?
L. Porcher, dans l'extrait proposé dans l'argumentaire de ce volume, souligne essentiellement les phénomènes de changement survenus dans le domaine de l'évaluation en DLC. On ne peut en effet nier que des évolutions de fond (sociétales, politiques, etc.) ont engendré des modifications profondes du paysage de l'évaluation, ne serait-ce que par la forte montée en puissance de celle-ci et de ce qu'il est désormais convenu de nommer une « culture de l'évaluation » (ce que D. Martuccelli a très récemment nommé une « philosophie de l'évaluation »), qui touche de manière protéiforme des domaines très divers de la société : université, santé publique, monde de l'entreprise, etc. Le champ éducatif, et donc la DLC, n'échappe pas à cette tendance de fond, puisqu'on observe depuis 2001 un développement particulièrement fort des certifications[1], ce qui se répercute entre autres sur les contenus des manuels, qui comprennent désormais quasiment tous une page d'entrainement aux certifications (au DELF notamment, dans le domaine du FLE).
Toutefois, peut-on pour autant dire que cette montée en puissance est vectrice de véritables changements ? Les épistémologies qui fondent les pratiques actuellement dominantes de certification relèvent essentiellement de la psychométrie, et, plus largement, d'un paradigme objectiviste et positiviste. Au niveau méthodologique, ce parti-pris épistémologique a abouti à une centration sur les descripteurs de la compétence langagière, notamment sur les dimensions linguistiques et pragmatiques de celle-ci, au détriment des dimensions socioculturelles, interculturelles, plurilingues, de médiation, etc., au sein de l'évaluation. Cette marginalisation est d'autant plus forte que l'évaluation est largement présentée comme devant (étant censée) s'appuyer sur des traces, voire des observables pour être « objective », et que les dimensions langagières sont réputées être plus « objectivement observables » que les dimensions interculturelles, plurilingues, etc.
En outre, cette domination de la psychométrie, ou, en tout cas, d'une philosophie psychométrique, a abouti à la mise en œuvre généralisée (Huver, 2009a) de tests constitués de QCM étalonnés et standardisés (TCF, TEF, DIALANG, TOEFL, TOEIC, WIDAF, etc.).

[1] Ainsi, le DELF fait office de certification dans un certain nombre de systèmes éducatifs étrangers, mais aussi dans les sections bilingues ou encore aux enfants migrants nouvellement arrivés, de même que réciproquement, se manifeste la tendance à proposer aux apprenants français une certification externe de langue (la certification par le CLES au niveau du Master constituant ici une illustration d'actualité).

Ainsi, même s'il existe un certain nombre d'outils et de pratiques plus divers dans les objectifs qu'ils visent, les modalités qu'ils mettent en œuvre, les épistémologies qui les fondent et les histoires dont ils émanent (histoire de l'évaluation, histoire de la DLC, histoire de l'éducation, entre autres)[2], on ne peut que constater l'actuelle « uniformisation des pratiques évaluatives par sur-valorisation des certifications et des approches objectivantes et quantitatives de l'évaluation » (HUVER 2009a : 79).
Ainsi, la notion d'évaluation traverse désormais le champ social de manière massive, sans pour autant que les outils et les pratiques mis en œuvre aient, eux, changé. En d'autres termes, du point de vue des outils et des pratiques, si changement il y a, celui-ci est surtout, et fondamentalement, d'ordre technique (techniques d'échantillonnage, constitution de banques d'items, etc.), mais pas d'ordre méthodologique ni épistémologique. De ce point de vue, il me semble donc nécessaire d'interroger la notion même de changement, non seulement au regard des critères qui en fondent l'interprétation, mais aussi en termes de visée et d'objectif, notamment de formation.

1.2. Changer ou diversifier l'évaluation ?

Interroger la notion de changement en tant que visée de l'intervention didactique et de la recherche a des conséquences sur la manière dont nous – chercheurs et/ou formateurs et/ou praticiens – devons concevoir la formation mais aussi la recherche. En effet, les formations qui ont pour objectif « l'innovation pédagogique » ou l'amélioration des pratiques (cf. la part belle actuellement faite aux « bonnes pratiques ») véhiculent de fait l'idée que toute nouveauté est « meilleure » que l'existant et relèvent à ce titre de ce qu'on pourrait appeler une « idéologie de l'innovation ». A l'inverse, une formation accueillant la diversité des pratiques et visant à leur diversification amène les bénéficiaires de la formation à réfléchir sur leurs propres pratiques, à les confronter à d'autres pratiques ainsi qu'aux théories qui les fondent, pour choisir en conscience de les changer et/ou de les revendiquer. De même, la notion de diversification semble plus heuristique que celle d'innovation au regard de l'impératif actuel de « contextualisation », qui, par définition, ne pose pas certaines pratiques comme intrinsèquement meilleures que d'autres, mais comme plus ou moins pertinentes selon les contextes dans lesquels elles s'inscrivent. Enfin, étant donné le contexte actuel de domination des certifications et, plus largement, d'une conception objectiviste et positiviste de l'évaluation, travailler à la diversification des pratiques me semble être une stratégie à envisager pour remettre en question cette domination elle-même, par le biais de l'introduction de pratiques « autres » (plus formatives, plus centrées sur l'apprenant et ses pratiques d'auto-évaluation, plus inclusives de dimensions réputées « inévaluables »).
Du point de vue de la recherche, explorer la diversité des pratiques d'évaluation suppose premièrement d'analyser les processus de diversification ou, à l'inverse, d'uniformisation dans lesquels celles-ci s'inscrivent. Ainsi, l'uniformisation constatée des certifications en langues doit s'analyser non pas seulement sous un angle méthodologique, mais surtout d'un point de vue socio-historico-politique (inscription dans un cadre plus large de pilotage politique par l'évaluation) et épistémologique (domination actuelle du paradigme positiviste et objectiviste en évaluation). Ce type de démarche suppose également de travailler à une déconstruction des démarches d'évaluation actuellement dominantes, et notamment au fait que celles-ci se fondent sur la production de données comparables en fonction d'indicateurs présentés comme objectifs et qui restent peu questionnés, notamment du point de vue de la pertinence et de la définition des catégories mobilisées[3]. Enfin, il s'agit de travailler à la

[2] Je pense ici plus particulièrement à l'auto-évaluation et à l'évaluation formative, mises en avant dès le début des années 80 dans les travaux du CRAPEL sur l'auto-apprentissage.
[3] Pour un exemple d'analyse critique des évaluations menées dans le domaine de l'illettrisme, cf. Blum et Guérin-Pace (2000).

construction argumentée d'un paradigme « alternatif » légitimant et reconnaissant la pluralité et la subjectivité des pratiques d'évaluation – travail de plus longue haleine qui suppose d'une part, une remise en question forte de quelques allants de soi sur l'évaluation et, d'autre part, une intégration des représentations aux recherches dans ce domaine.

2. Intégrer les représentations

Les représentations liées à l'évaluation ont fait l'enjeu d'un certain nombre de travaux dans le domaine des sciences de l'éducation. Ainsi, VESLIN et VESLIN (1992) proposent au début de leur ouvrage une analyse de quelques discours d'enseignants afin de montrer que ceux-ci se concentrent essentiellement sur la note et la fonction sommative de l'évaluation. Dans une perspective similaire, ANTIBI (2003) met en évidence la persistance des mauvaises notes, phénomène nettement affirmé en France, qu'il nomme la « constante macabre », et qui selon lui, émane de représentations sociales très largement partagées de l'évaluation. Dans ces deux cas, l'objectif est essentiellement méthodologique et « prescriptif » (DE KETELE éds., 1986) : il s'agit en effet d'analyser les freins représentationnels à la diffusion de pratiques d'évaluation formative (Veslin et Veslin) et/ou de pratiques d'évaluation sommatives plus équitables (Antibi), pour légitimer des propositions de démarches – des propositions méthodologiques donc – allant dans ce sens.

D'autres travaux s'inscrivent quant à eux dans une démarche plus compréhensive, dont l'objectif n'est plus tant de proposer, voire de préconiser des démarches, que de mieux comprendre les liens entre représentations et pratiques. Ainsi, BARLOW (2003) s'intéresse à l'évaluation scolaire et à ses représentations sous un angle essentiellement symbolique, au moyen d'un corpus littéraire et cinématographique varié. JORRO (2000 : 51) pose, elle, la question des gestes évaluatifs qu'elle met en relation avec des imaginaires « qui témoignent de la mise en discours et de la circulation de théories mais aussi de valeurs, de concepts, de principes, de lois scientifiques, de mentalités, de discours et d'actes imbriqués et indissociables ».

En DLC, ce travail d'exploration des représentations de l'évaluation reste en grande partie à mettre en œuvre. En effet, peu de travaux ont jusqu'à présent abordé cette problématique : par exemple, BARBE (1990) ou VELTCHEFF (2009) ne font que mentionner les représentations, en invoquant leur intérêt pour la formation à l'évaluation (Barbé) ou la diffusion d'une perspective actionnelle (Veltcheff). Ainsi, bien que la notion de représentation soit maintenant largement diffusée dans le champ de la DLC, elle n'a pour l'instant pas réellement été mobilisée dans le domaine de l'évaluation. Le choix de ce volume d'intégrer explicitement la question des représentations aux recherches sur l'évaluation marque donc à cet égard un pas important, qui me semble particulièrement bienvenu et pertinent.

2.1. Complexifier l'analyse des représentations

Dans ce cadre, et comme le propose ce volume, il s'agirait tout d'abord de travailler à une complexification et à un affinement des analyses des représentations. Ainsi, Barbé souligne que les représentations de l'évaluation sont, en DLC, largement négatives, du fait de pratiques « à la fois trop technocratique[s] et trop subjective[s], donc doublement suspecte[s] » (BARBE, 1990 : 121). De même, hors du champ de la DLC, un certain nombre de travaux visent à dénoncer l'actuel poids de l'évaluation dans les entreprises et les institutions publiques, en la présentant uniquement sous l'angle de ses effets pervers et de son instrumentalisation dans un cadre économique néolibéral (prises de position de R. Gori, dernier numéro des *Cahiers internationaux de sociologie*, etc.). Or, si on s'attache à la manière dont les acteurs (enseignants notamment) s'approprient la question de l'évaluation, on se rend compte que leurs représentations sont effectivement négatives (l'évaluation étant alors opposée à l'apprentissage, vécue – par l'apprenant, mais aussi par l'enseignant – comme

anxiogène, inutile, injuste, arbitraire, etc.). Mais on s'aperçoit surtout que ces représentations sont beaucoup plus complexes et en tension, ces tensions étant, dans le cadre d l'enseignement/apprentissage des langues, construites autour de trois constellations principales :

- *Evaluation sommative vs formative* : Dans le cadre de divers entretiens que j'ai pu mener (HUVER 2010a, HUVER et SPRINGER 2010, HUVER et GOÏ à paraitre 2011), les témoins envisagent très majoritairement l'évaluation comme un bilan et, *conjointement*, comme une aide à l'apprentissage, donc comme un vecteur de formation, complexifiant en cela le clivage classiquement établi entre évaluation formative et évaluation sommative (SPRINGER 2001). De même, si on s'attache aux discours actuellement produits sur l'évaluation, la dénonciation quasiment unanime dont elle fait l'objet renvoie à des représentations essentiellement sommatives. Or, cette focalisation sur le sommatif est certainement, au moins en partie, contre-productive, mais surtout, elle ne rend pas compte de la manière dont certains acteurs s'approprient ces évaluations pour mettre en œuvre des « processus locaux diversifiés » dans le cadre de l'évaluation des politiques publiques d'éducation (DEMAILLY 2006) ou encore pour en faire un vecteur de développement personnel et/ou professionnel, pour eux-mêmes, pour leurs élèves ou pour leur institution (PAQUAY et al. 2010). Ceux-ci laissent paraître des représentations bien plus complexes et subtiles, qui expriment la crainte des conséquences psychologiques et/ou sociales parfois extrêmement graves[4] que ces évaluations peuvent occasionner, et, *en même temps*, le fait que ces évaluations (notamment ces auto-évaluations) leur ont permis de réfléchir à leurs pratiques, d'améliorer les dispositifs auxquels ils participent, etc. (HUVER et CADET 2010). ;
- *Objectivité vs subjectivité* : Du point de vue des représentations, la subjectivité de l'évaluation est massivement envisagée comme un écueil, une bonne évaluation étant censée être une évaluation objective et un bon évaluateur un évaluateur neutre. Ces représentations sociales sont très largement partagées et ont légitimé la mise en œuvre de différentes procédures visant à éliminer la subjectivité5. Toutefois, un certain nombre de chercheurs, très sceptiques face à cette lame de fond objectiviste, cherchent à réhabiliter le « jugement professionnel » (MERLE 1998 et 2007), qui, s'il est moins « objectivé » n'en est pas pour autant moins fiable. En DLC, l'enjeu est de taille au regard d'une part de la nécessité de (re)légitimer des pratiques « autres » que les pratiques certifiantes (auto-évaluation évaluation formative, etc.) et d'autre part des nouvelles finalités éducatives (approche par compétences, compétence interculturelle, plurilinguisme, formation tout au long de la vie, apprentissage collaboratif, etc.), qui renvoient à des processus complexes et des itinéraires singuliers, d'une certaine manière « inaccessibles », en tout cas par le biais d'évaluations externes standardisées (HUVER et SPRINGER 2011, ZARATE et GOHARD-RADENKOVIC 2004, etc.) ;
- *Norme vs usage* : Les représentations de l'évaluation restent très généralement centrées sur la faute, celle-ci devant être systématiquement relevée et corrigée. En DLC, elles sont plus particulièrement centrées sur les fautes de langue : les procédures d'évaluation soulignent en effet essentiellement les erreurs linguistiques (morphologie, syntaxe, orthographe, etc.). L'avènement de l'approche communicative aura permis de

[4] Ces réflexions en tensions sur les liens entre évaluation et pratiques (voire identité) professionnelles apparaissent notamment au détour et en filigrane de différents documentaires sur le travail, et, significativement, sur la souffrance au travail. La labellisation des centres de FLE en France ou l'évaluation des formations et des équipes de recherche par l'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur) en constituent d'autres exemples, plus proches de nos champs d'intervention.

[5] Corrections multiples, procédures psychométriques, etc. Dans le domaine de la DLC, la vague psychométrique actuelle est également représentative de ce phénomène.

légitimer également des critères d'ordre pragmatique (la réalisation des actes de parole est désormais communément admise dans les grilles des évaluations certificatives par exemple). Toutefois, les dimensions culturelles et interculturelles restent pour l'instant très peu prises en compte, de même que les phénomènes plurilingues, qui restent essentiellement envisagés comme des erreurs dans les productions de classe, alors qu'ils sont envisagés comme des stratégies de communication dans les usages (HUVER 2010b). Ainsi, intégrer ces dimensions, et notamment les stratégies plurilingues, aux référentiels d'évaluation, suppose de reconsidérer la notion d'erreur, et, partant, les représentations de ce qui est « correct », « juste », etc.

2.2. Analyser les liens réciproques entre représentations et pratiques

Par ailleurs, il s'agirait d'explorer les liens réciproques entre pratiques et représentations, de manière à mieux comprendre un certain nombre de décalages entre la recherche et les pratiques d'enseignement, ainsi qu'entre les pratiques d'enseignement et les pratiques d'évaluation. Ce type de recherche constitue un enjeu majeur dans une perspective de formation des enseignants, dès lors que l'on cherche à identifier et/ou à anticiper les nœuds de résistance du public en formation et à s'appuyer sur les représentations pour construire la formation elle-même. Du point de vue de la recherche, cette mise en lien permettrait de questionner – au sens d'interroger et de remettre en question – l'idée selon laquelle, pour reprendre les mots de Porcher, « qui tient l'évaluation tient tout le système ». En effet, l'évaluation (notamment sommative et certificative) est souvent pensée comme un levier de renouvellement des pratiques d'enseignement, du fait de la nécessaire mise en cohérence des objectifs d'enseignement et des objectifs d'évaluation. Dans le domaine du FLE, l'introduction du DELD/DALF dans les années quatre-vingts a par exemple relevé de cette perspective, dans la mesure où cette certification a été pensée (entre autres) comme un vecteur de diffusion des approches communicatives. L'hypothèse sous-jacente était que l'introduction, au niveau de l'évaluation sommative (i.e. de la certification), d'objectifs, d'activités et de critères de type « communicatif » entraine(rait), en amont, une mise en cohérence entre enseignement et évaluation par nécessité de garantie de la validité de l'évaluation, et, partant, une évolution des objectifs, des activités et des pratiques d'enseignement. Or, dès lors que l'on considère les représentations comme un des principes moteurs de l'action sociale, cette hypothèse somme toute assez déterministe ne peut plus tenir, en tout cas de manière aussi ferme. En effet, le format de l'évaluation (contenus, activités proposées, etc.) joue certes indéniablement sur les pratiques de classe en amont. Ainsi, le DELF a bien évidemment joué un rôle de diffusion de l'approche communicative en tant qu'approche identifiée, avec ses construits épistémologiques propres et des principes méthodologiques englobant l'ensemble de la chaine enseignement / apprentissage / évaluation. Cependant, les principes théoriques sous-tendant cette certification se sont entretissés avec des représentations très diverses – de l'évaluation, mais aussi de l'approche communicative ou de l'enseignement / apprentissage – pour produire, au final, des pratiques de classes très hétérogènes. De même dans le cadre du français langue seconde, la mise en œuvre des évaluations diagnostiques proposées aux enfants nouveaux arrivants est traversée par les représentations des enseignants, ce qui engendre des pratiques diverses, intégrant le répertoire verbal des enfants selon des modalités très variées, et se posant comme plus ou moins étayantes, fondées sur l'observation et/ou les résultats, souples/rigides, etc. (cf. à ce propos : HUVER et GOÏ 2010 et à paraitre 2011).

2.3. Evaluation et identité

Enfin, au-delà du jeu d'influences réciproques entre représentations et pratiques, le lien est également assez peu fait entre évaluation et identité – identité d'apprenant, identité

d'enseignant. Or, les élèves se perçoivent comme « bons en langue » parce qu'ils ont de bonnes notes en langue à l'école (bien plus que pour les usages qu'ils peuvent en faire en dehors de l'école, notamment par le biais des outils numériques) : l'évaluation constitue donc ici un vecteur de construction identitaire non négligeable (HUVER 2009b) et pourtant peu évoqué en DLC. Du côté des enseignants, les modalités de recrutement (donc d'évaluation de leur professionnalité) auxquels ceux-ci sont soumis jouent également vraisemblablement sur la construction de leur identité professionnelle (HUVER, 2010a). De ce point de vue, il serait par exemple intéressant de se demander dans quelle mesure les évaluations auxquelles sont soumis les enseignants (évaluations par les inspecteurs de l'éducation nationale, mais aussi procédures de recrutement) influent sur leurs représentations ainsi que sur leurs pratiques : dans quelle mesure les épreuves de production orale aux CAPES de langues, souvent très normatives et faisant peu de place à la variation, influent-elles sur la reconnaissance et la prise en compte de la variation dans les classes de langue ? Dans quelle mesure la récente suppression de cette épreuve joue-t-elle sur la place de l'oral dans les classes ? Ou encore, dans quelle mesure la certification désormais obligatoire du niveau de langue dans les masters, et notamment dans les masters « Métiers de l'enseignement », par le biais du CLES (Certificat de Compétences en Langues de l'Enseignement Supérieur) joue-t-elle sur la construction de la professionnalité des futurs enseignants (cette certification pouvant constituer un vecteur de formation professionnelle et/ou d'insécurisation linguistique) ?

Conclusion

La recherche en évaluation reste pour l'heure essentiellement pensée sous un angle méthodologique, dans la mesure où il s'agit d'élaborer des protocoles de construction d'outils, ainsi que des outils et des dispositifs d'évaluation, l'objectif de la recherche étant alors de discuter, voire de prouver, leur validité, leur objectivité, leur pertinence, notamment au regard d'un courant méthodologique particulier. Cependant, ces recherches envisagent pour la plupart la classe de langue, et les activités qui y sont proposées, en elles-mêmes et pour elles-mêmes. Par conséquent, elles ne font aucune place à la manière dont se configurent et se pluralisent les pratiques (d'enseignement et d'apprentissage), au regard d'articulations toujours contextualisées et historiquement situées entre des orientations politiques, des considérations sociales, des paradigmes épistémologiques, des préconisations institutionnelles et des parcours biographiques individuels (Huver et Goï à paraître 2012). Or, s'attacher à la pluralité des pratiques d'évaluation et aux représentations qui les sous-tendent, en replaçant ces configurations multiples dans un cadre sociétal, historique, politique, épistémologique plus large, suppose de penser des modalités de recherche qui permettent de travailler non seulement sur cette diversité, mais avec elle. De ce point de vue, il s'agit donc de construire des démarches de recherche qui permettent de prendre en compte et de comprendre cette pluralité, et, plus largement, l'irréductible altérité et les objectifs diversifiés des acteurs parties prenantes de la recherche (dont les chercheurs eux-mêmes).

Bien sûr, il serait tentant de constituer l'évaluation en domaine de recherche autonome au sein de la DLC, ce qui contribuerait en outre à la visibilité et à la légitimation des chercheurs qui s'intéressent à ce sujet. Toutefois, dès lors que l'on élargit le champ d'investigation en l'ouvrant à des dimensions non strictement méthodologiques, il me semble qu'on ne peut plus envisager l'évaluation comme un sujet de recherche spécifique : il s'agit au contraire de la considérer comme une entrée permettant de continuer à se poser, sous un angle différent, un certain nombre des questions qui se posent à la DLC dans son ensemble.

Bibliographie

ANTIBI A., 2003 : *La constante macabre*, Cahors, Math'adore.

BARBE G., 1990 : « Se former à l'évaluation ou l'évaluation pour se former ? », In : Porcher L. (Coord.), « L'évaluation en didactique des langues et des cultures », *Etudes de linguistique appliquée*, n°80, pp.119-131.
BARLOW M., 2003 : *L'évaluation scolaire, mythes et réalités*, Paris, ESF.
BLUM A. et GUERIN-PACE F., 2000 : *Des lettres et des chiffres – Des tests d'intelligence à l'évaluation du « savoir lire » : un siècle de polémiques*, Paris, Fayard.
Cahiers internationaux de sociologie, 2010 : « Ce qu'évaluer voudrait dire », n°128-129.
DE KETELE J.-M. (eds.), 1986 : *L'évaluation : approche descriptive ou prescriptive ?*, Bruxelles, De Boeck.
DEMAILLY L., 2006 : « En Europe : l'évaluation contre la crise des systèmes scolaires, l'évaluation en crise », *Education et sociétés*, n° 17, pp.105-120.
HUVER E., 2010a : « Quelles représentations de l'évaluation chez les enseignants de langues dans l'enseignement supérieur ? » In F. Dervin et E. Suomela-Salmi (Eds.) *New approaches to assessing language and (inter-)cultural competences in higher education / Nouvelles approches de l'évaluation des compétences langagières et (inter)culturelles dans l'enseignement supérieur*. Bern : Peter Lang, pp.81-96.
HUVER E., 2010b : « "J'ai pas pu corriger parce que c'est le mélange des langues" - L'évaluation de la compétence plurilingue entre résistances et tensions », *Cahiers de l'ACEDLE*, n°7, pp.113-137. http://acedle.org/spip.php?article3099.
HUVER E., 2009a : « Le CECR : au service de l'évaluation ? », *Education et sociétés plurilingues*, n°27, 71-82.
HUVER E., 2009b : « Evaluation des enfants nouvellement arrivés en France : évaluer des enfants non francophones, allophones et/ou plurilingues ? », *Cahiers de l'ASDIFLE*, n°20, 149-168.
HUVER E. et CADET L., 2010 : « La formation professionnelle des enseignants : réflexivité et évaluation sont-elles compatibles ? », in : Paquay L., Van Nieuwenhoven C. et Wouters P., *Evaluation et développement professionnel*, Bruxelles : De Boeck, pp.117-129.
HUVER E. et GOÏ C., 2010 : « Evaluer – accueillir – insérer : quel(s) prisme(s) pour quelle(s) projection(s) ? », *Cahiers de Sociolinguistique*, Rennes : PUR.
HUVER E. et GOÏ C., à paraitre 2011 : « Des enseignants aux chercheurs en didactique des langues. Postures épistémologiques entre identités, cultures et imaginaires professionnels. », Presses universitaires de Dijon**.**
HUVER E. et GOÏ C., à paraitre 2012, « FLE, FLS, FLM : Continuum ou interrelations construites, négociées, imaginées ? », in : Cadet L. et Guérin E. coord. « Enseigner le FLM/FLS/FLE : convergences et divergences », *Le français aujourd'hui*, n°176.
JORRO A. 2000 : *L'enseignant et l'évaluation. Des gestes évaluatifs en question*, Bruxelles, De Boeck.
MARTUCCELLI D. 2010 : « Critique de la philosophie de l'évaluation », *Cahiers internationaux de sociologie*, décembre 2010, pp.27-52.
MERLE P., 2007 : *Les Notes - Secrets de fabrication*, Paris, PUF.
MERLE P., 1998 : *Sociologie de l'évaluation scolaire*, Paris, PUF.
PAQUAY L., VAN NIEUWENHOVEN C. et WOUTERS P., 2010 : *Evaluation et développement professionnel*, Bruxelles : De Boeck.
SPRINGER C. 2001 : « Diagnostic, bilan de compétences, certification : les nouveaux habits de l'évaluation », *Langues modernes*, n°2, pp.48-60.
VELTCHEFF C., 2009 : « Evaluer par les tâches : une évaluation formative accessible aux enseignants et aux apprenants », *Le Français dans le monde recherches et applications*, pp.133-141.
VESLIN O. et VESLIN J., 1992 : *Corriger des copies : évaluer pour former*, Paris, Hachette.

ZARATE, G. et GOHARD-RADENKOVIC, A. (2004) *La reconnaissance des compétences interculturelles : de la grille à la carte*, Paris, Didier.